TRENDS AND CHALLENGES IN SMART METERING

M. Khazaei, University of Strathclyde, Glasgow, UK (Mohammad.khazaei@strath.ac.uk)
L. Stankovic, University of Strathclyde, Glasgow, UK (lina.stankovic@strath.ac.uk)
V. Stankovic, University of Strathclyde, Glasgow, UK (vladimir.stankovic@strath.ac.uk)

ABSTRACT

With strong policy support globally, it is expected that the total amount of smart energy meters installed worldwide will reach 780 million by 2020, including 200 million in the EU and 30 Million in the UK alone. Smart metering can improve grid operation and maintenance of distribution networks through load forecasting, improve demand response measures, and enhance end-user experience through accurate billing and appliance-level energy feedback via Non-Intrusive Load Monitoring (NILM). In this paper, we review trends of smart metering applications and challenges in large-scale adoption, and provide case studies to demonstrate application of NILM for meaningful energy feedback.

Keywords: *Smart meter, NILM, Clustering, Smart grid*

INTRODUCTION

Climate change, awareness of energy efficiency, new trends in electricity markets, and the changing role of consumers as prosumers for microgrid generation, are promoting the use of Renewable Energy Resources, Distributed Generation and Distributed Storage. Evolution towards an electricity grid model capable of managing numerous generation and storage devices in an efficient and decentralized manner determines the core of the Smart Grid (SG), enabled by communications across all components of the electrical grid. Smart meters provide the interface between the grid and the end-user (e.g., household), through communications between the meter and supplier either directly or via a neutral third party. The choice of communications technology will depend on ease of installation, accessibility, integration within the existing infrastructure, economic impact to the utility customer, legacy equipment, technical requirements, and functionality. Wireless communications can be an issue for remote area coverage, whereas wired Power Line Communication (PLC) are cost-effective for rural lines over long distance. However, PLCs suffer from long latencies for data transmission (compared to wireless), less bandwidth and higher cost in cities (Al-Waisi, Z., & Agyeman, M. O. 2018), (Uribe-Pérez et al., 2016). A mixed approach is sometimes adopted, e.g., in The Netherlands and Germany, 20% of meters use GPRS and 80% used PLC for communication (ICCS-NTUA &AF Mercados EMI, 2015).

Direct smart meter to collector communications enables large bandwidth and enhanced throughput but suffers from long distances and interface with distribution automation devices. Hence a mesh approach is adopted, which extends over three hierarchical tiers (Emmanuel & Rayudu, 2016): Wide Area Network (WAN), neighbourhood area network (NAN) and home area network (HAN). HAN provides connectivity within the home or business space. It consists of smart meter, interconnected with in-home display and possibly with gas meter, and other sensors. The main technologies used for HAN include wireless Personal Area Networks (PAN) IEEE802.15.4 (ZigBee), IEEE802.11 (WiFi), and wired PLC. The coverage is short (1-10m) and the data rates low, in the order of kbps. A group of smart meters are often connected to one data concentrator, within NAN, and then all data concentrators are connected to the control center through the base station. Having data concentrators, increases numbers of smart meters that can be connected to a base station. The communication between a control center and a base station is usually wired, such as fiber optic. The connection between a base station and data concentrators can be WiMAX/LTE, and that between a data concentrator and smart meters can be RF 900 MHz. WAN includes backbone/core communications for all types of distributed area networks that exist at different segments of the grid. The main standards include 2.5G (GPRS, EDGE)/3G(HSPA, EVDO)/4G(LTE, WiMax), wired DSL, PON, with long distance coverage (10-100km) and high rates 10-100Mbps. NAN collects all the energy usage data from various HANs to the backbone via its gateways. It is deployed within the distribution system for pricing messages, monitoring, and controlling power delivery. The main standards used for NAN include cellular networks (2G/3G/4G LTE,WiMax (IEEE 802.16)) and PLC, wired Ethernet, DOCSIS. The coverage is around 10m-10km, and rates from 10kps to 1Mbps. [Bian et al. 2019] highlights preferred communications technologies implemented for AMI deployments in various states in the U.S, namely fiber optics and WiMAX/LTE for the AMI backbone network, 900MHz RF mesh network for NANs compared to PLC and WiMAX/LTE and PLC, Zigbee and WiFi for HANs.

