Complex mixed-methods research design approaches for data mining to meet utility and trust requirements for three-phase energy end-use customers

A thesis presented for the degree of Doctor of Philosophy

Apostolos Vavouris

Department of Electronic and Electrical Engineering
University of Strathclyde, Glasgow

May 14, 2025

Abstract

This thesis explores complex mixed-methods research design approaches for data mining to meet utility and trust requirements for three-phase energy end-use customers (residential and industrial) of electricity networks. The developed methods explore how much information can be inferred about individual load usage from single- and three-phase installations through Non-Intrusive Load Monitoring (NILM) using granular metered data, from low-frequency (1-sec) power data obtained through special metering equipment to very-low frequency (30-min) energy data obtained through the available smart metering infrastructure. Generalisability, i.e., training and testing in different premises that exhibit similar load patterns, and transferability, i.e., training and testing in different premises that exhibit different load patterns, has been explored in order to co-design scalable downstream NILM applications with end-users. Quantitative data, i.e., metered physical quantities, are complemented by qualitative data obtained from interviews and time-of-use surveys in order to develop complex mixed-methods data mining approaches to inform: (i) the usage component of lifecycle assessment (LCA) models of electric vehicles (EVs); (ii) the evaluation of energy-efficiency in net-positive energy households; and, (iii) load-scheduling of energy intensive activities for the residential and industrial sectors. Therefore, the goals of the thesis can be summarised as follows: (i) explore the effects of different levels of smart metering data granularity on the load disaggregation accuracy and robustness; (ii) improve the load disaggregation accuracy when using very-low frequency data obtained through the smart metering infrastructure by exploiting three-phase information in residential and industrial settings; (iii) generalise and transfer the trained models across different settings at scale; (iv) inform the usage component of LCA models of EVs and compare them fairly with the

fossil-fuelled equivalents, through the disaggregation of EV charging loads in residential customers and integration of drivers' charging routines by combining quantitative — i.e., energy consumption and production timeseries and granular spatio-temporal carbon intensity of the electricity network — and qualitative data including interviews and time-of-use surveys; (v) develop complex mixed-methods data-driven energy centric evaluation methods of net-positive households' methodology to answer the "what", "why" and "how" of energy prosumption in net-positive energy neighbourhoods through the disaggregation of energy-intensive load to the activity level and explore the potential of load scheduling on an activity basis based on different households profiles and flexibilities; and, (vi) develop a co-created NILM-enabled data driven methodology to improve load scheduling in the dairy sector and reduce the utility costs and the carbon footprint of farms, through the collaboration with various stakeholders during the design, data collection, implementation, and feedback process. By achieving these goals, this thesis aims to address all three levels (biosphere, society, and economy) of the Sustainable Development Goals (SDGs). By enhancing LCA of EVs, national policies can be directly influenced (SDG 13.2), while at the same time reducing pollution for sustainable cities (SDG 11.6), ensuring access to clean energy (SDG 7.1), and promoting responsible electricity consumption (SDG 12.2). The mixed-methods approach for net-positive buildings supports climate action policies (SDG 13.2), sustainable urbanisation (SDG 11.3), clean energy access (SDG 7.1), and efficient resource use (SDG 12.2). In the agricultural sector, research improves renewable energy generation, self-consumption, and sustainability (SDGs 7.2, 11.5, 12.2, 13.2), enhances resource efficiency and clean technology adoption (SDG 8.4, 9.4), and reduces inequalities by supporting income growth in Less Favoured Areas (SDG 10.1).

Contents

Contents

Contents

Contents

List of Figures

List of Figures

ix

List of Figures

List of Tables

List of Tables

Acknowledgements

Funding

Chapter 0. Acknowledgements

Nomenclature

AC	Air conditioning	DNO	Distribution network operator
ADAM	Adaptive moment estimation	DR	Demand response
AI	Artificial intelligence	DT	Decision tree
ANN	Artificial neural network	DTW	Dynamic time warping
BEIS	Department of business energy and industrial strategy	DW	Dishwasher
		ENTSO-E	European network for transmission system operators for electricity
BEV	Battery electric vehicle		
BHI	Beam horizontal irradiance	ESO	Electricity system operator
BNI	Beam normal irradiance	EU	European Union
CM	Coffee machine	EV	Electric vehicle
CO_2	Carbon dioxide	FN	False negatives
COP	Coefficient of performance	FP	False positives
CTC	Cradle to cradle	FRD	Fridge
CTGa	Cradle to gate	FRZ	Freezer
CTGr	Cradle to grave	GAN	Generative adversarial network
DHI	Diffuse horizontal irradiance	GB	Great Britain
DHW	Domestic hot water	GDPR	General data protection regulation
DL	Deep learning	GHG	Greenhouse gases
DNN	Deep neural network	GHI	Global horizontal irradiance

Nomenclature

GRU	Gated recurrent unit	Mt	Mega-tonne
GSEE	Global solar energy estimator	MW	Microwave
GSP	Graph signal processing	NI	Northern Ireland
HB	Electric hob	NILM	Non-intrusive load monitoring
HMM	Hidden markov model	NM	Noisiness metric
HT	Heating appliance	NOK	Norske Krone
HVAC	Heating, ventilation, and air conditioning	NY	New York
ICA	Independent component analysis	OV	Electric oven
ICEV	Internal combustion engine vehicle	PC	Principal component
IEA	International energy agency	PCA	Principal component analysis
IHD	In-home displays	PED	Positive energy districts
ILUC	Indirect land-use change	PEN	Positive energy neighbourhood
KET	Kettle	PHEV	Plug-in hybrid electric vehicle
kNN	k-nearest neighbour	POWER	Prediction of worldwide energy resource
L1	Least absolute deviation	PV	Photovoltaic
LCA	Lifecycle assessment	REF	Fridge-freezer
LCT	Low carbon technology	ReLU	Rectified linear unit
LFA	Less favoured areas	RES	Renewable energy sources
LSTM	Long short-term memory	RF	Random forest
MAE	Mean absolute error	RMSE	Root mean square error
ML	Machine learning	RNN	Recurrent neural network
MLP	Multi-layer perceptron	SAE	Signal aggregate error
MR	Match rate	SDG	Sustainable development goals
MSE	Mean square error	Seq2point	Sequence-to-point

Nomenclature

Seq2seq	Sequence-to-sequence
Seq2subseq	sequence-to-subsequence
SGD	Stochastic gradient descend
SHT	Smart home technologies
SMETS2	Smart meter equipment technical specifications version:2
SP	Submetering point
SVM	Support vector Machine
TD	Tumble dryer
TECA	Total energy correctly assigned
TN	True negatives

TP	True positives
TTW	Tank to well
UK	United Kingdom
UTC	Coordinated universal time
V2G	Vehicle-to-grid
VAT	Value-added task
WD	Washer-dryer
WM	Washing machine
WTT	Well to tank
WTW	Well to wheel

Chapter 1

Introduction

The growing adoption of smart metering technologies and the increase of energy-intensive equipment and devices through the introduction of low carbon technologies (LCTs), such as electric vehicles (EVs) and electrical heating appliances (HTs), have underscored the importance of accurate energy monitoring and management. Non-Intrusive Load Monitoring (NILM) has emerged as a promising solution to disaggregate energy consumption from aggregated signals, enabling appliance-level monitoring without requiring extensive submetering [2]. This capability, which has been extensively studied, is critical for addressing key challenges in energy efficiency, renewable energy sources (RES) integration, and demand-side management [3].

In three-phase installations, traditionally, load disaggregation has focused on the sum of three phases, which can obfuscate the individual load patterns of appliances, particularly those with sparse activations or variable load profiles [4]. Recent updates in national smart meter rollouts [5] to include three-phase metering provide an opportunity to refine load disaggregation methodologies and enhance the accuracy of appliance identification. This is especially timely as the adoption of high-power loads, such as EV chargers, continues to rise. Moreover, the per-phase disaggregation approach offers a way to mitigate false positives (FPs) caused by overlapping signals from similar devices or unknown appliances [6], enabling better energy insights and optimisation.

In smart housing developments, the growing emphasis on net-zero and net-positive energy designs highlights the need for robust energy monitoring systems that align with

households' day-to-day practices, renewable energy production, and dynamic energy pricing [7,8]. Beyond residential applications, NILM and energy management systems hold significant potential for other domains, including the agriculture sector [9]. In agricultural settings, such as dairy farms, energy-intensive, largely non-standardised processes like milking, cooling, and feeding can benefit from load scheduling and optimisation informed by NILM.

Despite these advances, challenges remain in the areas of data granularity, computational efficiency, and generalisability across diverse settings [10]. For instance, the move toward lower-frequency smart meter data, driven by privacy concerns, introduces new complexities in maintaining disaggregation accuracy [3,10]. Furthermore, while NILM algorithms have shown success in specific use cases, their performance across geographic regions, varying appliance configurations, and emerging technologies, such as voluntary milking systems, requires further investigation [11].

This thesis aims to address these challenges through co-creation with the different stakeholders by exploring advanced, co-designed, scalable NILM methodologies for carbon-intensive, but often overlooked applications, such as EV charging, emerging net-positive energy communities, and agriculture. By leveraging transfer learning, improved data mining through per-phase load monitoring, mixed-method evaluations, and scalable approaches, this work contributes to the broader goal of optimising energy usage, reducing carbon footprint, and supporting the transition to sustainable energy systems.

1.1 Research motivation and aims

The motivation of this research was the exploration of complex mixed-methods research design approaches for data mining to meet utility and trust requirements for three-phase energy end-use customers. The question of how much meaningful information we can infer from low and very-low frequency energy data to meet utility and trust requirements and enable downstream applications of NILM was explored. Due to the introduction of residential high-power loads (such as EVs) on the power grids as well as the electrification of various largely non-standardised industrial processes, data mining

of load-specific information gains traction as an enabler technology for load shifting, flexibility services, and demand response (DR). These lead to the following research questions (RQ):

R.Q.1 How can we improve the accuracy of NILM-based data mining methods by exploiting three-phase data? How can we efficiently label datasets for further analysis by exploiting the three-phase information?

R.Q.2 How can we adapt NILM-based data mining methods to EV load disaggregation? How can we quantify generalisability and cross-domain transferability of NILM-based data mining approaches?

