Supported in part by European Commission under Horizon2020 MSCA-ITN-2020 Innovative Training Networks programme, Grant Agreement No 955422.

Trustworthy NILM in context of Demand Response and Sustainability

by Dr Lina Stankovic

Workshop: NILM for demand response – solutions for the energy crisis

11th May 2023

DEPARTMENT OF ELECTRONIC & ELECTRICAL ENGINEERING

University of Strathclyde Glasgow

THE QUEEN'S ANNIVERSARY PRIZES 2019 & 2021

For Higher and Further Education

UNIVERSITY OF THE YEAR 2012 & 2019

Times Higher Education

SCOTTISH UNIVERSITY OF THE YEAR 2020

The Times & The Sunday Times

University of Strathclyde Engineering

Research excellence in energy

In 2019, Strathclyde was selected as the winner for its excellence in energy innovation.

The University is a long-standing leader in energy research. We have more than 250 researchers working at any given time on energy systems innovations and over 300 PhD graduates over the past decade with expertise in the field.

The University also has a roster of spinout companies making valuable contributions to the enhancement of energy provision and we play a prominent role in informing energy policy in Scotland, the UK and beyond.

Energy crisis & Demand Response

- Recent rise of energy bills, by 67% in the UK between Feb'22-Feb'23
- 53% of adults in GB reducing energy consumption over winter, GECKO field study users not using their EV
- Energy transformation to meet NetZero targets: growth in renewable electricity generation
- Peak loading not aligned with peak renewable generation, renewable generation curtailed due to low demand and transmission capacity constraints limit generation capacity to be exported over larger distances
- Trials by National Grid Electricity System Operator in 2022 have shown it is domestic flexibility can be delivered at national level with small financial incentives,, e.g., variable 30-minute wholesale pricing, to maintain system power balance
- For smaller households with <1MW, smart domestic load shifting can be run 24/7
- Requires human interaction to adopt more sustainable consumption habits
- Products emerging is to remove dependence on end user and maximise convenience via an automated scheduler, e.g. running high loads such as EV charging at lowest import tariffs or times of high renewable generation….but still assume user input to program scheduler

Personalised ICT-tools for the Active Engagement of Consumers Towards Sustainable Energy

University of Strathclyde Engineering

- Piloting behavioural model and NILM methods to break down energy consumption to the appliance level and providing tailored recommendations on energy behaviour via a chatbot interface
- The technology was validated through large-scale residential and commercial pilots across Europe.
- "*Despite the fact that the COVID-19 pandemic affected the users' energy consumption in all pilots, results showed that Eco-Bot motivated both residential and commercial users to improve their energy consumption behaviour by implementing the provided recommendations and investing in more energy-efficient appliances*"
- Featured in Results in Brief

Taking away control from the user

University of Strathclyde Engineering

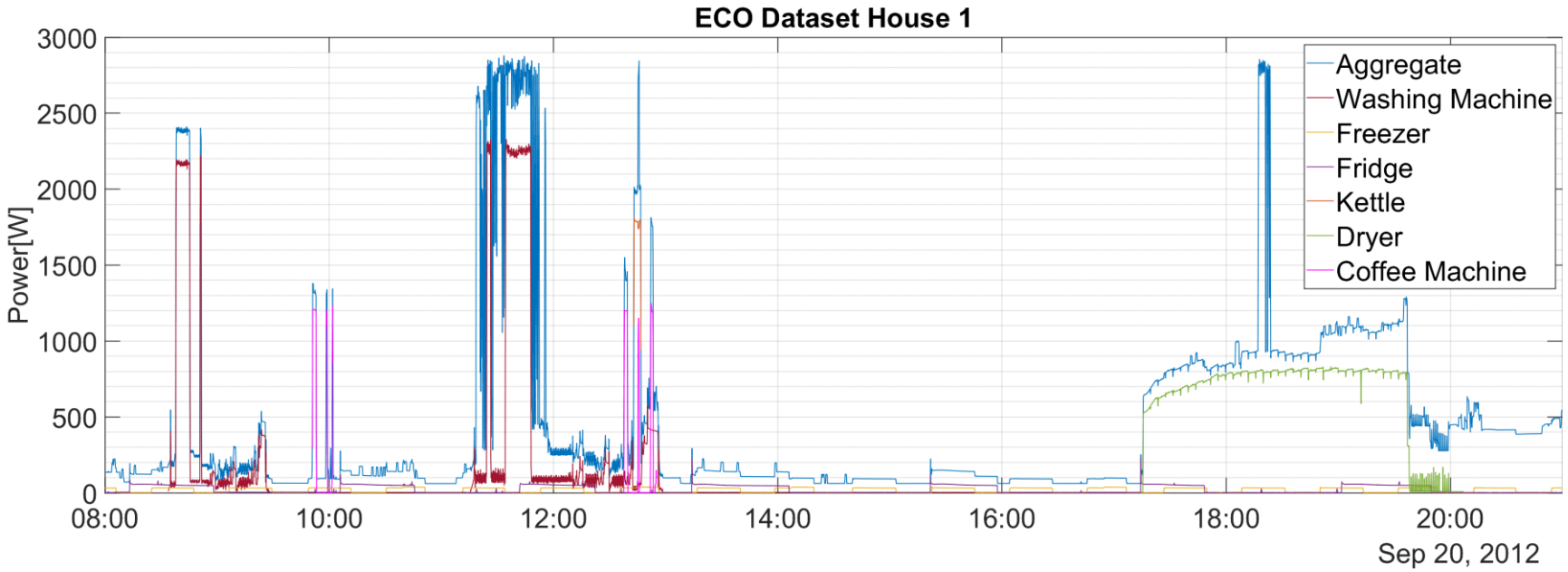
"*By introducing new ways of automatically and remotely controlling domestic environments smart technologies have the potential to significantly improve domestic energy management. It is argued that they will simplify users lives by allowing them to delegate aspects of decision-making and control - relating to energy management, security, leisure and entertainment etc. - to automated smart home systems. Whilst such technologically-optimistic visions are seductive to many, less research attention has so far been paid to how users interact with and make use of the advanced control functionality that smart homes provide within already complex everyday lives. What literature there is on domestic technology use and control, shows* ***that control is a complex and contested concept****. Far from merely controlling appliances, householders are also concerned about a wide range of broader understandings of control relating, for example, to control over security, independence, hectic schedules and even over other household members such as through parenting or care relationships*"

Hargreaves, T., Hauxwell-Baldwin, R., Coleman, M., Wilson, C., Stankovic, L., Stankovic, V., Liao, J., Murray, D., Kane, T., Hassan, T., & Firth, S. (2015). *Smart homes, control and energy management: how do smart home technologies influence control over energy use and domestic life?*. Paper presented at ECEEE-2015, Toulon/Hyères, United Kingdom.