Driving factors for widespread deployment of smart meters revolve primarily around accurate billing, customer retention via enhanced customer experience and improved demand side management (DSM). With the end-user or customer in focus, smart metering has been shown, together with the provision of In-Home Displays (IHDs) providing real-time energy consumption feedback, to lead to reduction in energy demand. However, the potential of smart metering has in the past decade been demonstrated to go much beyond that of accurate billing and DSM. Through smart analytics of the meter readings, namely load disaggregation (referred to as Non-Intrusive Load Monitoring or NILM) providing appliance-level consumption feedback, further energy savings of up to 4.5% (Kelly, J. & Knottenbelt, W. 2016) can be achieved. Furthermore, at network level, smart metering has been shown to improve operations and maintenance of the grid, low voltage network load forecasting and infrastructure planning. Thus, smart meters are also beneficial to energy suppliers and Distribution Network Operators (DNOs) to maintain and improve network operation and efficiently manage electricity generation.

Thus, the paper is organized as follows. We first provide a detailed review of emerging applications of smart metering, including NILM, briefly reviewing the state-of-the-art for each. This is followed by a case study where we describe the roll out of NILM via a chat bot to evaluate energy recommendations at scale in residential and commercial settings. Finally, we implement and compare four NILM algorithms (Artificial Neural Networks, Decision Tree, K-means and DBSCAN) in order to assess their effectiveness at identifying common appliances such as the oven and kettle from the aggregate smart meter data.

SMART METERING APPLICATION TRENDS

In this section, we review smart metering applications in two categories: from the DNOs perspective regarding network operation, energy provider focusing on technical and non-technical losses and the end-user focusing on NILM.

Network Topology and Maintenance: Using measurements from smart meters we can extract network topology to identify loading and voltage profiles, connectivity issues, and distributed energy resources (DER) impact measurement, that is helping High Voltage (HV) network, also smart meters data can help forecast future needs and trends of demand management, and maintenance requirements, for example, calculate working duration of transformers (Alahakoon, D., & Yu, X. 2015).

Detecting Voltage Deviation: Analysis of smart meter events helps reinforcement of low-voltage (LV) network operation by the detection of voltage deviations or even outage, as shown in (Prado et al., 2017) where a set of strategies have been undertaken to manage big data measurements: (1) event gathering, filtering and sorting processes; (2) isolate event types that are strongly related to Quality of service (QoS) (overvoltage, under voltage and loss of neutral); (3) measurements from distribution transformer in each substation are taken into account to complement event analysis, to improve the detection of hot spots in the LV network.

Network Line Outage and Fault Detection: Analyzing in real-time multivariate smart meter data is used to predict an outage before it occurs and to detect/isolate the location of each outage and its severity after it occurs (Moghaddass & Wang. 2017). It was concluded that defining pre-outage status as an anomaly level and then efficiently predict it using smart meter data can help the utility predict outages before they occur, which results in better planning and preparation for maintenance and avoiding costly outages. Defining post-outage behavior as another level of anomaly will also help the utility better perform post-outage activities including finding the location (isolation) and severity of the outage and faster and more efficient restoration. Descriptive analytics' results from historical data for different types of cause-specific outages verified that there is no fully deterministic and consistent relationship between each smart meter measurement and outage events, thus more advanced analytics frameworks are needed to integrate and analyze multi-dimensional and multi-source smart meter count data to generate useful insights and decision-making intelligence while accounting for critical factors, such as missing points.