R.Q.3 How can we inform the usage component of current EV LCA models by incorporating quantitative and qualitative data that reflect on the time of use of the charging infrastructure?

R.Q.4 How can we infer the "what", "why", and "how" of energy prosumption in residential settings and enable energy efficiency initiatives through complex mixed-methods approaches? How can we incorporate qualitative data into NILM-based data mining methods?

R.Q.5 How can we co-create, with different stakeholders, a cost-efficient load scheduling approach for the largely non-standardised dairy sector through NILM-enabled cross-domain transferability to reduce the utility costs and the carbon footprint?

Chapter 3 provides an answer to R.Q.1, where the open-source three-phase ECO data set [4] is used to explain the effects in the accuracy of data mining when using aggregate and per-phase readings (under different granularities scenarios), while a previously unlabelled EV dataset from a household in Germany is labelled. The labelling of prior unlabelled household energy data is further discussed in Chapter 5 & Appendix A where six prior unlabelled household datasets are labelled through the usage of information derived from the intersection of three-phase data and qualitative surveys.

Chapter 4 discusses R.Q.2 by investigating the performance of a regression-based approach in estimating the EV load under different granularities scenarios. Generalis-

ability and cross-domain transferability are also discussed in this chapter, which leads to the investigation of R.Q.3 where results obtained through transferability and qualitative data informed the usage component of lifecycle assessment (LCA) models for EVs.

Chapter 5 provides a mixed-methods approach as an answer to R.Q.4 where transferability is intersected with qualitative data obtained from surveys and semi-structured interviews to understand energy prosumption and enable energy efficiency initiatives.

Lastly, Chapter 6 discusses three-phase NILM-based load disaggregation and transferability, for the largely non-standardised dairy sector, as an enabler for a co-created load scheduling approach and thus answering R.Q.5.

1.2   Contribution of thesis

This research investigates how NILM-enabled data mining methods can be co-created to meet the utility and trust requirements of three-phase energy end-use customers in electricity networks. In summary, the main contributions of this thesis are as follows:

- Quantification of the benefits arising from three-phase energy monitoring and load disaggregation under different granularity scenarios; and a methodology to label a previously unlabelled electricity dataset through the exploitation of three-phase information (Chapter 3).
- The adaptation of a sequence-to-subsequence (seq2subseq) deep neural network (DNN)-based NILM approach for EV load disaggregation and the rigorous evaluation of the approach under different testing scenarios; the quantification of generalisability and cross-domain transferability of the proposed methodology (Chapter 4).
- The augmentation of existing EV LCA models through incorporation of usage factors that impose specific time of charging patterns obtained through a combination of NILM outputs and qualitative data (Chapter 4).
- A new complex mixed-methods energy-centric data-driven NILM-enabled approach to provide tools to explain the "what", "why", and "how" of energy prosumption in residential settings and enable energy efficiency initiatives (Chapter 5).

- Labelling, curation, and release of the first granular energy load consumption and production dataset (NorPEN) for a positive energy neighbourhood (PEN) in Norway, accompanied by metadata and interview transcripts with the homeowners (Chapter 5 & Appendix A).

- A co-created NILM-enabled framework that enables load scheduling in the dairy sector to reduce the utility costs and carbon footprint under different monitoring scenarios and the introduction of a post-processing step that improves NILM regression accuracy across similar equipment with different manufacturers' specifications (Chapter 6).

- The curation and release of the most extensive, comprehensive, labelled, open source dataset (FIELD) of dairy electrical equipment from 30 three-phase dairy farms in Germany (Chapter 6 & Appendix B).

1.2.1   Contribution to trustworthy data mining methods

The co-creation of user-centric data mining methods to meet utility and trust requirements in energy networks enabled the advancement of machine learning (ML) research in relation to trustworthy artificial intelligence (AI) principles and sustainability goals. More specifically, through the process of co-design (see Chapter 5) and co-creation (see Chapter 6) the end-user was brought to the forefront of the development of the AI methods, directly influencing them and therefore increasing the transparency of the AI systems with their decisions explained in a manner adapted to the stakeholder concerned. Further to that, human agency and oversight have been enabled through the co-design process by allowing end-users to make informed decisions based on the outcomes of the AI models. Through the research carried on quantifying the amount of information that can be extracted from smart meter data streams and the parameters that can affect the trustworthiness of a data mining ML method (see Chapters 3 & 4) the technical robustness of the solutions has been fostered through more accurate, reliable and reproducible experiments. Societal and environmental well-being, as well as ensuring diversity, non-discrimination, and fairness, has been achieved through the community-wide research involving different groups, including vulnerable people and businesses

located in severely disadvantaged Less Favoured Areas (LFAs) (see Chapter 6).

1.2.2 Contribution to sustainable development goals (SDGs)

Sustainability was at the forefront of this research by simultaneously tackling all three levels of SDGs, including the biosphere, society, and economy. Accurate disaggregation of EV charging and the subsequent uncovering of end-users' routines enabled the more accurate LCA of different vehicular technologies (see Chapter 4), which can directly impact national policies, strategies and planning (environment: SDG 13 Climate Action, target 13.2). On the same time EV research impacts cities and human settlements and make them more sustainable by reducing the annual mean levels of fine particulate matter (society: SDG 11 Sustainable Cities and Communities, target 11.6) while ensuring access to affordable, reliable, sustainable, and modern energy for all members of the community (society: SDG 7 Affordable and Clean Energy, target 7.1). EV research ensures more sustainable consumption patterns in terms of electricity consumption patterns (economy: SDG 12 Responsible Consumption and Production, target 12.2).

Through the mixed-methods approach for the evaluation of net-positive energy buildings (see Chapter 5) SDGs were again targeted, the results of which can directly impact national policies, strategies and planning (environment: SDG 13 Climate Action, target 13.2). Further to that, through a mixed-methods approach with the end-user and its needs at the forefront, inclusive and sustainable urbanisation and capacity for participatory, integrated and sustainable human settlement planning and management (society: SDG 11 Sustainable Cities and Communities, target 11.3) while ensuring access to affordable, reliable, sustainable, and modern energy for all members of the community (society: SDG 7 Affordable and Clean Energy, target 7.1) was enabled. The better management of available RES and more responsible consumption enabled through the research accelerates the efficient use of natural resources (economy: SDG 12 Responsible Consumption and Production, target 12.2).

Lastly, the research carried out in the agricultural sector (see Chapter 6), that enhances the renewable energy generation estimation, increases the consumption of self-generated energy, and fosters more sustainable consumption patterns also tackles SDGs

7 (7.2), 11 (11.5), 12 (12.2), and 13 (13.2). Further to that, the resource efficiency in consumption and production and the decoupling of economic growth from environmental degradation is improved with developed countries taking the lead (economy: SDG 8 Decent Work and Economic Growth, target 8.4) through the upgrade of infrastructure and retrofit of the dairy sector to make it more sustainable, with increased resource-use efficiency and greater adoption of clean and environmentally sound technologies and industrial processes, with all countries taking action in accordance with their respective capabilities (economy: SDG 9 Industry Innovation and Infrastructure, target 9.4). Lastly, as the research focused on designated LFAs the reduction of inequalities within countries by progressively achieving and sustaining income growth of the bottom 40 per cent of the population at a rate higher than the national average through the better handling of the available energy has been achieved (economy: SDG 10 Reduced Inequalities, target 10.1).

1.2.3   Wider impact to the broader community beyond academia

Through the research carried out in this thesis, a wider impact has been achieved across society. Improvement of the usage component of LCA models that can better inform policymaking towards deployment of EVs (see Chapter 4) in different areas within and across countries has been achieved. Co-design and co-creation of user-centric mixed-methods approach to evaluate energy consumption practices and optimise load scheduling to maximise self-consumption directly benefited both residential end-users, through more affordable energy and reduced environmental carbon footprint of their consumed energy (see Chapter 5) with outputs scaled up to national level [12], as well as industrial (agricultural) end-users with a focus on the part of the communities that need the most support including farmers in areas that face lower than average production with a limited potential which cannot be increased except at excessive cost, and with low and dwindling population (see Chapter 6). Research results have been disseminated to non-academic parties, including educational outreach events and society/business engagement events; more specifically, research findings have been disseminated in the following events:

Chapter 1. Introduction

    I. "User-centric insights from low-frequency smart meter data analysis towards flexibility potential", Green Homes Exhibition networking event, 8 May 2024 (Barclays, Glasgow).

    II. "Efficient load scheduling for dairy farms with renewable energy provision", Engage with Strathclyde: Transforming the Dairy Supply Chain with Innovation & Collaboration, 1 May 2024 (University of Strathclyde, Glasgow).

    III. "Let's talk about AI: ethics, risks & sustainability", speaker/workshop facilitator: an online seminar/workshop for secondary education teachers, about advances in AI, ethics and risks, accompanied by a workshop on sustainability projects and AI. Organised in collaboration with the Directorates for Secondary Education (Imathia, Rethymno, Argolida), Ministry of Education, Religious Affairs, and Sports (Greece), 21 Mar 2024 (online).

    IV. "Revolutionising Dairy with Digital Innovation", how can we decarbonise the agricultural sector? A co-design approach on reducing carbon footprint/utility bills of dairy industry with the use of renewables, load flexibility and energy storage, 23 Jan 2024 (Prosper: Peer Works – Scottish Council for Development and Industry, online).

Lastly, the co-created load scheduling framework that revolutionises sustainable agriculture with digital innovation has been included in the University of Strathclyde Images of Research 2024 exhibition, winning the "Capturing Collaboration" award. The image, namely "Cow in the loop", that shows the hourly energy consumption & renewable energy production for a dairy farm located in Rhins of Galloway, a designated LFA of Scotland, has been exhibited in various places across Glasgow, including Glasgow Central Station as part of the Glasgow Science Festival. In summary, the value of developing ethical AI models that mitigate the risks that AI could involve has been demonstrated, with the focus placed on achieving the SDGs while achieving a prosperous and inclusive society.