Human-in-the-loop vision

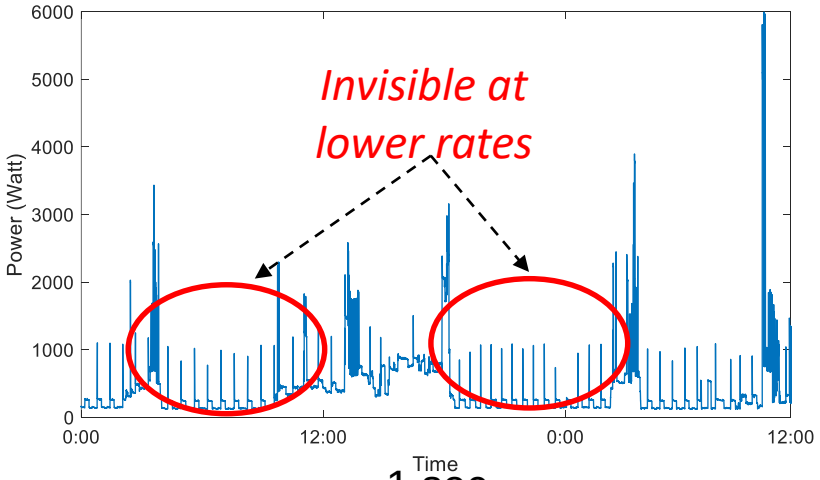
- Integrated, transparent system showing (for a household) live generation/export, import, import/export tariffs, consumption (at appliance level) and transparent scheduling
- NILM can bridge the gap between the DR schemes and automated schedulers that aim to minimise end-user input and user-led behavioural changes in energy consumption
- NILM learns a user's patterns of use, or energy consuming practices, and can inform a user-specific schedule for DR
- But…assumes NILM models are accurate across households and reliability of NILM predictions

What is NILM?

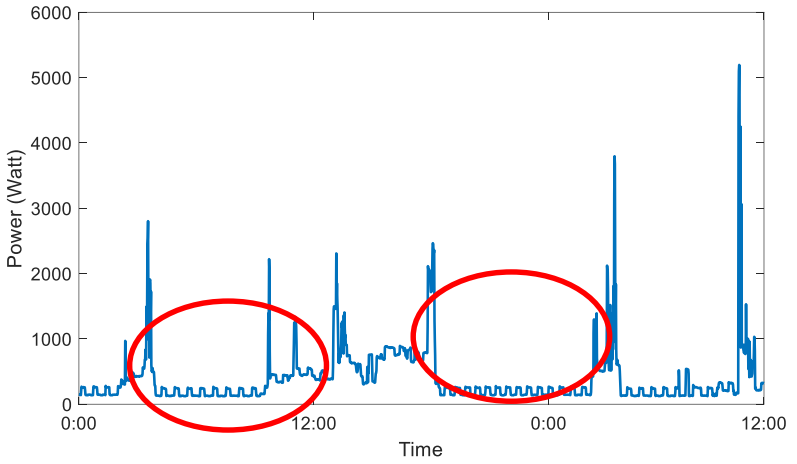


- *Detecting individual appliance usage and consumption from a building's total aggregate, smart meter reading*

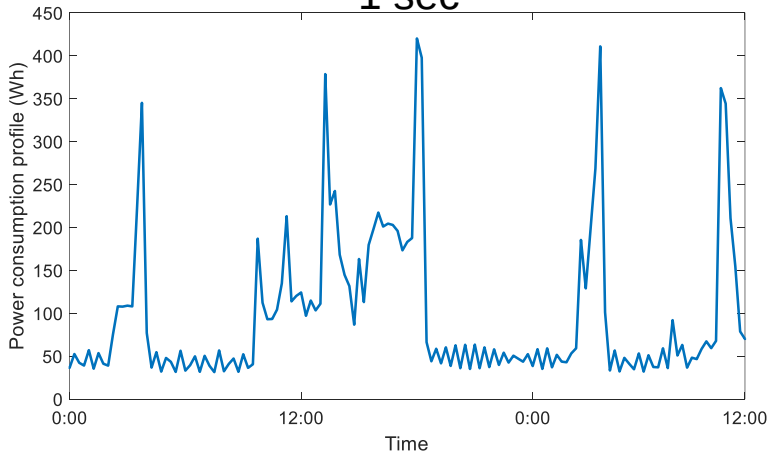
Importance of sampling resolution



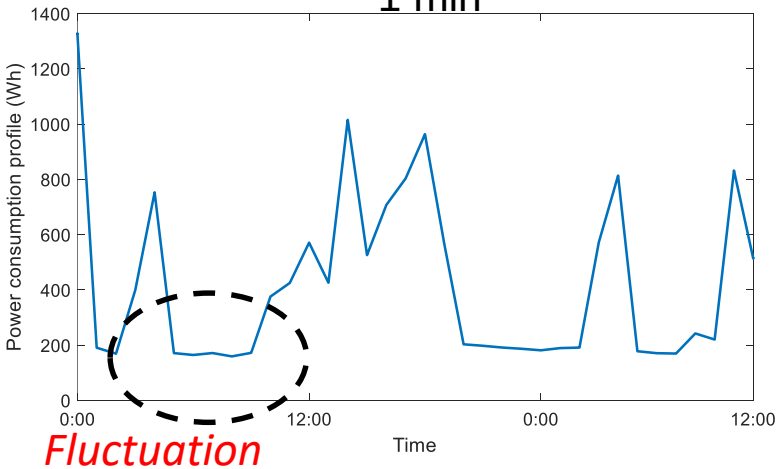
1 sec



1 min



15 min



Fluctuation completely gone

1 hour

Signal processing & AI challenges

- Power & SNR varies considerably, i.e., most loads fluctuat… …e mean value by typically 1-5%
- Sensor noise is present and many appliances ar…
- Different appliance settings and/or operatio… …ifferent load signatures
- Many appliances are '*Multi-state appli*… …g., washing machine, dishwasher)
- Some appliances are always on (… …ier, heating) – but load signatures are not periodic
- Only steady-state signa… …used at low sampling rates -> transients act as noise
- Average househo… …e owns over 40 electric appliances (and this number is increasing)!
- Many ap… …iten run simultaneously
- Huge … wattage: ~10W (phone charges) to ~5kW (heaters)
- There are more and more appliances with similar power profiles (e.g., stove and iron)