Technical and Non-Technical losses: Energy theft detection (Wang et al. 2018) can be implemented using smart meter data and power system state data, such as node voltages. The energy theft detection methods with only smart meter data can be based on supervised learning (if labelled training data are available) and unsupervised learning (in the absence of training data). Supervised methods rely on training the classifier using known, labelled data, to discriminate between customers' normal and malicious consumption patterns, e.g., (Jokar et al. 2015) first use distribution transformer meters to identify areas with a high probability of energy theft, and then applies anomaly detection to identify suspicious customers. (Jindal et al. 2016) use a top-down scheme based on decision tree and

SVM. The decision tree is used to estimate the expected electricity consumption based on the number of appliances, occupants, and outdoor temperature. Then, the output of the decision tree is fed to the SVM classifier to classify the consumer as normal or malicious. Since obtaining the labeled dataset for supervised energy theft detection can be difficult and expensive, unsupervised energy theft detection methods are more desirable. (Depuru et al. 2013) developed an optimum-path forest clustering algorithm, where each cluster is modelled using Gaussian distribution, after which the load profile is identified as an anomaly if its distance to the cluster head is greater than a threshold. The method shows competitive performance to some popular methods, including k-means, Birch, affinity propagation, and Gaussian mixture model. In (Smart Meters and Losses, 2016), BC Hydro indicated in 2013 that, before introducing smart meters, electricity theft amounted to about 7% of residential load in their networks. With 1.9 million customers, they estimated that they could increase revenue by $802 million per annum due to smart meter information assisting in their theft detection efforts. To mitigate non-technical losses, (Buzau et al., 2018) propose a methodology which uses the smart meter data and auxiliary databases to formulate various characteristics of a customer's consumption behavior and geography. These characteristics are afterwards introduced into supervised machine learning algorithm (XGBoost) for model selection and evaluation for non-technical losses.

Load forecasting: Algorithms at MV/LV level need to handle high consumption variability, captured via smart meters that provide more granular time-interval data opening up possibilities of trend and cycle analysis and different time of day consumption analysis (Hong et al., 2014). Time-interval-based consumption also enables consumer behavior profiling and potentially relate consumption to external data such as weather, geography, and consumer information. The main limitation of short-term forecasting models is lack of inclusion of consumer consumption profiling. End-users can be categorized using available smart meter data into groups that show significant differences in demand profiles including peak seasonal demand. Customer categorization has been studied extensively and is performed using machine learning, namely feature extraction followed by clustering and classification. Depending on the data resolution and quantity available, various methods can be performed for feature extraction and classification, including Fast Fourier Transform (FFT), followed by SVM, decision trees, or (deep) neural networks. For example, short-term forecasting with hourly load data from a Belgian grid substation is reported in (Espinoza et al. 2005) highlight that forecasting and customer profiling are interrelated and proposing a unified framework which incorporates both. The initial modelling is based on seasonal time-series analysis, using the periodic auto-regression model. The stationary properties obtained from these models are run through K-means clustering to capture different customer profiles.