1.3   Organisation of thesis

The thesis is organised as follows. Firstly, Chapter 2 provides background and preliminaries information across the different topics covered in the thesis. Chapter 3 introduces the benefits of three-phase NILM in residential settings for improved load disaggregation and provides a methodology for the labelling of high-power, three-phase loads while releasing the related dataset. Three-phase disaggregation presented in Chapter 3 forms the basis for Chapters 4, 5, and 6. Chapter 4 discusses a detailed and robust methodology for large-scale evaluation of EV load disaggregation from households with smart metering data, leveraging on prior NILM algorithms, while exploring the data mining potential under different granularities and transferability scenarios. Further to that, in Chapter 4, the usage of NILM for EV load disaggregation as an enabling technology to inform the usage component of LCA models of EVs is presented. Chapter 5 proposes a complex mixed-methods NILM-based approach to provide tools to explain the "what", "why", and "how" of energy prosumption in three-phased residential settings in Norway and enable energy efficiency initiatives, and introduces the NorPEN dataset. Chapter 6 presents a co-created NILM-enabled load scheduling approach that leverages three-phase disaggregation (see Chapter 3) and mixed-methods approaches (see Chapter 5) to reduce utility costs and carbon footprint of the dairy sector, and introduces the FIELD dataset. Appendix A contains further information regarding NorPEN dataset that informed the study in Chapter 5, whereas Appendix B includes additional information pertaining to FIELD dataset that has been used as the training dataset for the study in Chapter 6. Rather than a single literature review chapter, each chapter is self-contained, with the related background and literature review being included within the corresponding chapter.

1.4   Publications

Journals

 I. Vavouris, A., Stankovic, L., & Stankovic, V. A Non-Intrusive Load Monitoring-

Enabled Framework for Load Scheduling in the Dairy Industry. (Under review at Elsevier Applied Energy, Apr. 2025).

II. Vavouris, A., Stankovic, L., Stankovic, V., & Shi, J. FIELD: A comprehensive FarmIng Electrical LoaD measurements dataset from 30 three-phase dairy farms in Germany. (Under review at Nature Scientific Data, Jan 2025).

III. Vavouris, A., Guasselli, F., Stankovic, L., Stankovic, V., Gram-Hanssen, K., & Didierjean, S. (2024). Descriptor: A Norwegian Positive Energy Neighbourhood Dataset of Electrical Measurements and Interviews on Energy Practices (NorPEN). IEEE Data Descriptions, 1, 113-121. `https://doi.org/10.1109/IEEEDATA.2024.3483154`

IV. Vavouris, A., Guasselli, F., Stankovic, L., Stankovic, V., Gram-Hanssen, K., & Didierjean, S. (2024). A complex mixed-methods data-driven energy-centric evaluation of net-positive households. Applied Energy, 367, Article 123404. `https://doi.org/10.1016/j.apenergy.2024.123404`

V. Guasselli, F., Vavouris, A., Stankovic, L., Stankovic, V., Didierjean, S., & Gram-Hanssen, K. (2024). Smart energy technologies for the collective: time-shifting, demand reduction and household practices in a Positive Energy Neighbourhood in Norway. Energy Research and Social Science, 110, Article 103436.`https://doi.org/10.1016/j.erss.2024.103436`

VI. Vavouris, A., Garside, B., Stankovic, L., & Stankovic, V. (2022). Low-frequency non-intrusive load monitoring of electric vehicles in houses with solar generation: generalisability and transferability. Energies, 15(6), Article 2200. `https://doi.org/10.3390/en15062200`

Conference Proceedings

I. Du, D., Vavouris, A., Veisi, O., Jin, L., Stevens, G., Stankovic, L., Stankovic, V., & Boden, A. (2024). Time and money matters for sustainability: Insights on user preferences on renewable energy for electric vehicle charging stations. In A. Maedche, M. Beigl, K. Gerling, & S. Mayer (Eds.), MuC '24: Proceedings of Mensch und Computer 2024 (pp. 269-278). Association for Computing Machinery (ACM). `https://doi.org/10.1145/3670653.3670677`

II. Vavouris, A., Stankovic, L., & Stankovic, V. (2023). Integration of drivers' routines into lifecycle assessment of electric vehicles. Transportation Research Procedia, 70, 322-329. `https://doi.org/10.1016/j.trpro.2023.11.036`

III. Vavouris, A., Stankovic, L., Stankovic, V., & Shi, J. (2022). Benefits of three-phase metering for load disaggregation. In Proceedings of the 9th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (pp. 393-397). (BuildSys '22). Association for Computing Machinery (ACM). `https://doi.org/10.1145/3563357.3566149`

Datasets

I. Vavouris, A. (Creator), Stankovic, L. (Supervisor), Stankovic, V. (Supervisor), Shi, J. (Data Collector) (embargoed until publication of relevant article) Data for: "FIELD: A comprehensive FarmIng Electrical LoaD measurements dataset from 30 three-phase dairy farms in Germany". University of Strathclyde. Aggregate(.zip), Submetering(.zip), ReadMe(.txt). `https://doi.org/10.15129/1211ae7c-9b70-4a39-b3ce-318d81583749`

II. Vavouris, A. (Creator), Guasselli, F. (Contributor), Stankovic, L. (Supervisor), Stankovic, V. (Supervisor), Gram-Hanssen, K. (Supervisor), Didierjean, S. (Data Collector) (14 Oct 2024). NorPEN: A Norwegian Positive Energy Neighbourhood Dataset of Electrical Measurements and Interviews on Energy Practices. University of Strathclyde. NorPEN(.zip). `https://doi.org/10.15129/7d3ac671-2b97-439b-92cf-ce4021e804d2`

III. Vavouris, A. (Creator), Stankovic, L. (Supervisor), Stankovic, V. (Supervisor) (6 Oct 2022). Appliance Phase Identification on ECO Dataset. University of Strathclyde. AppliancePhaseInformation(.csv), ReadMe(.txt). `https://doi.org/10.15129/deddd9a7-0cff-4db2-8478-42abc93fba9f`

IV. Vavouris, A. (Creator), Stankovic, L. (Supervisor), Stankovic, V. (Supervisor) (29 Aug 2022). Smart meter electricity of a Household in Germany with Electric Vehicle Charging Annotation. University of Strathclyde. ReadMe(.txt), discoverygTesla1min(.csv). `https://doi.org/10.15129/c41a6a02-5df5-4ed7-b8e6-6488895d43f7`

1.5   Author's Contribution to Publications

Journals

I. Research, design, methodology, and testing of the proposed co-created NILM-enabled load scheduling approach. Paper writing. Supervisory input (validation & paper editing) from Lina Stankovic and Vladimir Stankovic. Referenced in Chapter 6.

II. Dataset management, processing, curation, and distribution (with data collection input from Jiufeng Shi). Write-up of curation process and background. Supervisory input (validation & paper editing) from Lina Stankovic and Vladimir Stankovic. Referenced in Chapter 6 & Appendix B.

III. Dataset management, processing, curation, and distribution (with data collection input from Sébastien Didierjean). Technical write-up of the curation process and background. Qualitative data collection, curation, and analysis from Fernanda Guasselli. Supervisory input (validation & paper editing) from Lina Stankovic, Vladimir Stankovic and Kirsten Gram-Hanssen. Referenced in Chapter 5 & Appendix A.

IV. Research, design, methodology, and testing of the mixed-methods data-driven approach (quantitative data collection support from Sébastien Didierjean). With qualitative data input from Fernanda Guasselli. Paper writing (with input in the qualitative section from Fernanda Guasselli). Supervisory input (quantitative data validation & paper editing) from Lina Stankovic and Vladimir Stankovic; (qualitative data validation & paper editing) from Kirsten Gram-Hanssen. Referenced in Chapter 5.

V. Supporting author in the quantitative analyses. Quantitative methodology, data analysis, organising (with input from Sébastien Didierjean), processing, visualising, and examining the dataset. Paper reviewing and editing. Supervisory input (quantitative data validation & paper editing) from Lina Stankovic and Vladimir Stankovic.

VI. Research, design, methodology, and testing of the EV load disaggregation using

DNN regression. Random forest (RF) classification input from Benjamin Garside. Paper writing. Supervisory input (validation & paper editing) from Lina Stankovic and Vladimir Stankovic. Referenced in Chapter 4.

Conference Proceedings

I. Supporting author, contributing to the design of the study, including background and problem formulation on granular renewable energy generation, EV charging patterns, and grid congestion. Paper editing.

II. Research, design, methodology, and testing of the proposed integration of drivers' charging routines in the usage component of LCA models of EVs. Paper writing. Supervisory input (validation & paper editing) from Lina Stankovic and Vladimir Stankovic. Referenced in Chapter 4.

III. Dataset curation (with support from Jiafeng Shi in the data collection process) and analysis, methodology, adaptation of NILM algorithms for three-phase disaggregation. Paper writing. Supervisory input (validation & paper editing) from Lina Stankovic and Vladimir Stankovic. Referenced in Chapter 3.

Chapter 2

Preliminary material

This chapter introduces background and preliminary material that are discussed in the rest of this thesis. Firstly, single and three-phase installations are introduced. Then, the electricity quantities collected through the metering infrastructure that are required for NILM approaches and recent changes in advanced smart metering specifications are introduced. Lastly, a literature review of NILM methods is included, focusing on NILM performance evaluation metrics, NILM for three-phase installations and EVs.

2.1   Single-phase & three-phase installations

The electricity distribution grid delivers power to the end-customers, usually either through a single-phase or a three-phase installation. Single-phase installations are a two-conductor alternating power circuit and are the simplest form of an electrical power system, as a single phase carries the total current. A neutral wire then completes the circuit by returning the current. Single-phase systems typically deliver power either at 220–240V or at 100–130V through a transformer. Single-phase systems are usually found in the majority of residential customers and small businesses with low power demand in the United Kingdom (UK) as they are simple, cost-effective, and easy to install and maintain. Though single-phase systems work for low energy requirements, very high consumers (such as electric HTs and EVs) or heavy-duty appliances (such as industrial equipment) cannot be supported by single-phase installations.

An approach taken to power higher energy consumers and industrial sites requires the usage of three-phase systems. A three-phase system directly connects to the three phases of the distribution system, where energy is supplied through three alternating current waveforms, each with a phase offset of 120 degrees, that ensure a continuous, balanced power flow. Three-phase systems and consumers are either connected through a three-wired, delta configuration or a four-wired, star configuration with an extra wire acting as the neutral. The line voltage, i.e., the potential difference between any two lines, is either 190 – 220V or 380 – 415V, depending on the country offering an approx. 173% increase in voltage (and thus an increase in power). The phase voltage, i.e., the potential difference between any phase and the neutral, remains the same as in the single-phase systems. Due to the increased potential difference (and thus the reduction of current for the same power transfer) between any two phases, power transmission over longer distances is more efficient, as less energy is lost compared to a single-phase system. Though three-phase systems do offer significant advantages, these are tied up with increased installation and maintenance costs compared to single-phase systems.