Dynamic and uncertain electrical measurements

Progress in tackling these challenges

- 40 years since the concept was introduced, with SP and ML approaches, at high frequencies and many load features
- Recent emergence of smart buildings and nationwide roll-out of smart meters and growing availability of public datasets
- Many supervised and unsupervised approaches for solving both classification and regression problems in detecting individual appliance usage and their energy consumption

https://xkcd.com/1838/

Common mistakes in literature:

- selectively choose a subset of "good" data from a dataset that is less noisy or gives best results
- only publish results of one experiment, or average results or even worse the best set of results
- selectively report metrics that show "good" performance of experiments

Responsible & sustainable NILM

University of Strathclyde Engineering

- Lots of code, even more papers out there with different methodologies, do we need to invent more?
- Same for data – lots of datasets publicly available. Do we need another dataset if there is no complementarity?
- Can we re-use? If so, ensure repeatability and reproducibility.

➢ Interpretability of the models: how does the algorithm come up with results, explaining the internal functions

➢ Explainability of outcomes from choice of dataset to end results

➢ Meaningful performance metrics

Barriers to NILM adoption

University of Strathclyde Engineering

- What is missing to ensure reliability of NILM:
  - Reliable supervised NILM models that are **transferable** to ‘unseen’ datasets or reliable **unsupervised** NILM methods that can operate on any dataset
  - Reliable NILM methods that focus on accurate disaggregation of **challenging loads**;
  - NILM feasibility for **non-residential sectors**, esp. hard to decarbonize sectors
  - **Fair and explainable metrics** for the evaluation of different NILM algorithms;
  - **Interpretable and explainable** algorithms for NILM;
  - Practical NILM deployments (scalability, transferability, privacy preserving) and **large-scale trials**
  - Novel datasets, data models, and toolkits for NILM research especially for **emerging appliances** such as EVs, heat pumps
  - **Applications** leveraging on NILM disaggregated data (e.g., flexibility estimation, life cycle analysis);

Agriculture, still far from decarbonization
Sustainable dairy farming – How can NILM help?

University of Strathclyde Engineering

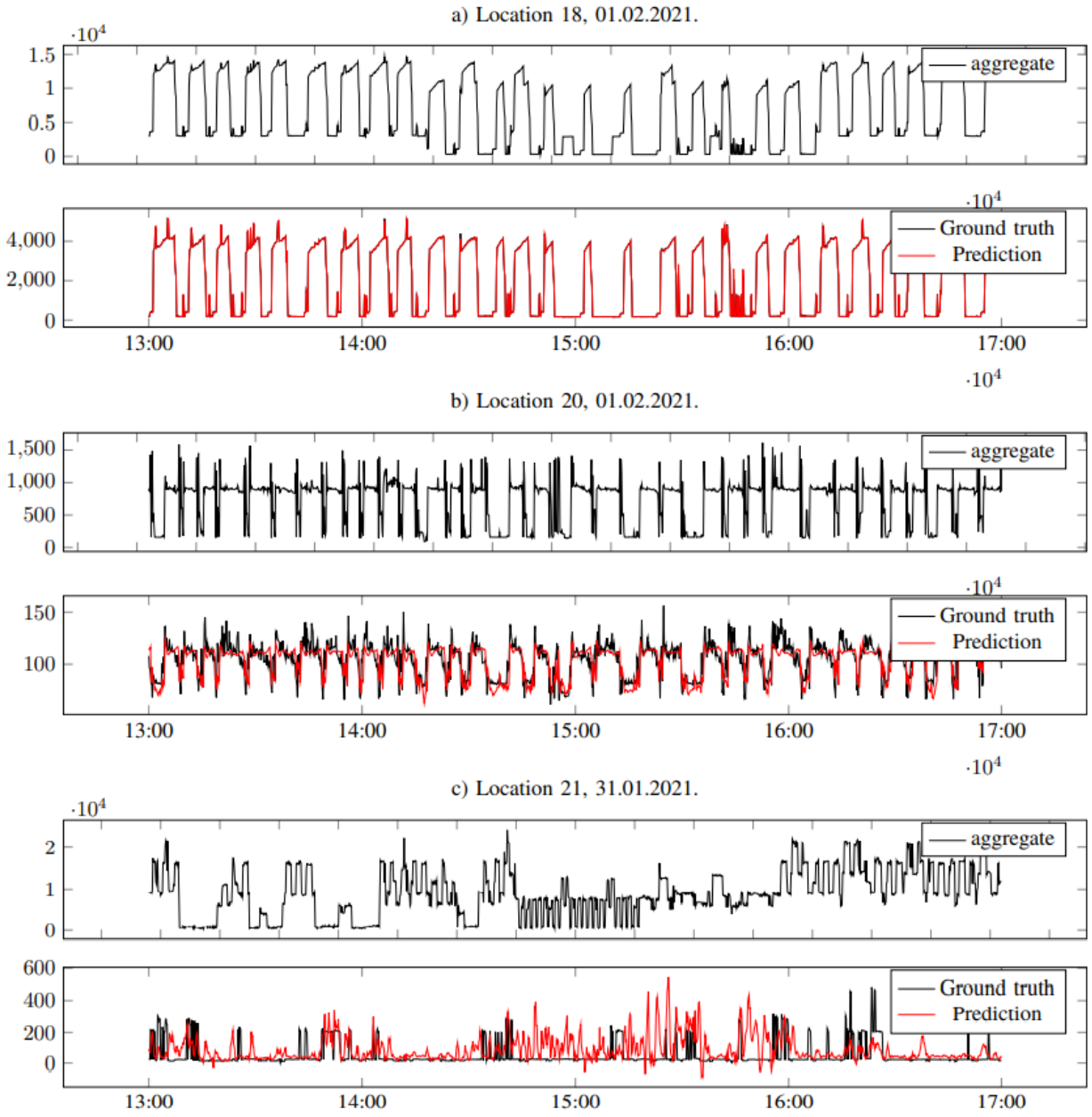
- Dairy industry is a large contributor to ***GHG emission***
- Although methane is the main problem, electricity consumption cannot be neglected due to various machinery that use ***electricity*** / fossil fuels
- 20 cows consume approximately 2000 kWh per year, equal to the consumption of an average mid-terrace UK house/flat