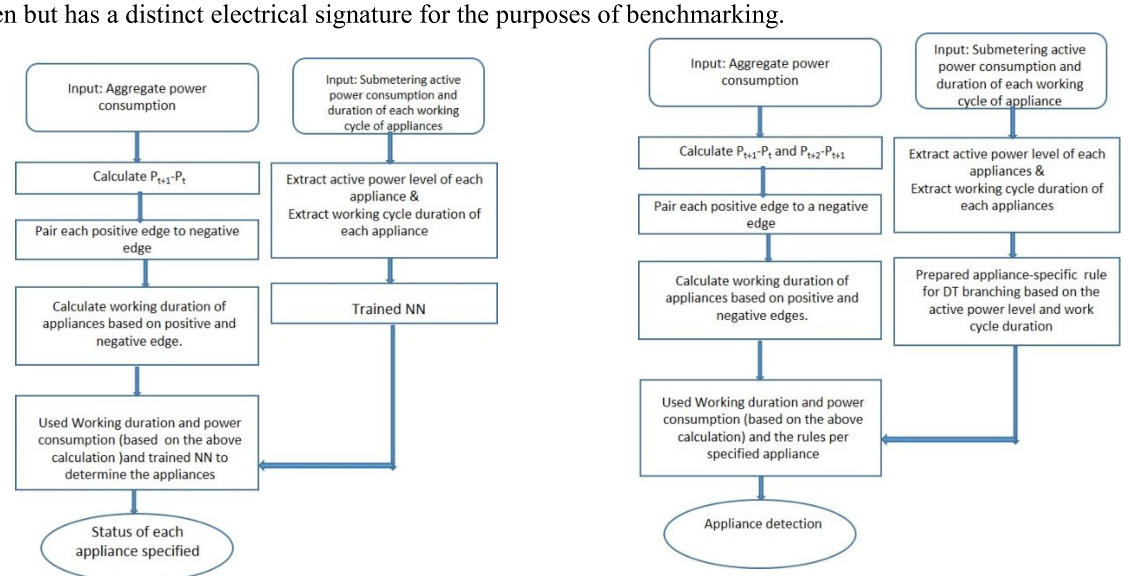
Load Disaggregation: Monitoring energy consumption of individual appliances using individual appliance sensors in a house or a commercial building is often impractical and expensive, especially since the number of electrical devices at home/ commercial buildings is rapidly increasing. On the other hand, energy disaggregation via NILM offers a non-intrusive, purely computational, software-based approach to separate aggregate load obtained from a single electricity meter into individual appliance loads. While research on NILM has primarily revolved on high frequency load measurements, the lower rates provided by smart metering initiatives (1 sec to 1 hour) are driving research into low-rate NILM methods. Low-rate NILM is particularly challenging due to noise from unknown appliances, signal transients that act as noise, load fluctuations, and the fact that the average household owns over 40 appliances. NILM methods can be grouped into supervised and unsupervised techniques (though hybrid, semi-supervised approaches are also possible). Supervised NILM techniques require a labelled dataset of appliance consumption data for training and are commonly based on low-complexity event detection, i.e., appliance switching on/off, then feature extraction from these events followed by classification or pattern matching to match events to pre-defined categories, each corresponding to one appliance, that are learned during training. Different supervised approaches based on Graph Signal Processing (GSP) and Decision Tree (Zhao et al., 2018), Support Vector Machines with K-means (Altrabalsi et al., 2016) and Deep Neural Networks (Murray et al., 2019) have been proposed. Unsupervised approaches showing good results with no requirement of training have been based on Dynamic Time Warping (Elafoudi et al, 2014) and unsupervised GSP (Zhao et al. 2016).

CASE STUDY: LOAD DISAGGREGATION AND CHATBOT FOR ENERGY FEEDBACK

The scalability and effectiveness of load disaggregation from low-resolution smart meter measurements for the purpose of energy feedback are currently being developed and tested on residential and commercial pilots across multiple countries in Europe within the Eco-Bot project (Eco-Bot). Eco-Bot aims to combine recent advances in non-intrusive appliance load monitoring techniques and chat-bot technology in order to engage residential and commercial energy consumers, with the goal of raising awareness in their behaviour in regard to energy efficiency. This personalized "eco-bot" can, among other things, deliver for the first time, information on itemized (i.e., appliance-level) energy usage, achieving what is called the "holy grail of energy efficiency". Furthermore, Eco-Bot proposed solution aims to achieve a measurably higher level of engagement with consumers than previous efforts by utilities, governments, appliances, property managers, Non-Governmental Organization (NGO) etc. (including serious games, gamification apps, competitions or other current interactive ICT tools already being deployed in the energy field, as well as other domains), by adding a more engaging form of interaction within existing platforms that has been proven in different market settings. In Eco-Bot, NILM will be performed on residential datasets with sampling resolutions of 10 seconds, 1 minute and 1 hour. This is the first time NILM will be tested at scale, on unseen datasets, and for a range of sampling rates.