In summary, the selection of a single-vs a three-phase installation depends on the specific power needs of the end-user. While single-phase systems are sufficient for low-power applications (such as residential customers with reduced energy requirements), three-phase systems might be necessary, especially for larger energy consumers (commercial and/or industrial) as well as for residential customers with high load profiles, such as owners of LCTs. In fact, in multiple countries in Central and Northern Europe the majority of new installations as well as a relatively high percentage of existing installations are by default three-phased; in the UK, on the other hand, mainly due to older standards still in use, households are almost entirely powered by single-phase installations [13].

2.2 Apparent, active and reactive power

In an electrical system, complex power can be calculated as:

$$S = \frac{1}{2}\vec{V}\vec{I}^* = \tilde{V}\tilde{I}^* = P + jQ \tag{2.1}$$

where $\vec{V}$ is the peak phasor voltage, $\vec{I}$ is the current, $\tilde{V} = \vec{V}/\sqrt{2}$ is the RMS phasor voltage, P is the active power and Q is the reactive power. The angle difference between the current and voltage is given as:

$$\phi = arg\left(\vec{V}\right) - arg\left(\vec{I}\right) \tag{2.2}$$

and the power factor is defined as:

$$PF = cos(\phi). \tag{2.3}$$

Therefore the apparent power, which is measured in Volt-Ampères [VA], can be calculated as:

$$|S| = \sqrt{P^2 + Q^2}. \tag{2.4}$$

Active power, measured in Watts [W], is the actual power consumed by the electrical load connected to the source, i.e., the energy that is converted from electricity to another form of energy. It is given by:

$$P = V \times I \times cos(\phi), \tag{2.5}$$

whereas reactive power, measured in Volt-Ampères reactive [VAR], and corresponds to the portion of power that oscillates between the source and the load, is:

$$Q = V \times I \times sin(\phi). \tag{2.6}$$

The relationship between active, reactive and apparent power can be visualised in a power triangle (see Figure 2.1), where the cosine of the angle between the apparent and reactive power vectors is the power factor. Given a circuit that has a load connected, both the voltage and the current will be sinusoidal at the same frequency. If the load connected is purely resistive, i.e., the reactance is insignificant compared to its resistance, the two aforementioned quantities will reverse their polarity at the same time, and therefore the energy flow will always go from the source to the load. In this case,
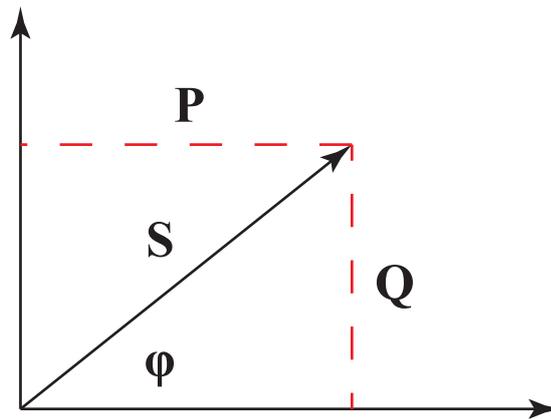
Figure 2.1: Power triangle; relationship between apparent (S), active (P), and reactive power (Q).

apparent power is equal to active power, and the power factor is 1. Contrary, if the load is purely inductive or capacitative, i.e., the resistance is insignificant compared to its reactance, the two aforementioned quantities will reverse their polarity with a 90-degree difference. This results in energy flowing for half of the circle from the source to the load and for the other half of the circle from the load to the source, and the power factor is 0. In general, a higher power factor indicates a higher energy efficiency of the system.

In summary, apparent power is the overall power required for the system to operate. Active power is the part of power that generates the required Work, and reactive power is the part of power that oscillates from the generator to the load. Although energy-efficient systems aim to minimise reactive power to reduce energy losses, multiple devices that include capacitive and inductive components require reactive power to maintain their magnetic and electric fields. Therefore, the study of the reactive power in parallel with active power is essential, especially with the introduction of energy-intensive LCTs in the grid.

2.3 Advanced metering infrastructure

The end-users' electricity metering infrastructure plays a crucial role in monitoring energy consumption across residential, commercial, and industrial settings. This infrastructure has evolved significantly over the years, driven by the need for accurate billing and

improved energy services [14]. In the past, metering infrastructure was based on analogue meters that required manual readings for billing purposes. Though these meters are reliable, they lack the ability to provide real-time data or support dynamic services. In order for governments around the world to meet the net-zero carbon targets, smart meters have been introduced [15] as a way to reduce carbon emissions from energy consumption by better managing the available energy generation and influencing on a real-time basis the demand.

In the UK, the UK government with an aim to modernise the energy sector, launched the Smart Metering Implementation Programme to replace over 53 million meters in UK homes and businesses with smart ones [5]. Although smart meters can provide consumers with greater control over their energy usage, eliminate the need for manual readings by automatically transmitting usage data to energy suppliers, and enable demand-response and energy flexibility services, there is still scepticism over the installation of smart meters, mainly due to privacy concerns.

Despite the concerns imposed by the end-users [16, 17], the smart metering infrastructure is playing a pivotal role in achieving net-zero targets. Through smart metering, time-of-use and agile tariffs are being provided to end-users that encourage energy usage during off-peak hours by dynamically changing the pricing of the energy used. Further to that, smart metering infrastructure is already playing a crucial role in the integration of LCTs such as EV charging [18] and heat pumps [19], through granular metering that supports dynamic charging /operating schedules when the grid is not congested.

In summary, the electricity metering infrastructure is undergoing a transformative shift towards smart, interconnected systems that enable the near real-time collection and processing of highly granular (1 second – 30 minutes) power consumption data [5, 20, 21]. For example, the Smart Meter Equipment Technical Specifications Version:2 (SMETS2) framework in the UK, permits regular smart meter readings to be taken at a 30-min resolution [5]. This transformation, apart from empowering end-users, enhancing grid reliability, and supporting the integration of RES, is providing an immense amount of load consumption data that can be used to extract information that can enable ancillary services and provide an improved and more energy-efficient grid operation.

2.4 Non-intrusive load monitoring

NILM is a technology that enables the extraction of individual load consumption patterns from aggregate readings without the need for individual appliance-level metering. By leveraging data from a single point of measurement, NILM aims to disaggregate the energy usage of various activities and/or devices to produce a breakdown of the total energy consumption. An analogy of the disaggregation problem is illustrated in Figure 2.2. Similarly to how buildings with distinct outlines can be identified from the skyline of a city, NILM methodology aims to identify the distinct patterns of connected appliances only from the aggregate energy consumption. NILM can either be seen as a
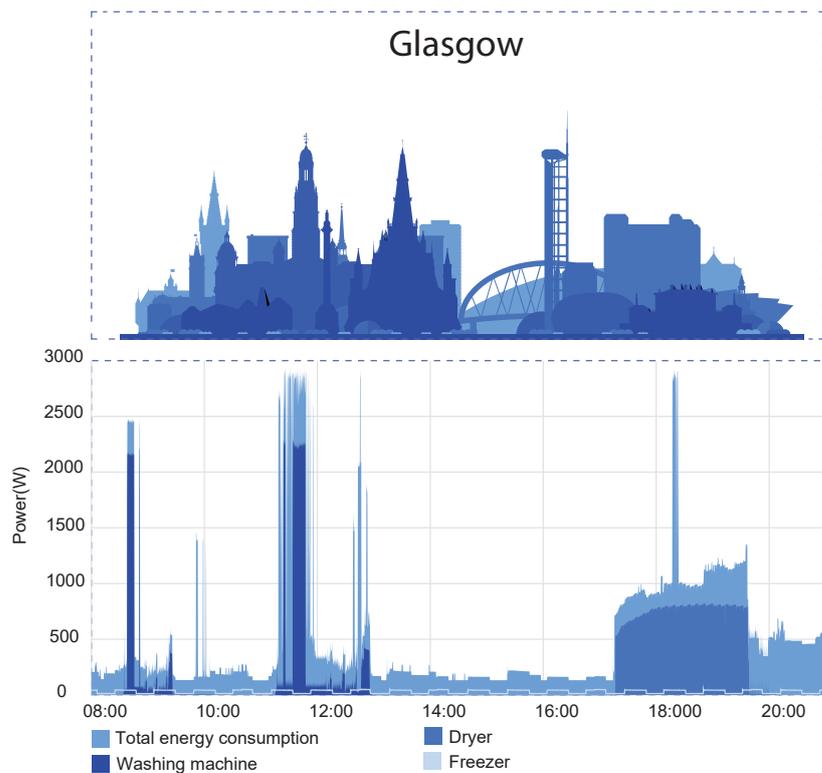
Figure 2.2: City skyline – NILM analogy.

regression problem, where the algorithm is assigning a part of the aggregated energy to a specific load/activity or as a classification problem, where the algorithm is matching the time-of-use with specific loads/activities.

Chapter 2. Preliminary material

The NILM regression problem can be formulated as in [22] as:

$$P(t) = \sum_{i=1}^{n} \alpha_i(t) P_i(t) + e(t) \qquad (2.7)$$

where $P(t)$ is the aggregate power at time t, $\alpha_i(t)$ is the status of the i-th appliance at time t, with 0 being the OFF state and 1 being the ON state, $P_i(t)$ is the power level of the i-th appliance at time t and $e(t)$ is the error term. Therefore, NILM aims to disaggregate all i appliances while reducing the remaining error term.

NILM has been an active area of study for over 30 years — with recent review papers [3, 23, 24] summarising general NILM approaches — but with the ongoing roll-out of millions of smart electricity meters globally, the deployment of large-scale residential NILM systems is emerging. The term NILM was first coined by Hart in 1992 [22], but since then, there have been significant advancements, both in the residential [2, 25] and industrial sector [26, 27]. NILM methods either involve the use of advanced signal processing techniques [28, 29] or ML / deep learning (DL) approaches [23] to analyse the aggregate electrical signal and either classify or estimate the energy consumption of the individual activities.