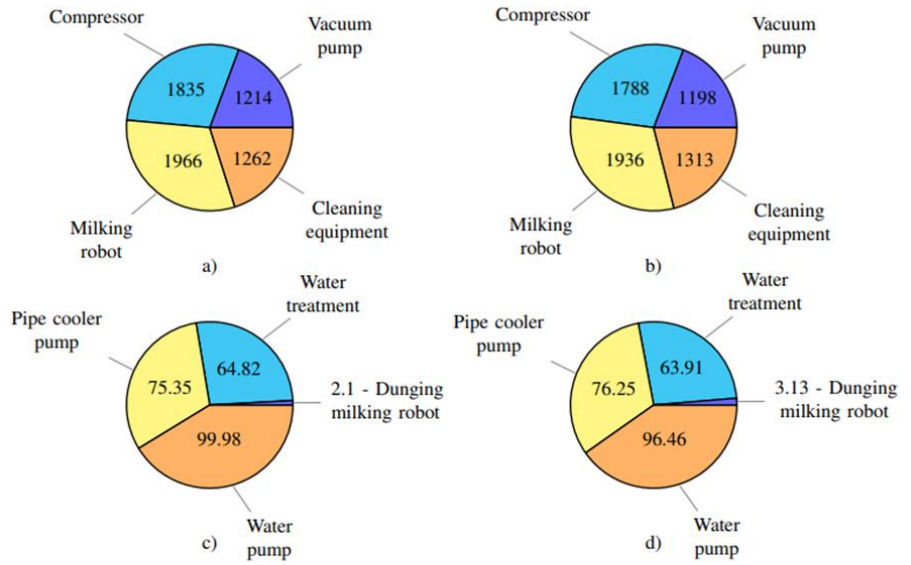
➢ Limited progress on non-intrusive load monitoring (NILM), i.e., energy disaggregation, for dairy farms where energy consumption of milk production equipment ***has not been quantified from actual data***

➢ Quantifying electricity consumption of machines present on dairy farms can be a ***valuable measure for decision support and planning***

Todic, T, Stankovic, L, Stankovic, V & Shi, J 2022, 'Quantification of dairy farm energy consumption to support the transition to sustainable farming', Paper presented at International Conference on Smart Computing 2022, Espoo, Finland, 20/06/22 - 24/06/22.

Milking robot active power signatures



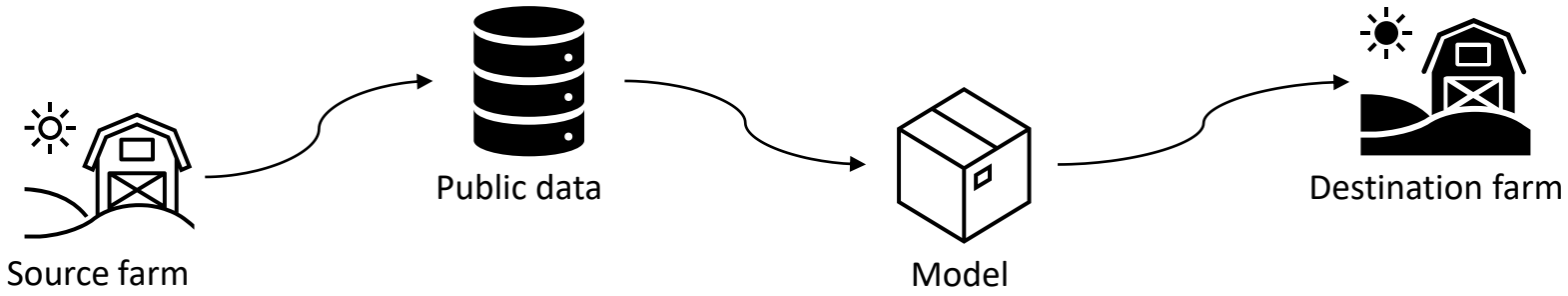


Consumption by device during the test period in kWh.
a) Actual consumption at location 18. b) Predicted consumption at location 18. c) Actual consumption at location 21. d) Predicted consumption at location 21.

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

Disaggregation results

Transferability between farms



		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

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

Transfer of models from one farm to another was not ***successful***, due to ***differences in used equipment***, as well as ***different labelling approaches of farmers.***

Research questions

- Is WaveNet-based deep learning, shown to perform best for energy disaggregation in residential settings, capable of performing disaggregation in ***industrial settings***, i.e., on dairy farms?
- Which ***metrics*** are meaningful for measuring the success of energy disaggregation?
- Are models trained on one farm ***transferrable*** to other dairy farms

Conclusions

- Milking-related devices can be ***successfully disaggregated*** from aggregate measurements using the algorithm initially developed for residential data
- ***Mean absolute error*** is ***not a reliable metric*** for regression tasks – it depends on the range of values in the data, so can be high even if the performance is good and vice versa.
- ***Match rate*** is more informative – gives the percentage at which true consumption overlaps with model's prediction.
- ***Transfers*** of models between farms were ***not successful***

NILM end uses => virtual submeter

- Appliance and food life cycle analysis
- Anomaly detection: detecting anomalous use of appliances or unusual usage patterns
- Load shifting, i.e., exploiting flexibility in time-of-use of appliances to manage peak demand, with the incentive of lower tariffs and improve demand response
- Retrofit advice, i.e., installing replacement appliances or energy savings measures
- Smart home automation, for improving energy conservation, comfort and security in the home
- Office of National Statistics – quality of life inferences via smart meter analytics
- Activity recognition, i.e., analysing energy consumption through the lens of activities, potentially more meaningful to users
- Monitoring of health and wellbeing, e.g., vulnerable population, elderly living alone