NON-INTRUSIVE LOAD MONITORING (NILM) ALGORITHM

To illustrate operation and performance of NILM, we design and evaluate two supervised, Artificial Neural Network (ANN) and Decision Tree (DT), and two unsupervised, DBScan and K-Means, methods to disaggregate two appliances: Kettle and Oven. The methodology is shown as flowcharts for ANN, DT, K-means and DBScan in Figures 1(a), 1(b), 2(a) and 2(b), respectively. We are primarily motivated to disaggregate the oven in this study because it has not been tackled in depth in the NILM literature, compared to other appliances such as refrigeration, washing machines, dishwashers, toasters, microwaves, kettles etc., while it is one of the large home consumers. This is because ovens are rarely submetered, not included in most public electrical measurements datasets and hence cannot provide ground truth to the research community. We also included the kettle in our study, which also includes a heating element like the oven but has a distinct electrical signature for the purposes of benchmarking.

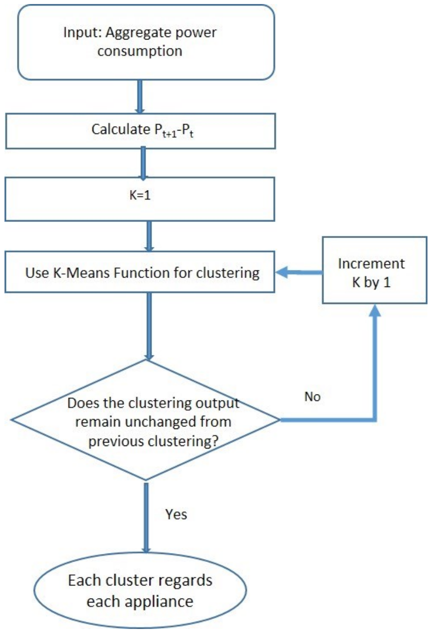


(a) (b)

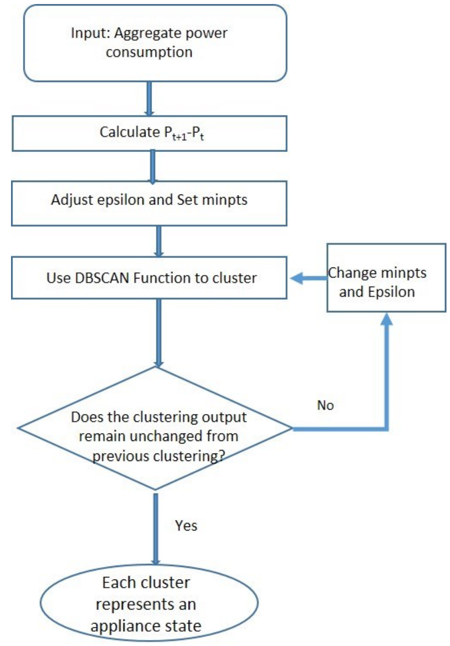
Figure 1: Flow chart of proposed supervised algorithms, each showing testing on the left and training on the right for: a) ANN, b) DT. P_t denotes a power measurement at time instance *t*.

We demonstrate the performance of the proposed algorithms using three datasets containing aggregate and sub-metering active power: REFIT H17 (Murray et al., 2017), REDD H1 (Kolter & Johnson, 2011) and a house in Norway (NOR), all for a month's duration. Sampling rate for aggregate power are 10 secs, 1 sec, and 1 minute, for REFIT, REDD and NOR, respectively. For the supervised algorithms, we use 80% of the sub-metering data to train the model and perform testing of the remaining 20% of labelled aggregate data. Features used were active power and duration of each ON-state. There are multiple states in an oven signature due to the thermostat. There were 28 and 5 oven activations in NOR and REDD datasets, and 17 and 261 kettle activations for NOR and REFIT datasets, respectively.

Results are shown in Table 1. For training ANN, we use Levenberg-Marquardt backpropagation with 25 hidden layers to train the kettle model, and for the oven model, the conjugate gradient backpropagation with 50 hidden layers provides the best results. With the K-means unsupervised algorithms, we observed that best results were obtained with k=10, indicating 10 distinct clusters. Input is the change in power level, i.e., $\Delta P=P_{t+1}-P_t$. With DBSCAN, using the same ΔP feature for clustering, we obtained best results with Epsilon=10 and Minpts=3.