A typical NILM system requires a granular data stream collected either from the mains or from sub-mains distribution boards in larger applications (such as industrial settings) that contains information about the instantaneous power, energy consumption, and/or voltage and current. NILM algorithms then process the aggregate granular time series to generate the load consumption time series of individual equipment that can then be presented to end-users or used as an actuator for other services.

NILM has been used in various applications such as identification of appliances that consume the most energy such as EVs [30], enabling end-users to make informed decisions to reduce energy demand [31]; detection of anomalous energy consumption patterns [32, 33] that could indicate faulty appliances or a precursor of a fault; and as an enabler technology for DR, where users' awareness of consumption practices can lead to energy saving behaviours and reduce peak loads [34]. The minimal hardware requirements and the non-invasive nature of NILM make it a cost-effective and scalable

solution that can be deployed in multiple residential, commercial, and industrial settings with minimal hardware intervention.

Although NILM may appear a straightforward procedure with multiple benefits, still NILM faces several challenges, including accurately disaggregating appliance signals (especially for very-low data frequencies), handling the data variability (i.e., neither all installations are the same nor all similar appliances have the same load consumption signal) and lastly ensuring end-user privacy [10].

Early work in the area often assumed the availability of mid- to-high frequency power measurements in the region of 1 Hz and above, as well as current and voltage measurements. However, due to storage limitations and potential privacy concerns, current real-world smart meter readings are only available at 15 to 60 min intervals, and provide only aggregate consumed power (see Section 2.3). This motivates the need for low- (1–60s) to very low-resolution (15–60 min) NILM algorithms operating on power measurements only [35]. Recent years have seen an explosion of low-frequency NILM approaches, mostly based on DNNs. Indeed, according to [23], which provides a thorough literature review of DNN approaches for NILM, there were 87 DNN-NILM publications in the period 2018 to 2020. However, these DNN-NILM approaches focus primarily on typical household appliances, excluding EVs, and do not report results with meaningful performance metrics to truly evaluate consumption estimation. This is partly because of the limited availability of EV charging consumption datasets and generic, non-application-specific regression and classification metrics for evaluating DNN approaches for benchmarking.

Besides DNN approaches, the low-frequency NILM problem has also been tackled via other supervised and unsupervised approaches over the years, the former requiring training on labelled data, unlike the latter — an up-to-date review can also be found in [3]. Examples of supervised NILM approaches are graph signal processing (GSP) approaches [29], support vector machines (SVMs) [36], decision trees (DTs) [37] and k-nearest neighbour (kNN) [38]. Some unsupervised approaches include combinational optimisation, unsupervised GSP [35], hidden Markov models (HMMs) [39, 40] and dynamic time warping (DTW) [37]. Unsupervised methods have the advantage of

not being limited by the appliances available in the training data, but achieving good performance is challenging. Supervised approaches could equally be viable for practical large-scale deployment as long as sufficient labelled training data are available, and generalisability to similar unseen data and cross-domain transferability to other data can be demonstrated [41]. Most supervised approaches have mainly focused on NILM on seen houses and, more recently, unseen houses on the same dataset, and even fewer on cross-domain transferability [3, 23].

2.4.1 NILM performance evaluation metrics

To evaluate classification performance of NILM, *Accuracy* (Equation 2.8), *Precision* (Equation 2.9), *Recall* (Equation 2.10), and $F-score$ (Equation 2.11) metrics are used. More specifically, the metrics are given by:

$$Accuracy = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{2.8}$$

$$Precision = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2.9}$$

$$Recall = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2.10}$$

$$F-\text{score} = \frac{2 \times \text{TP}}{2 \times \text{TP} + \text{FP} + \text{FN}} \tag{2.11}$$

where TP are the true positives (TPs), TN are the true negatives (TNs), FP are the FPs and FN are the false negatives (FNs). All these metrics are calculated on a sample-by-sample basis, e.g., TP is the number of samples when a load is used that were classified as load ON. $F-score$ captures both FPs and FNs, while giving a more rigorous evaluation than *Accuracy*, particularly for datasets with a significant imbalance. For example, suppose data are split into windows for a given house and only 10% of windows are labelled as containing a specific load. A model that classifies every window as load OFF would achieve an *Accuracy* of 90% but an $F-score$ of 0%.

Typical regression metrics, mean absolute error (MAE) (Equation 2.12) and normalised signal aggregate error (SAE) (Equation 2.13) are used for evaluation of NILM

performance [23]. Additionally, the following two NILM-specific metrics are used to estimate how well the load estimation from both classification and regression learning approaches estimate the energy consumed by a specific load: Acc — sometimes referred to as total energy correctly assigned (TECA) (Equation 2.14) and match rate (MR) (Equation 2.15).

$$\text{MAE} = \frac{\sum_{d=1}^{D} |E_{est_d} - E_{true_d}|}{D}, \tag{2.12}$$

$$\text{SAE} = \frac{\left| \sum_{d=1}^{D} E_{est_d} - \sum_{d=1}^{D} E_{true_d} \right|}{\sum_{d=1}^{D} E_{true_d}}, \tag{2.13}$$

$$Acc = 1 - \frac{\sum_{d=1}^{D} |E_{est_d} - E_{true_d}|}{2 \times \sum_{d=1}^{D} E_{true_d}}, \tag{2.14}$$

$$\text{MR} = \frac{\sum_{d=1}^{D} \min\{E_{est_d}, E_{true_d}\}}{\sum_{d=1}^{D} \max\{E_{est_d}, E_{true_d}\}}, \tag{2.15}$$

where E_{est_d} and E_{true_d} refer to estimated and ground truth consumption for day d, and D is the number of days in the test dataset. MR is generally considered to be a better load estimation metric [24] for the same reason $F-$score is considered a better measure of performance than classification $Accuracy$, as it can better indicate the match between the estimated and the true energy. When compared to MAE, MR is more robust and deviates less when an experiment is repeated. It ranges from 0 to 1, with values closer to 1 indicating a strong correlation between the two, while values closer to 0 suggest a weak match. A MR of 0 occurs only when both true and estimated energy values are 0. On the other hand, the MAE metric calculates the average of errors, i.e., a large error for one subsequence or point would significantly affect the value.

2.4.2   NILM for three-phase installations

With plans for three-phase smart metering in residential settings only emerging now, NILM approaches on three-phase smart meter data are scarce for residential buildings in the literature, but different NILM approaches are proposed for industrial buildings.

Authors in [42] propose an event detection approach based on composite window analysis for three-phase industrial metering, concluding that a significant improvement in classification performance is obtained compared to detection on the cumulative-sum of phases approach.

In addition, due to the inherent complexity of measuring electricity load consumption in three-phase installations and the requirement of additional hardware, three-phase installation datasets are limited. In [23], where an in-depth review of DNNs applied to low-frequency NILM is presented, only 3 three-phase datasets were presented namely iAWE [43], ECO [4] and BLOND [44] whereas more than 10 single-phase installation datasets were presented, including the widely used: UK-DALE [45], REFIT [46], REDD [47] datasets. Of the three-phase installation datasets iAWE [43] consists of a single house in Delhi where electricity, gas, and water were monitored for 73 consecutive days, ECO [4] contains electricity measurements from 6 households in Switzerland spanning 8 months, and lastly BLOND [44] contains energy measurements from a typical office environment for 213 days.

ECO [4] dataset is a widely used NILM dataset that contains information about voltage, current, angle between phases, active and reactive power etc. of each household as well as sub-metering of several household appliances. Authors in [4] provided a comparison of the performance of four different NILM algorithms on the ECO dataset. For the experimental results, the aggregate of the three phases was used. The algorithms that were tested are Parson's, Baranski's and Voss', Weiss', and Kolter's algorithms. Parson's algorithm was used for classification and disaggregation of the energy consumption of fridges (FRDs) and microwaves (MWs), with the use of 1-minute data, achieving an $F1-score$ that ranged, for FRDs between 0.54 and 0.84, and for MWs between 0.29 and 0.14. The root mean square error (RMSE) for FRDs ranged from 17 W to 75 W. RMSE was not reported for MWs. The use of Baranski's and Voss' algorithm on 1-second data produced 11 different clusters that were then manually labelled. Devices whose electricity consumption varied between different activations, such as stoves, were spread across different clusters, as stove activations differ greatly in terms of duration and maximum power level. In addition, several kettle (KET) events

were clustered together with stove events as the load signal of a KET greatly resembles that of a very short stove activation, whereas FRDs and freezers (FRZs) were also spread across multiple clusters, but not mixed with other resistive appliances' events. The next algorithm that was used was Weiss' algorithm, with data sampled every 1 second. $F1-score$ when using the aforementioned algorithm on cooling appliances was in the range of 0.92. On the contrary, energy intensive appliances, such as dishwashers (DWs), KETs and stoves, were correctly recognised with almost no FPs — precision values ranging between 0.95 and 1.00; however, a lot of activations of these devices were missed and therefore the $F1-score$ was relatively poor — ranging between 0.25 and 0.75. Lastly, Kolter's algorithm was used to identify and cluster different appliances' signatures. However, authors in [4] claim that the cluster centroids that were produced could not be matched with the consumption patterns of the ground truth data. It was concluded that to achieve adequate results, a supervised method was required. Weiss' and Parson's algorithm appeared to perform better than the unsupervised approaches that were tested. Lastly, in [4] it is stated that unsupervised approaches require manual labelling of the outputs — i.e., manual identification and matching of the appliances.

Another approach for disaggregation of the ECO dataset was presented in [48] where authors used artificial neural networks (ANNs) to tackle the problem. Measurements were resampled to a frequency of 1/600 Hz, i.e., granularity of 10 minutes. However, although the signal of all three phases was available in the dataset, only the aggregated signal — i.e., the sum of the three phases — was used as an input to the neural networks. A multi-layer perceptron (MLP) type ANN with one hidden layer, as well as a DNN with three dense layers, were used to disaggregate appliances' signals. It was claimed that, when disaggregating ECO dataset, a DNN with the use of a rectified linear unit (ReLU) as an activation function, when compared with other — i.e. SoftPlus, SoftSign, Softmax, Sigmoid and Tanh — produced the best results. MLP-type ANNs appeared to perform poorly in the disaggregation problem, and therefore, only DNNs were further explored. An adaptive moment estimation (ADAM) optimiser and the mean squared error (MSE) were used as the loss function. Results varied greatly between the different appliances that were studied — i.e. FRD, FRZ, personal computer and washing machine

(WM), with the best performance being observed in FRDs and the worst one in personal computers.