➢ NILM reveals a lot about you

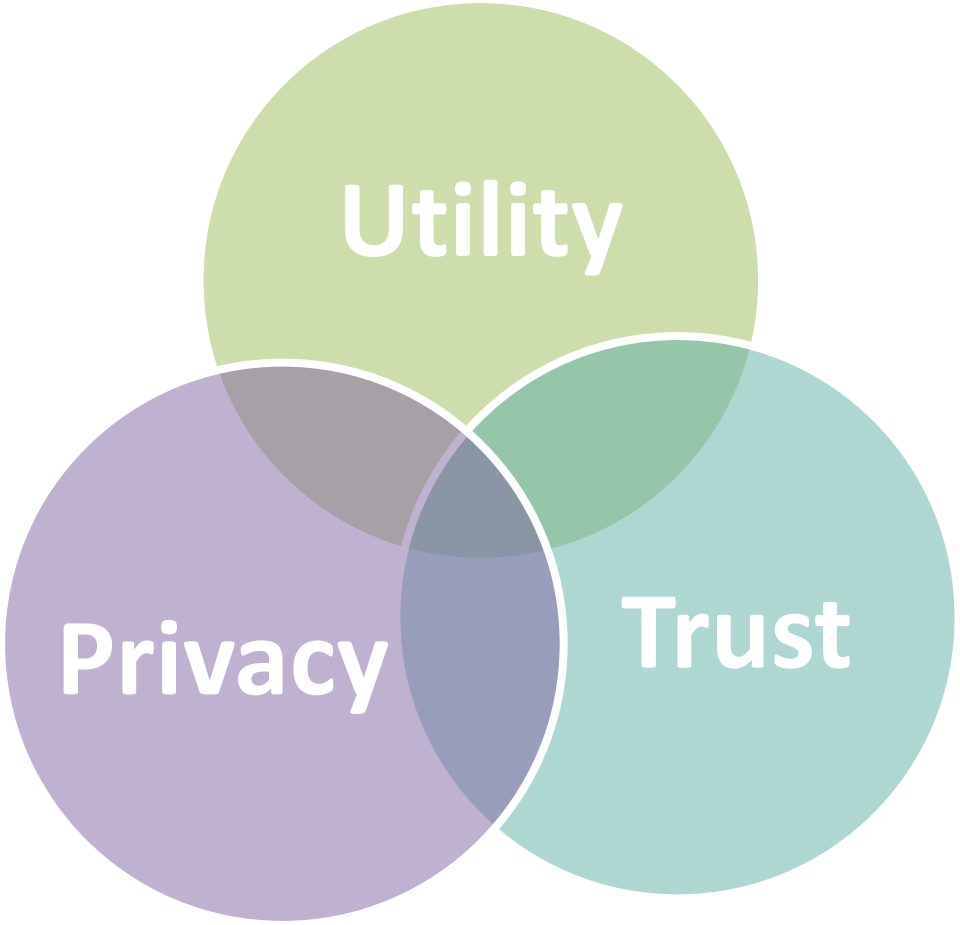
Risks & benefits of NILM from smart meters

University of Strathclyde Engineering

"*further effort is needed to clearly identify the benefits and risks that are opened up by AI engines leveraging machine learning algorithms before adopting them for decision making. The machine learning algorithms must be made transparent to users to ensure trust. Investigation into appropriate data transformation tools to alter the data is needed, so that it is useful but not harmful, and recommendations can be inferred, depending on the level of individual consent to sharing their data. It is also important to explain clearly to individuals what the benefits and risks are when their personal data is being shared for analysis. As the technology evolves, more can be inferred from smart meter data, so the review of benefits and risks is a continuous process. This should be performed by a range of stakeholders and competencies, including those understanding the sensing elements of the meter and measurement signals to those making inferences using machine learning, and those managing and offering the services. While machine learning technology opens up risks of misuse and misinterpretation of data, this technology is also the solution to provide accountability as personal meter data flows from the home, where it is captured to the cloud, to the technology provider(s) and finally end-user.*"

Stankovic, L., & Stankovic, V. (2020, May 11). The risks and benefits of AI smart meters. Apolitical.

To guarantee the ethical application of AI in NILM and foster trust in the technology, it is crucial to:
Ensure AI NILM algorithms do not infringe upon user privacy and security.
Encourage transparency and comprehension of the prediction process by adopting a clear and understandable approach to AI NILM algorithms



Utility

Accuracy via fair metrics, reliability (inc. repeatable, reproducible) => robustness

Privacy

NILM analyse highly sensitive and personal data that can reveal daily activities and habits of a household. User privacy and protection against malicious use of data

Trust

Above plus transparency, fairness, accountability

Trustworthy AI

- European Commission has recently published seven principles of Trustworthy AI, inc.:
  - Human Agency and Oversight, Technical Robustness and Safety, Privacy and Data Governance, Transparency, Diversity, Non-discrimination and Fairness, Societal and Environmental Well-Being and Accountability.
- Trustworthy AI-based NILM system design implications
  - a human-in-the-loop approach which enables human intervention during the design cycle of the NILM system and monitoring the system's operation,
  - designing NILM systems for high accuracy but also highlighting how likely errors are for occasional inaccurate predictions,
  - reliability of design by ensuring that the NILM design works for a range of inputs and situations, by demonstrating performance on different houses with different appliance ownership,
  - traceability to enable transparency by leveraging on public datasets, where data gathering, labelling and performance with different algorithms are well documented,
  - explainability for transparency by providing explanations of the NILM system's decision making process,
  - communication for transparency by clearly identifying the level of accuracy and limitations,
  - a low complexity methodology to ensure implementation is environmentally friendly without resorting to large data centres since the system can run locally,
  - keeping NetZero and sustainability goals in mind, e.g., UN Sustainable Development Goal 7 by enabling responsible consumption of energy and an affordable and modern energy service.