(a)



(b)

Figure 2: Flow chart of proposed unsupervised algorithms: a) K-Means, b) DBSCAN, where Epsilon: A scalar value for Epsilon-neighborhood threshold and Minpts: A scalar value for minimum points in Epsilon-neighborhood that holds the core-point condition. P_t denotes a power measurement at time instance t.

Table 1: F1-score accuracy in % for NILM of 2 appliances for 2 supervised and 2 unsupervised ML algorithms

Algorithms		Kettle		Oven	
		REFIT	NOR	NOR	REDD
Supervised	ANN	73	**84**	45	57
	DT	**92**	61	57	**67**
Unsupervised	K-Means	76	51	79	66
	DBSCAN	**92**	46	**85**	60

The best performing disaggregation algorithm for the oven is the unsupervised DBSCAN for 1 min sampled NOR dataset at 85% accuracy. Best results for disaggregation of kettle are observed with supervised ANN for NOR dataset and DT for REFIT datasets. Supervised algorithms, ANN and DT, disaggregate the kettle better than the oven in general since the kettle has a very distinct always-on profile for all activations, unlike the oven whose electrical signature does not have a distinct profile, i.e., the variations among the many activations of the oven for active power and duration vary too much due to opening/closing oven variability. DT performs better with the relatively higher sampling rate datasets (REFIT and REDD) than the NOR dataset. Results obtained are in line with previous studies, with 65% accuracy for the Oven for the REDD dataset obtained with unsupervised GSP in (Zhao et al., 2016) and 94% obtained with DT for kettle in the REFIT dataset in (Zhao et al., 2018).

CONCLUSIONS

The value of smart metering analytics has been demonstrated to have tremendous potential for both consumer and network operators and can reduce costs of supplier and DNO and end users. NILM in particular helps end-users effectively manage their energy consumption and reduce their energy bills through appliance-level feedback. The accuracy of different NILM algorithms varies tremendously and has not been tested at scale yet, but a case study highlighting the Eco-Bot project will enable this with sampling rates varying from a few seconds to hourly resolutions. No single algorithm has been reported in the literature, which can disaggregate the whole range of appliances typically found in residential buildings. We demonstrate results for the oven, since that is a challenging appliance to obtain in public datasets and to disaggregate, and report our results. We show that unsupervised NILM approaches perform better than supervised NILM approaches for this appliance due to its variability in unique electrical signatures.

ACKNOWLEDGEMENTS

This project was partly supported by the European Commission under the ``H2020-EU.3.3.1. - Reducing energy consumption and carbon footprint by smart and sustainable use" program topic, according to the Grant agreement No. 767625. Furthermore, the authors would like to thank Oss Norge AS, Norway, for provision of smart metering data that was used for our results.