Authors in [49, 50] proposed the use of DNNs to disaggregate the electric signal of several datasets. A 4-layered bidirectional long short-term memory (LSTM) model was trained for each target appliance, and some post-processing steps were implemented to increase the accuracy of the proposed methodology. ECO dataset was also used for evaluating the performance of the proposed methodology, where all phases were summed, and the aggregate signal was used as an input to the network. It was concluded that devices such as DWs and WMs were hard to classify due to different signatures that were a result of the different appliances' cycles. SAE ranged from 1.3% up to 64.9%, with DWs having an SAE of 28.8%, FRDs 12.1%, WMs 64.1% and MWs 43.7%.

In summary, the NILM problem on three-phase installations has not been extensively studied in the literature, mainly due to the complexity of monitoring three-phase loads and the lack of publicly accessible open datasets. With the introduction of high-powered, three-phase loads in the residential sector (such as EVs and heat pumps) and the subsequent upgrade of current single-phase infrastructure to three-phase, it is expected that NILM applications for three-phase installations will increase.

2.4.3   NILM for EV load disaggregation

While NILM models have been developed for disaggregation of most conventional household appliances, NILM for the disaggregation of EV loads is still an emerging area of study. Although at first glance, EV load disaggregation may seem a relatively simple problem due to its high power level and being a single state load, houses nowadays use many electric devices with complex electrical signals and high energy consumption that make the separation of the EV signal a challenge. These include households with electric heaters, heat pumps, electric showers, air conditioning (AC) and heating, ventilation and air conditioning (HVAC) units, or prosumers, i.e., consumers that also produce electricity through solar panels and/or other RES — tend to have quite complex load signals. Load disaggregation of EV charging is first tackled in [51], where a training-free approach, based on time-series signal thresholding, filtering and denoising, is proposed

that uses knowledge of known appliance signatures to remove contributions from other loads and estimate power consumption of EVs. The approach is validated with over a year of 1 min data from Dataport [52] between 2012 and 2013, across 11 houses, randomly picked out of hundreds of houses fromthe Austin area. Monthly consumption error and MSE were used to evaluate the performance of the method, and results were benchmarked against the HMM algorithm of [39]. Results outperformed HMM, which had difficulty distinguishing between EV loads and AC "spike trains", which becomes particularly challenging in the summer months. However, the calculation of error in terms of monthly consumption is not as rigorous as the *Acc* metric and MR that have emerged more recently and are often calculated based on daily consumption estimates [24, 35]. The authors of [51] do not provide the IDs of the households that were used, and therefore, the results cannot be reproduced and compared.

Another unsupervised approach is proposed in [53], where independent component analysis (ICA) is used to extract EV loads from aggregated signals. This is followed by a series of complex processing steps to remove interference from appliances with similar load characteristics and rebuild an estimated EV load profile. Validation of the method is carried out on 1-minute Dataport [52] from 34 houses, and on 5-minute resolution samples, obtained by resampling the measured 1-minute readings. Results were evaluated using EV load reconstruction error, calculated sample-by-sample and a modified $F-$score that takes Accurate/Inaccurate TPs into account as used in [28]. However, when tested on 5 min resolution data, performance was significantly reduced. As with [51], the authors of [53] do not provide the IDs of the households that were used, hence the results cannot be reproduced.

Apart from unsupervised approaches, supervised approaches for the EV load disaggregation problem have been proposed. In [38], a mean sliding window algorithm is used to detect and extract features from ON/OFF events, i.e., an appliance switched ON and OFF — which are subsequently classified as AC and EV charging, using a kNN classifier. The method is validated on 1 min data collected by Dataport [52], from June to August 2014. The classifier was trained on 15 days of data collected at house 26 and tested on 4 days from the same house achieving $F-$scores of 83% and 91% for EV charging

ON and OFF events, but $F-$scores fell to 86% and 75%, respectively, for 5 min data. Generalisability to unseen house 3036 in the Dataport dataset was attempted using a pre-trained model, but optimal k-values were chosen based on misclassification error rate for each house individually; this requires labelled data for both houses and therefore fails to fully test generalisability to house 3036. Although classification results are promising, a testing period of only 8 days from 2 houses is inadequate to fully evaluate the effectiveness of the method. It is also unclear how the test days and houses were chosen. No energy consumption estimations were calculated from the classifications, and hence no consumption-based metrics were used for evaluation.

Another low-complexity supervised method is proposed in [54] where active power data are split into overlapping windows that are fed into an RF classifier. Principal component analysis (PCA) is used for feature extraction. Once again, the method is validated on 1 min data from Dataport [52] — 6 houses were considered over the period January 2016 to December 2017. The data for each house were split into 10-minute overlapping windows and used directly as input to the RF classifier, achieving an $F-$score of 92.61%. In [54], PCA is applied to the windows and all 10 principal components (PCs) are used, resulting in a reported improvement in classification performance. However, the $F-$score was only changed by 0.08%, which is far from a significant increase. The authors discuss the use of PCA for removing redundant information and show that over 95% of the variance in their dataset is explained by 2 PCs, but no attempt is made to reduce dimensionality. It is also unclear whether PCA was applied to the train and test datasets separately, which is important for ensuring that no bias is imparted on the training data through implicit knowledge of the test data. The application of PCA to the windows resulted in a small reduction in FNs, with marginally improved $F-$score of 92.69%. According to [54], this outperforms the ICA unsupervised approach of [53]. However, the direct comparison with [54] is hard to make for two reasons. Firstly, the RF classifier is given a balanced dataset for testing, i.e., 50% of windows contain EV charging and 50% only contain other household appliances, which is achieved by random under-sampling. This does not represent the real proportion of EV charging vs. non-charging windows, which is reported to be 6% in the initial, unbalanced dataset. As

a result, it is not demonstrated how robust the classifier is against FPs that may arise from interference from other large loads. Secondly, labelled data from all 6 houses were used for training, and therefore, generalisability to unseen houses is not demonstrated.

Building upon [54], in [55] RF is evaluated alongside kNN and ANN approaches for EV load disaggregation. Models are trained and tested with a selection of 18 houses for a month from Pecan Dataport [52] — however, which houses are used exactly for training and testing is not specified. Therefore, the results cannot be reproduced or compared. In the pre-processing step, since only one month was considered with insufficient EV charging events, the authors simulated additional EV load charging patterns instead of using real data from other days. Training and testing sets were created by selecting only one month of data from selected houses that included both EVs and photovoltaics (PVs). Generic classification and regression metrics are presented, without taking into account NILM-specific metrics such as MR or Acc [24]. The RF model outperformed the other two models, with results presented only for two set-ups — classification ($F-$score $= 93\%$ and 75% for training and testing on a selection of houses and testing on one unseen house, respectively) and regression (MAE $= 500$ W and 630 W for training and testing on a selection of houses and testing on one unseen house, respectively).

Compared to the general NILM problem (for which multiple popular electrical measurements' datasets on which NILM approaches are generally validated exist: REDD [47], UK-DALE [45], REFIT [56] and Pecan Street Dataport [52] — see [3] for some other examples of commonly used datasets), EV specific datasets are scarce. From the aforementioned datasets, only Dataport includes EV sub-metering and aggregate meter readings for multiple houses for a few months. A thorough review of available EV load datasets, including charging point locations, historical and real-time charging sessions, which refer to the period an EV is charged, traffic counts, travel surveys and registered vehicles, is presented in [57] to improve EV load modelling. However, none of the vehicle-centric data contains actual consumption readings from charging points, but rather spatial and temporal EV charging sessions to artificially reconstruct synthetic house-level and aggregated load consumption. This is not used in this study since synthetic loads do not reflect true consumption from the grid, and are not integrated

into the household overall mains metering with other interfering loads and prosumers.

In summary, there is limited previous work on EV classification and load consumption estimation using a range of signal information processing methods. The main gaps are the lack of transparency in reporting sufficient details — such as specifics and number of houses and days used for training and testing — for reproducing and comparing results, lack of transparency in the choice of experimental data — including quality and quantity metrics. Furthermore, performance evaluation in current literature tends to be non-rigorous, especially on generalisability and cross-domain transferability of the methods, which is needed for practical deployment.

Chapter 3

Load disaggregation of three-phase residential settings

The content of this chapter has been published in: Vavouris, A., Stankovic, L., Stankovic, V., & Shi, J. (2022, November). Benefits of three-phase metering for load disaggregation. In Proceedings of the 9th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (pp. 393-397).

3.1   Introduction & background

With the ever-increasing pace of introduction of energy-intensive devices and services, such as EV charging [58] and self-generation from renewables — e.g., solar panels — the transition to smart metering for three-phase electric installations for nationwide smart meter roll-outs is underway. Three-phase installations, mostly predominant in commercial and industrial installations, are emerging across parts of the world — e.g. in the UK [13] — to support the surge in end-user demand as well as households and devices that require three-phase installations. With an aim to achieve a reduction of 80% of greenhouse gas (GHG) emissions in the UK — as it is outlined in the Climate Change Act [59] — industry is adapting to meet the new governments' regulations that focus on the reduction of the carbon footprint [13]. The imminent ban of new fossil-fuelled means of transportation, the introduction of EVs, as well as the usage of

other energy-intensive appliances such as heat pumps [60] instead of petrol and gas, are measures taken in order to reduce greenhouse emissions.

Although the electrification of different aspects of daily life in combination with energy provided from RES will reduce the emissions, current electric installations [13] are not able to cope with the exponential growth in energy demand. One-phase EV chargers have varying power levels in the range of $1 - 8kW$ with 3-pin charging cables usually providing 2.3 kW and therefore making the charging of an average battery a lengthy procedure, that can last several hours. On the contrary, with the use of residential three-phase EV chargers, which are able to supply vehicles with a power level of $10 - 30kW$, it is possible to greatly minimise the EV charging window and therefore allow users to fast charge their car. According to preliminary results from the Electric Nation project [61], a staggering 87% of commuters prefer to charge their vehicle when at home, whereas only 8% prefer charging at work, 4% at service stations and 1% in other locations, such as shopping centres. Therefore, fast, three-phase EV chargers installed in end users' houses would provide the benefits of fast and reliable charging.