European Commission, Directorate-General for Communications Networks, Content and Technology, Ethics guidelines for trustworthy AI. Publications Office; 2019. Available from: doi/10.2759/346720

Utility - Robustness

- Robustness refers to the capacity to handle unseen, noisy, or incorrect data. This can be addressed at the levels of data, evaluation metrics, or algorithms:
  - Data level – Typically targets issues related to inadequate or imbalanced data, often employing data augmentation techniques. Methods involve random appliance activation imputation in aggregate data, use of GANs to generate realistic power signatures, and active learning approaches.
  - Evaluation level – Focuses on developing new metrics that assess generalizability and transferability. Techniques include metrics that more accurately gauge performance on unseen data and energy-based metrics that quantify the proportion of predicted, missing, and extra energy.
  - Algorithm level – Aims to create algorithms that offer strong transferability or generalizability performance.

Utility: Example of good practice

Literature Review	Comparison of different methods & suitability assessment for model reuse Huber, P.; Calatroni, A.; Rumsch, A.; Paice, A. Review on Deep Neural Networks Applied to Low-Frequency NILM. Energies 2021, 14, 2390
Model Selection	Maximise code reusability: sequence-to-subsequence DNN trade-off between convergence speed of seq2seq and computational load of seq2point
Noisiness & Sparsity	Calculation of the noisiness and of the sparsity of the load signal could predict accuracy of inference in a dataset
Evaluation Metrics	Selection of robust and context-aware metrics Match Rate & *Acc* vs commonly used metrics such as MAE
Rigorous Experimental Evaluation	Generalisability on unseen households with similar EV load profiles Cross-domain transferability on unseen households with different EV load profiles
Labelling	Transferring knowledge on unlabeled datasets Verification of results via expert knowledge
DNNs Adaptation	DNNs for 3-phased installations (predominant on the Continent) Devices connected in arbitrary permutation

Table 4. Mean values and standard deviations of assessment metrics, training and testing on the same household, using sequence-to-subsequence, 1 min data.

Area	House	MAE (μ, σ)	SAE (μ, σ)	*Acc* (μ, σ)	MR (μ, σ)
Austin	661	(45.45 W, 6.28 W)	(8.41%, 0.25%)	(89.09%, 1.51%)	(79.55%, 1.24%)
	1642	(55.13 W, 11.52 W)	(4.19%, 0.14%)	(88.87%, 0.74%)	(79.59%, 0.79%)
	4373	(82.33 W, 14.09 W)	(1.72%, 0.07%)	(89.57%, 1.11%)	(80.94%, 0.98%)
	4767	(210.6 W, 41.86 W)	(5.47%, 0.27%)	(58.19%, 3.57%)	(48.52%, 1.89%)
	4767-1	(219.1 W, 48.55 W)	(59.94%, 0.88%)	(46.84%, 2.59%)	(41.22%, 1.26%)
	4767-2	(490.2 W, 74.57 W)	(83.36%, 0.86%)	(30.98%, 2.10%)	(24.78%, 0.97%)
	6139	(103.2 W, 13.97 W)	(4.15%, 0.11%)	(70.36%, 1.47%)	(53.53%, 1.05%)
	8156	(82.33 W, 14.09 W)	(1.72%, 0.06%)	(89.57%, 1.10%)	(80.94%, 0.98%)
NY	27	(86.78 W, 19.74 W)	(3.47%, 0.08%)	(79.14%, 0.49%)	(74.38%, 1.14%)
	1222	(148.1 W, 28.17 W)	(5.73%, 0.07%)	(53.27%, 1.97%)	(45.11%, 1.47%)
	5679	(73.65 W, 6.75 W)	(16.22%, 0.72%)	(85.29%, 1.72%)	(79.51%, 1.18%)

ID	MAE	SAE	Acc	MR	G-loss
			1 min		
NY5679	183.4W	24.32%	61.15%	58.67%	28.30%
			15 min		
NY5679	201.5W	29.60%	56.92%	48.83%	10.01%

Table 2. Mean values of performance and generalisation loss metrics for transferability test: (on an unseen house in NY and training on all other houses from Austin, regardless of EV charging level).

ID	MAE	SAE	Acc	MR	G-loss
			1 min		
AU1642	50.49W	4.54%	89.81%	81.12%	-1.06%
NY5679	178.4W	26.94%	62.21%	49.21%	27.06%
			15 min		
AU1642	78.66W	11.69%	84.13%	71.15%	-21.72%
NY5679	183.6W	33.00%	60.75%	46.04%	3.95%

Table 3. Mean values of performance and generalisation loss metrics, for transferability test (Testing on an unseen house with EV charge level of 6.6 kW and training on all houses with EV charge level of 3.3 kW, regardless of geographical area)

- Standard deviation for all metrics over 10 runs of the experiment, in identical conditions, is less than 4%, which indicates that all the experiments are repeatable. MAE metric has a 10-15% std
- Context-specific metrics, Match Rate and Acc, are more robust and meaningful instead of commonly used metrics such as MAE → improved trustworthiness of model outputs;
- Small, but acceptable loss in EV charging consumption estimation accuracy incurred when making inference on lower granularity smart meter data → utility preserved to a certain extent;
- Transfer quality between different datasets measured by generalization loss, ratio between the error from unseen houses and seen houses;
  - Testing on an unseen house in NY (6.6kW) and training on all other houses from Austin area (3.3kW): algorithm identifies EV loads but fails to assign the correct amount of energy
  - Testing on two unseen houses and training on a mix of houses (areas and charging levels): drop in performance in house NY5679 due to different power levels; increase in performance in AU 1642 at 15min due to greater data availability