REFERENCES

Alahakoon, D., & Yu, X. (2015). Smart electricity meter data intelligence for future energy systems: A survey. IEEE Transactions on Industrial Informatics, 12(1), 425-436.
Altrabalsi, H., Stankovic, V., Liao, J., & Stankovic, L. (2016). Low-complexity energy disaggregation using appliance load modelling. *AIMS Energy*, *4*(1), 884-905.
Al-Waisi, Z., & Agyeman, M. O. (2018, September). On the Challenges and Opportunities of Smart Meters in Smart Homes and Smart Grids. In Proc. 2nd Int. Symposium on Computer Science and Intelligent Control (p. 16). ACM.
Bian, D., Kuzlu, M., Pipattanasomporn, M., Rahman, S., & Shi, D. (2019). Performance evaluation of communication technologies and network structure for smart grid applications. *IET Commun.*, *13*(8), 1025-1033.
Buzau, M. M., Tejedor-Aguilera, J., Cruz-Romero, P., & Gómez-Expósito, A. (2018). Detection of non-technical losses using smart meter data and supervised learning. IEEE Transactions on Smart Grid, 10(3), 2661-2670
Depuru, S. S. S. R., Wang, L., Devabhaktuni, V., & Green, R. C. (2013). High performance computing for detection of electricity theft. International Journal of Electrical Power & Energy Systems, 47, 21-30.
Eco-Bot: http://eco-bot.eu/
Elafoudi, G., Stankovic, L., & Stankovic, V. (2014). Power disaggregation of domestic smart meter readings using dynamic time warping. In *6th Int. Symp. Communications, Control and Signal Processing (ISCCSP), 2014* IEEE.
Emmanuel, M., & Rayudu, R. (2016). Communication technologies for smart grid applications: A survey. *Journal of Network and Computer Applications*, *74*, 133-148.
Espinoza, M., Joye, C., Belmans, R., & De Moor, B. (2005). Short-term load forecasting, profile identification, and customer segmentation: A methodology based on periodic time series. *IEEE Trans. Power Systems*, *20*, 1622-1630.
Hong, T., Pinson, P., & Fan, S. (2014). Global energy forecasting competition 2012.
Jindal, A., Dua, A., Kaur, K., Singh, M., Kumar, N., & Mishra, S. (2016). Decision tree and SVM-based data analytics for theft detection in smart grid. *IEEE Transactions on Industrial Informatics*, *12*(3), 1005-1016.
Jokar, P., Arianpoo, N., & Leung, V. C. (2015). Electricity theft detection in AMI using customers' consumption patterns. *IEEE Transactions on Smart Grid*, 7(1), 216-226
Kolter, J. Z., & Johnson, M. J. (2011). REDD: A public data set for energy disaggregation research. In Workshop on Data Mining Applications in Sustainability (SIGKDD), San Diego, CA (Vol. 25, No. Citeseer, pp. 59-62).
Lipošćak, Z., & Bošković, M. (2013, July). Survey of smart metering communication technologies. In *IEEE Eurocon 2013* (pp. 1391-1400)
Moghaddass, R., & Wang, J. (2017). A hierarchical framework for smart grid anomaly detection using large-scale smart meter data. IEEE Transactions on Smart Grid, 9(6), 5820-5830
Murray, D., Stankovic, L., Stankovic, V., Lulic, S., Sladojevic, S. (2019). Transferability of neural network approaches for low-rate energy disaggregation. *ICASSP IEEE Int. Conf. Acs., Speech & Sig. Proc.* pp. 8330-8334
Murray, D., Stankovic, L., & Stankovic, V. (2017). An electrical load measurements dataset of United Kingdom households from a two-year longitudinal study. *Scientific Data*, *4*, 1-12. https://doi.org/10.1038/sdata.2016.122
Prado, J. G., González, A., & Riaño, S. (2017). Adopting smart meter events as key data for low-voltage network operation. CIRED-Open Access Proceedings Journal, 2017(1), 924-928
Smart Meters and Losses: Best Practice Review. UK Power Networks (Operations) Limited, 2016.
Uribe-Pérez, N., Hernández, L., De la Vega, D., & Angulo, I. (2016). State of the art and trends review of smart metering in electricity grids. *Applied Sciences*, 6(3), 68.
Wang, Y., Chen, Q., Hong, T., & Kang, C. (2018). Review of smart meter data analytics: Applications, methodologies, and challenges. *IEEE Transactions on Smart Grid*, 10(3), 3125-3148.
ICCS-NTUA &AF Mercados EMI, 2015, Study on cost benefit analysis of Smart Metering Systems, FINAL REPORT in EU Member states
Zhao, B., Stankovic, L., & Stankovic, V. (2016). On a training-less solution for non-intrusive appliance load monitoring using graph signal processing. *IEEE Access*, 4, 1784-1799
Zhao, B., He, K., Stankovic, L., & Stankovic, V. (2018). Improving event-based non-intrusive load monitoring using graph signal processing. IEEE Access, 1-15. https://doi.org/10.1109/ACCESS.2018.2871343