Furthermore, heating in houses and industry creates around 32% of the total yearly greenhouse emissions in the UK [62]. A strategy to reduce building emissions is the reduction of the emissions that are connected with buildings' heating. This can be achieved, as with transportation, by the electrification of heating, i.e. with the use of heat pumps. These devices, which can work using renewable energy instead of commonly used fossil fuels such as petrol and gas, usually require high nominal power levels that cannot be supplied by the common and widely used single-phase installations. Therefore, as electrification of heating is being rolled out widely, new installations that can support that surge in demand need to be implemented.

In addition, the introduction of three-phase installations in the new buildings, as well as the upgrade of existing single-phase buildings' installations to three-phase ones, will enable the installation of larger residential solar PV systems, when compared to current ones that are capped at a certain maximum power due to restrictions imposed by the maximum load that single-phase installations' can handle. According to the rules imposed in the export of energy from end-users, as they are underlined in the

Engineering Recommendation G83 [63], single-phase installations can export up to 3.68 kW AC of solar power per phase, without the requirement of a special permission from the Distribution Network Operator (DNO). This greatly limits the maximum amount of renewable solar energy that can currently be produced per household. With the introduction of three-phase installations, it is possible to export three times the power, without the need for any special permission and therefore greatly reducing the amount of time and sources that would otherwise be required to install the same solar PV power.

Besides, as it is highlighted in [13], the availability of more power at any given time to an end-user can further increase the ability of time-shifting of home charging or of energy-intensive loads. In other words, as end-users' installations will have the ability to supply more power simultaneously, when compared to single-phase installations, restrictions imposed by that limited energy supply per unit of time are lifted and therefore load shifting can be implemented more easily and efficiently. Therefore, users will be able to move some loads in different time windows, without having to worry about other energy-intensive appliances that may run at the same time.

Apart from the advantages for the end user, three-phase installations can be greatly beneficial for the stability of the grid. As it was highlighted in [13] current single-phase installations, where houses are connected to the grid on a rotational basis across the different phases, with each house being connected only on a single phase, can pose a great threat to the grid, as increased energy consumption may lead to imbalances between the phases on the local transformer. These imbalances, combined with high currents that may be drawn from a single phase, when one or more energy-intensive appliances are used, may lead to overheating of the grids' cables and therefore to greater losses for the network.

Compared to common one-phase installations, three-phase installations are able to deliver the same amount of power using only one-third of the required current per phase and subsequently reduce the conductors' diameter for AC wiring. Three-phase installations are also of paramount importance for the grid, as they can mitigate current balancing issues that would arise if all the current of an energy-intensive appliance were drawn from a single phase. Hence, it is recommended that loads are balanced across

the different phases, as otherwise neutral shifting may occur. The fact that three-phase installations are being deployed will lead to another collateral effect, that of distributing loads of energy-intensive appliances in different phases and therefore reducing the noise due to other appliances on the load signal of interest per phase.

Recent standardisation for three-phase smart metering in UK and Europe (see Section 2.3) are making it obligatory for manufacturers to produce metering devices that are capable of measuring and transmitting the power, the voltage, the current, and the angle between the different phases of a three-phase installation. This provides an opportunity to potentially improve disaggregation accuracy of NILM algorithms by exploiting the load distribution across the three phases.

In this chapter, a detailed and robust methodology for the evaluation of load disaggregation of energy-intensive appliances from three-phase household installations is presented. The main contributions of this chapter are:

- adapting a seq2subseq [64] DL-based NILM algorithm from [65], providing full details of the proposed pre-processing, hyperparameter tuning and post-processing steps for different appliances;

- quantifying gain in disaggregation accuracy when using per-phase and aggregate signal, taking into account noisiness and sparsity metrics, and different data granularities;

- proposing a method of appliance phase identification and releasing information regarding the phase on which each appliance of the ECO dataset [4] is connected to [1];

- labelling the EV usage of a three-phased household in Germany for a period of 1 year via transfer learning with manual verification, and releasing the labelled dataset[2]

---

[1]Appliance phase identification data for the ECO dataset can be accessed at: `https://doi.org/10.15129/deddd9a7-0cff-4db2-8478-42abc93fba9f`

[2]The research data supporting this study can be accessed at: `https://doi.org/10.15129/c41a6a02-5df5-4ed7-b8e6-6488895d43f7`

3.2   Methodology

Following the recent NILM review papers [2, 3, 23, 24, 66], the DL seq2subseq NILM approach of [64] is adapted, shortlisted in [23] as one of the best performing on standard household appliances and demonstrated on the PECAN [52] dataset in [67].

3.2.1   Data selection & preparation

As it was already discussed in Subsection 2.4.3, publicly available three-phase installations datasets are scarce, and only the ECO [4] dataset had residential data for over 6 months and for 6 houses. A wide variety of different appliances are available in ECO dataset, including washing and cooking activities as well as lighting and entertainment. A summary of the appliances that are available in ECO dataset is presented in Table 3.1.

This chapter focuses on energy-intensive appliances, including DWs, dryers, FRZs, FRDs, and WMs, as well as appliances used on a daily basis, such as CMs and MWs. ECO Houses 3 and 6 were discarded as there were long periods of missing readings. From the remaining houses, the following devices with the respective duration were taken into account:

- House 1: CM: 113 days, dryer: 231 days, FRZ: 231 days, FRD: 231 days and WM: 231 days;
- House 2: DW: 240 days, FRZ: 240 days and FRD: 240 days;
- House 4: FRZ: 192 days, FRD: 194 days and MW: 194 days;
- House 5: CM: 218 days and FRD: 218 days.

In houses 4 and 5, a second FRD and a second FRZ, respectively, were installed, which were not monitored using a separate smart plug. Therefore, disaggregation accuracy of the refrigeration appliances is expected to be affected as the signatures of these two devices are expected to be highly correlated with the respective monitored ones.

Furthermore, certain devices were discarded as either containing more than one device — e.g., kitchen appliances in house 4 — or containing not enough data/activations — e.g.,

Table 3.1: Summary of ECO Dataset

Appliance/House	1	2	3	4	5	6
Air exhaust		x				
Coffee machine (CM)	x		x		x	x
DW		x				
Dryer	x					
Entertainment		x	x	x	x	x
Fountain					x	
FRZ	x	x	x	x		
FRD	x	x	x	x	x	x
KET	x	x	x		x	x
Kitchen appliances				x		
Lamp		x		x		x
Laptop		x				x
MW				x	x	
PC	x		x		x	
Router						x
Stereo		x		x		
Stove		x				
Tablet		x	x	x	x	
TV		x				
WM	x					

stove in house 2 and MW in house 5. Lastly, certain appliances that were monitored, such as tablets, routers and laptops, were not further investigated, as these devices usually correspond only to a very low part of the daily energy consumption.

The original dataset had a sampling frequency of 1 Hz, i.e., mains meter readings as well as smart plugs were transmitting values every 1 second. However, for our research purposes, data were resampled to granularities of 10, 30 and 60 seconds to investigate the disaggregation accuracy of the seq2subseq algorithm when using data of different granularities in the low- and very-low-frequency domain. The decision to use data of

lower sampling frequency was made according to recent protocols (see Chapter 2.3), which take into account data collection and storage limitations as well as end-users' privacy issues; less granular data should be used. As the dataset contained missing data, missing values for short periods of time were linearly filled, whereas longer periods of missing data, either in the mains or in the ground-truth data, were discarded.

3.2.2 Sparsity & noisiness of the dataset

A fundamental factor that affects the disaggregation accuracy of the NILM algorithms is the amount of available activations of each appliance used to train the models. In order to measure this aspect, the sparsity of each appliance in a dataset is calculated using Equation 3.1 as:

$$S = \frac{T_{\text{On}}}{T_{\text{Total}}}, \tag{3.1}$$

where T_{On} is the duration that the appliance is on and T_{Total} is the total duration of the dataset.

As it was demonstrated in [67], noisiness in a dataset, measured by the noisiness metric (NM) [1], is positively correlated with the disaggregation performance. NM can be given by the Equation 3.2 as:

$$\% - NM^{(T)} = \frac{\sum_{t=1}^{T} \left| y_t - \sum_{m=1}^{M} y_t^{(m)} \right|}{\sum_{t=1}^{T} y_t}, \tag{3.2}$$

where T is the total monitoring duration — in the number of samples — y_t is the aggregated load measured at time sample t and $y_t^{(m)}$ is the submetered measurement of load/appliance m at time sample t. M denotes the number of appliances that are disaggregated. As it was discussed in [67], when using multiple binary classifiers, it is appropriate to apply the NM for $M = 1$ where the only appliance of interest is the target appliance that is to be disaggregated.

In Table 3.2, the sparsity and the NMs for each appliance, both for per-phase and aggregated signal, is presented, where all other loads contributing to the aggregate are considered as noise. NM on the aggregated signal — i.e., NM_{Agg} which is calculated using Equation 3.2 where y_t is the aggregate signal — is expected to be higher than the

equivalent of the phase — i.e., NM_ϕ which is calculated using Equation 3.2 where y_t is the signal of the phase that the appliance is connected to — due to the fact that the total load present in the household is spread in three different phases. This reduction in the NM on a per-phase basis is expected to increase the disaggregation accuracy of the seq2subseq algorithm.

Table 3.2: Sparsity and NM of ECO dataset. ϕ denotes the phase that the appliance is drawing current from. It can be seen that the noisiness measure, as expected, decreases if phase aggregates are considered.

ID	Appliance	ϕ	S	NM_Agg	NM_ϕ
1	CM	2	0.64%	98.56%	96.18%
	Dryer	3	3.25%	91.98%	56.30%
	FRZ	1	54.45%	93.54%	86.32%
	FRD	2	36.92%	92.66%	78.76%
	WM	1	6.70%	91.79%	82.64%
2	DW	1	1.43%	92.73%	87.78%
	FRZ	1	50.16%	87.36%	78.80%
	FRD	1	34.09%	88.63%	80.92%
4	FRZ	1	28.90%	96.76%	94.64%
	FRD	1	84.83%	79.80%	66.59%
	MW	1	1.39%	98.12%	96.90%
5	CM	3	1.44%	99.32%	98.48%
	FRD	3	35.70%	94.39%	87.37%

3.2.3 Regression Based on DNN

Seq2subseq learning is adapted and optimised, with a conditional generative adversarial network (GAN), using publicly available code [65], to disaggregate common residential loads. Seq2subseq network targets the middle part of a sequence, and therefore a shorter sequence compared to sequence-to-sequence (seq2seq) DNN, making convergence faster. Additionally, since the network targets a subsequence instead of a point, as is the case with the sequence-to-point (seq2point) DNN architectures [68,69], training is faster and less computationally expensive.