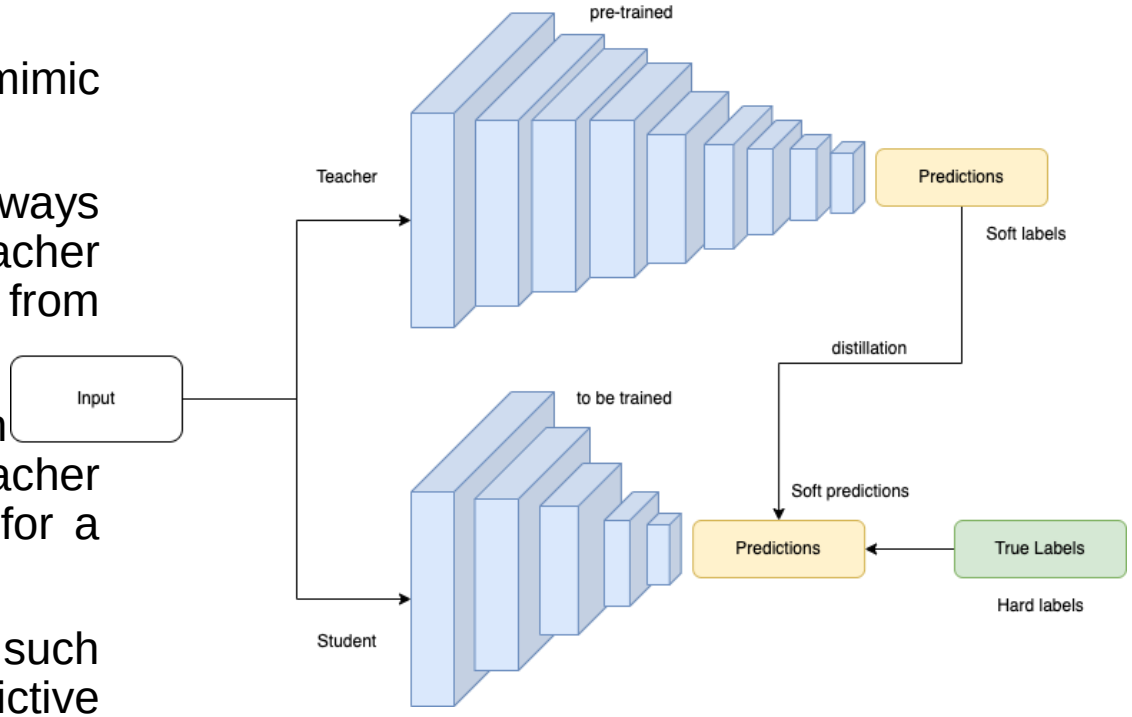
Vavouris, A., Garside, B., Stankovic, L., & Stankovic, V., Shi, J. (2022). Low-frequency non-intrusive load monitoring of electric vehicles in houses with solar generation: generalisability and transferability. Special Issue Digital Transformation in the Energy Sector: Data-Driven Analytics, Services and Business Models, Energies, 15(6).

Privacy

- Numerous household activities, such as laundry, dishwashing, cooking, and boiler heating, exhibit unique and discernible power signatures. This information can reveal the types of activities being performed by family members, their presence at home, or their sleep patterns, potentially leading to malicious exploitation, inc. targeted advertising.
- Privacy – design that ensures user privacy, and protects against malicious use of data.
- Unauthorized data access risks in NILM can occur through data breaches, AI model inference, or the AI model itself. Common risk mitigation strategies include:
  - Data anonymisation – aimed at modifying the data in a way that knowledge contained in the data cannot be attached to a person. Common approaches include k-anonymity, noise injection and differential privacy.
  - Federated learning (FL) – Contrasting with traditional AI learning approaches that gather data on a central server, FL trains models on edge devices (e.g., smart meters) without sharing data, communicating only model parameters to the central server.
  - Hardware protocols – Implementing hardware security protocols to safeguard data privacy and prevent attacks. A prevalent approach is the trusted execution environment (TEE).

Example of Explainability guided learning for Privacy

- Knowledge Distillation (KD) is a technique used in edge deployment for model compression
- It is achieved by conditioning a lightweight “student” model to mimic the behaviour of larger, more complex “teacher” model.
- However, we observe that in the NILM setting, KD might not always succeed in transferring explainable knowledge from the teacher (trained on large datasets) to the student (trained on soft labels from teacher and weak labels from target) model.
- In particular, we identify the main type of inconsistency in transfer of explainable knowledge: “Given identical inputs, teacher and student networks produce dissimilar output explanations for a given class of interest”.
- We propose a learning method for minimization of such inconsistencies that also leads to improvement in predictive performance – explainability guided learning

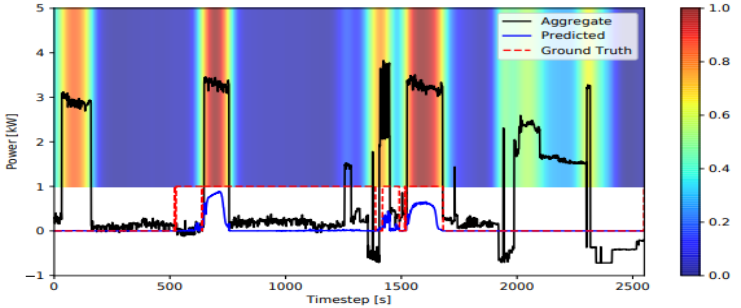


D. Batic, G. Tanoni, L. Stankovic, V. Stankovic, and E. Principi, “Improving knowledge distillation for non-intrusive load monitoring through explainability guided learning.” in 2023 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2023), IEEE, 2023.

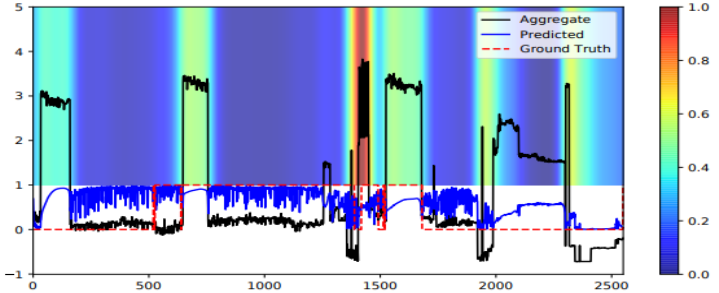
Explainability Guided Learning

- Methodology to reduce inconsistencies, enhancing the optimization of the distillation process and promoting more stable predictive performance
- by modifying the KD loss function and guiding the student network to minimize the discrepancy between teacher and student prediction explanations.
- Validated on UK-DALE and REFIT, and evaluated in two domain adaptation scenarios over 5 different appliances.
- Ablation studies reveal that our learning method can yield improvements ranging from 1.6% up to 22.6% improvement over the baseline.
- Address privacy, scalability, transparency/explainability, and transferability

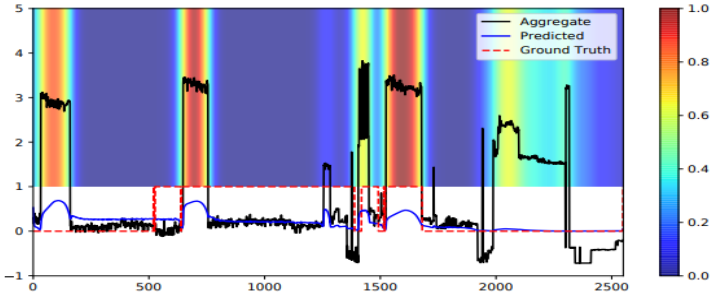
(a)



(b)

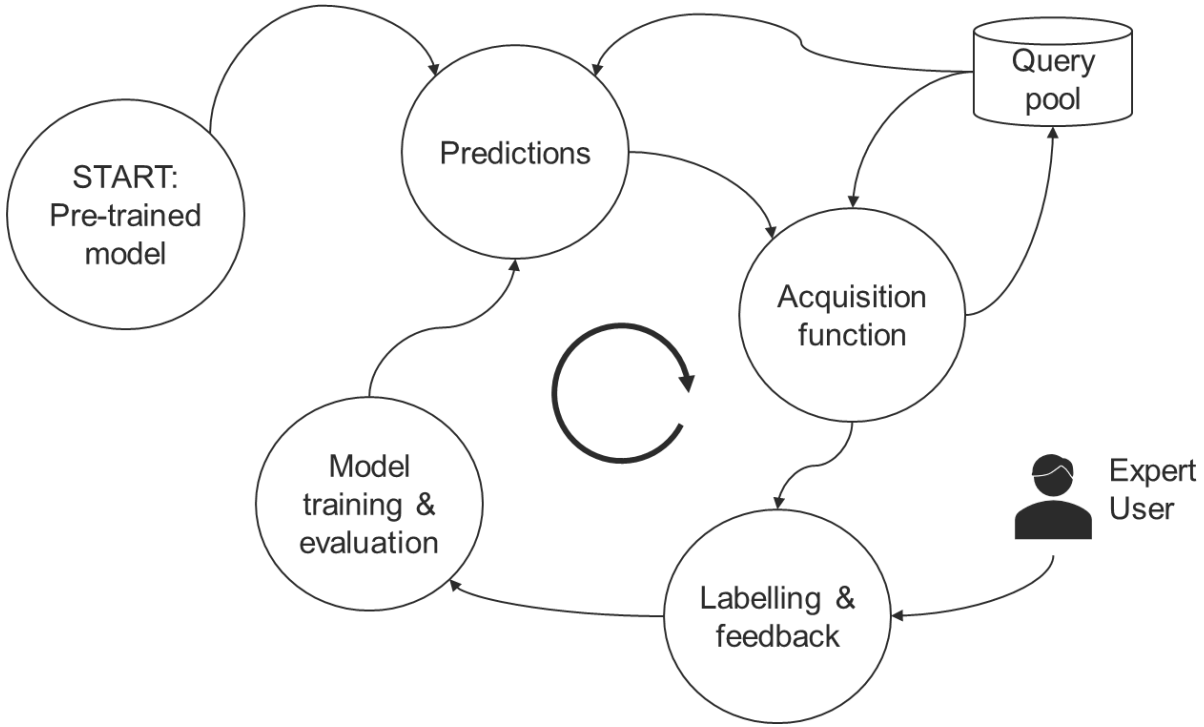


(c)



Active learning – human-in-the-loop

- Large labelled datasets are usually required to train a deep learning based NILM algorithm,
- Not all data samples are equally important for model training – usually there is high redundancy in data.



AL chooses the worthiest of unlabeled data samples (query pool) to label and include in model training via specially designed acquisition functions. AL runs iteratively => Acquisition function ranks all data samples by their informativeness and chooses the best combination to improve the model performance the most.=>After data samples are chosen, they are sent to an expert or a user for labelling=>After being labelled, they are included in model training=> Once all worthy samples are included in the training, the others won't contribute much to performance improvement.=> Active learning results in achieving peak algorithm performance with significantly reduced labelling effort.

Todic, T., Stankovic, V., & Stankovic, L. (2023). An active learning framework for the low-frequency Non-Intrusive Load Monitoring problem. Applied Energy, 341, [121078]. https://doi.org/10.1016/j.apenergy.2023.121078

How does AL comply with Trustworthy AI principles?

- Having an already developed NILM model, it can be transferred to a new environment – house of a new user, where it can be adapted to new conditions and performance can be boosted via active learning (technical robustness).
- Active learning includes users in the process of development of AI algorithms by asking them to provide labels and to give their feedback on model behavior (human agency).
- Adaptation to a new house can be done in that house so no sensitive data has to be exported to the outside (privacy and data governance).
- During the active learning process, users are involved in model training – the model asks them for help with the most challenging parts, so they get the sense of model's strengths and weaknesses (transparency). This helps them build realistic expectations of the model's behavior, and also calibrates their trust and confidence about using AI algorithms daily in their homes.

Conclusions

- Future of AI-based NILM adoption:
  - Reliable supervised NILM models that are **transferable** to ‘unseen’ datasets or **reliable unsupervised** NILM methods that can operate on any dataset
  - Reliable NILM methods that focus on accurate disaggregation of **challenging loads**;
  - Novel datasets, data models, and toolkits for NILM research especially for **emerging loads** such as EVs, heat pumps and **hard to decarbonise sectors**, e.g. agriculture, manufacturing
  - NILM feasibility for **non-residential sectors**, esp. hard to decarbonize sectors
  - **Fair and explainable metrics** for the evaluation of different NILM algorithms;
  - **Interpretable and explainable** algorithms for NILM;
  - Practical NILM deployments (**robustness, scalability, transferability, privacy preserving**) and large-scale trials
  - Adherence to European Commission principles of **Trustworthy AI**, inc.:
    - Human Agency and Oversight, Technical Robustness and Safety, Privacy and Data Governance, Transparency, Diversity, Non-discrimination and Fairness, Societal and Environmental Well-Being and Accountability.
  - Keeping **NetZero and sustainability goals** in mind
  - **Applications** leveraging on NILM disaggregated data (e.g., flexibility estimation, life cycle analysis);

University of Strathclyde Glasgow

DEPARTMENT OF ELECTRONIC & ELECTRICAL ENGINEERING