3.2.3.1   Window size selection

For the disaggregation algorithm to work successfully, the whole appliance event must be included in the targeted subsequence of the DNN. Thus, the optimal seq2subseq window size ω is set by Equation 3.3 as:

$$\omega \gtrsim \frac{2 \times L}{g}, \omega \in \{2^0, 2^1, ..., 2^n, ...\}, \tag{3.3}$$

where L [in seconds] is the usual length of the appliance cycle period, and g [in seconds] is the resolution of the data samples. Equation 3.3 is proposed based on the following criteria: the subsequence has a width that is equal to half of the window size, and the window size must be a power of 2 — a requirement of the DNN architecture. Note that different window sizes will be used on different granularities. A summary of the used window sizes for the different appliances is presented in Table 3.3. The used window lengths range from 64 to 2048 samples. It is expected that a carefully selected window size based on the characteristics of the appliances' signals can greatly increase the disaggregation accuracy of the seq2subseq algorithm.

Table 3.3: Seq2subseq window size selection [in number of samples]

Appliance	10 sec	30 sec	1 min
CM	256	128	64
DW	2048	512	256
Dryer	2048	512	256
FRZ	256	128	64
FRD	256	128	64
MW	256	128	64
WM	1024	512	128

3.2.3.2   Seq2subseq hyperparameters

The remaining neural network's hyperparameters were chosen based on the performance of the neural network on the validation set. The L1 (least absolute deviation) loss was

used in all setups. The optimisers chosen for the discriminator and generator filters were the stochastic gradient descent (SGD) and the ADAM optimiser, respectively. The initial learning rate for the SGD was 0.001, whereas for the ADAM optimiser, an initial learning rate of 0.0005 was chosen. In addition, the momentum term used for the ADAM optimiser was equal to 0.5, whereas the weights on L1 and GAN term for the generator gradient were 100 and 1, respectively.

For window sizes between 256 and 2048, a total of 7 layers were used. For a window size of 128, 6 layers were used, and finally, for a window size of 64, 5 layers were used. The number of generator and discriminator filters in the first convolutional layer for all the setups was calculated by dividing the window size by 4, and therefore, the number of filters ranged from 16 to 512. With the use of the early stopping criterion on the validation set, the number of epochs was chosen to be 120.

3.2.3.3 Post-processing procedure

Given the produced sub-sequences of residential loads, a simple correction procedure is applied. The seq2subseq algorithm produces either some very small negative values — indicating that the appliance is feeding power to the network — or sometimes produces values that are higher than the aggregate consumption — which is impossible as the aggregate power is always greater than or equal to the power of each appliance. As these consumption values are erroneous, a simple post-processing step was performed. All negative consumption values were replaced by zero values, and all the power consumption values that were greater than the aggregate were replaced by the maximum aggregated power. As it was expected, this post-processing step increased the accuracy of the disaggregation in the range of 0% to 2%.

3.2.4 Measuring improvement in accuracy and granularity loss

As the initial hypothesis of the research was that the usage of the signal only from the phase that the appliance is connected to, instead of the aggregate, would improve the disaggregation performance, a measurement of the gain in disaggregation was introduced. The improvement in accuracy, when using the signal only from the phase where the

44

appliance is connected, is given by Equation 3.4 as:

$$G_{\text{phase}} = \frac{Acc_{\text{phase}}}{Acc_{\text{aggregate}}} - 1, \tag{3.4}$$

Furthermore, the loss introduced by data of lower sampling rate was used in order to correlate the disaggregation loss with the use of less granular data. This loss was calculated using the Equation 3.5 and is calculated as:

$$Loss = \frac{Acc_i}{Acc_j} - 1, \tag{3.5}$$

where Acc_i and Acc_j are the accuracy metrics obtained by Equation 2.14, when performing load disaggregation using lower and higher frequency data, respectively.

3.2.5 Determining phase

Although in ECO dataset, ground truth data exist, there is no indication regarding the phase each appliance is connected to. Therefore, this information should be extracted per appliance and per household. This was performed by feeding the seq2subseq DNN all three phases per appliance. The hypothesis that was made is that considering the demonstrated accuracy of seq2subseq algorithm for different appliances [64, 67], the performance of the algorithm when measured using the accuracy metric given by the Equation 2.14 on the single phase on which each appliance is connected will be high whereas on the other two phases, where the load is not present, the algorithm will produce an empty consumption vector. During the experimental results, this hypothesis was confirmed for all the appliances that were being studied. In addition, a further manual cross-check was made by verifying the presence of each appliance's pattern on the corresponding phase. The phases to which each appliance is connected are presented in the summary Table 3.4 and denoted with letter ϕ.

3.2.6 Phase identification & load labelling

In addition to ECO dataset, three-phase smart meter readings from an unseen German household were used in order to demonstrate the value of disaggregating per phase in

Table 3.4: Seq2subseq performance measured using accuracy metric (Equation 2.14).

ID	Appliance	ϕ	Aggregated phases			Appliance phase			Gain (G_{phase})		
			10 sec	30 sec	1 min	10 sec	30 sec	1 min	10 sec	30 sec	1 min
1	CM	2	75.15%	58.54%	–	79.77%	71.97%	59.81%	6.15%	22.94%	–
	Dryer	3	40.73%	36.47%	16.82%	79.27%	78.06%	75.49%	94.62%	114.04%	348.81%
	FRZ	1	85.32%	83.59%	81.28%	93.53%	92.69%	90.69%	9.62%	10.89%	11.58%
	FRD	2	69.52%	69.05%	68.00%	83.79%	82.06%	80.51%	20.53%	18.84%	18.40%
	WM	1	80.33%	67.70%	47.12%	88.02%	83.30%	72.23%	9.57%	23.04%	53.29%
2	DW	1	62.70%	43.22%	81.56%	67.17%	60.27%	87.80%	7.13%	39.45%	7.65%
	FRZ	1	91.09%	87.89%	83.14%	92.05%	89.64%	85.32%	1.05%	1.99%	2.62%
	FRD	1	85.48%	82.49%	76.76%	87.32%	84.72%	79.58%	2.15%	2.70%	3.67%
4	FRZ	1	81.80%	80.05%	80.60%	85.32%	83.61%	83.98%	4.30%	4.45%	4.19%
	FRD	1	50.12%	47.73%	46.98%	54.70%	54.98%	54.19%	9.14%	15.19%	15.35%
	MW	1	64.88%	60.45%	58.13%	74.28%	74.22%	69.70%	14.49%	22.78%	19.90%
5	CM	3	74.56%	66.41%	54.52%	83.48%	79.67%	72.57%	11.96%	19.97%	33.11%
	FRD	3	75.10%	74.54%	70.20%	89.42%	89.31%	87.13%	19.07%	19.81%	24.12%

dataset labelling. Data from this household spanned one year from the 1st of January 2021 up to the 31st of December 2021. An EV was installed in this household, which used a three-phase symmetrical load charger, i.e., simultaneous and equal loading across all phases during charging. In the absence of submetering data for the EV, the dataset was labelled via transfer learning, i.e., training the seq2subseq algorithm as per [67] with the one-phase PECAN Dataport dataset [52], and disaggregating the load per phase. The training houses from PECAN Dataport were chosen such that they had a similar charging load profile with a 3kW load as one phase of the unseen German household.

Elimination of FPs was made possible by exploiting the fact that the load was completely symmetrical across all 3 phases, i.e., FP activations that were observed only in one or two phases were discarded as they corresponded to other one-or two-phase energy intensive appliances in the house, such as WMs, DWs, resistive kitchen appliances and heating devices, which had similar duration as EV charging and drawing loads of 2.2 – 2.5kW.

The disaggregation results were then manually validated. The entire period of the dataset was manually inspected, and symmetrical three-phase loads' start and stop times were annotated. The charger's signal was completely symmetrical across all phases, i.e.,

the same signal was available three times, starting and stopping at the same timestamp, distorted by different noise — i.e., the rest of the appliances that were connected on each of the three phases. The recovery of the signal was performed taking into account the symmetry and the load profile. Manual estimation of the EV's load was then used to calculate an approximate estimation of the disaggregation accuracy of the seq2subseq algorithm.

3.3 Results

3.3.1 The ECO Dataset Case-Study

Results obtained on the ECO dataset are summarised in Table 3.4. *Acc* metric is presented for data sampled at 10 seconds, 30 seconds and 1 minute when using the aggregate of the three phases as well as when using only the phase where the appliance is connected to. The improvement in *Acc* metric when using the signal from the appliance phase was calculated using Equation 3.4. In house 1, the CM signal consisted of very short pulses, mostly in the range of 30 seconds to 1 minute and therefore the seq2subseq algorithm was unable to disaggregate the signal when using a sampling rate of 1 minute.

As expected, due to the reduced noise in the signal — calculated using Equation 3.2 and summarised in Table 3.2 — experimental results demonstrate that per-phase accuracy is in general higher when compared to the accuracy obtained when disaggregating the sum of the three-phase readings. However, by manually observing the algorithm's output, it was observed that, apart from the reduced noisiness of the signal, the important aspect that influences the performance is the similarity of the loads connected to the same phase.

In house 1, disaggregation accuracy of the dryer, when using only the appliance phase, is greatly increased, especially for less granular data. This is partially due to the reduced noise in the signal of the specific phase when compared to the aggregate over all three phases. Secondly, the majority of the dryer's activations, also energy intensive with a multi-state load profile, were not mistaken for another appliance with a similar load profile since the dryer was connected to another phase. Therefore, the dryer's signal

could not be distinguished from the aggregate of the three-phase readings, whereas on the appliance phase, the similar load was absent, and the seq2subseq algorithm was able to accurately disaggregate the dryer's load.

The seq2subseq algorithm on refrigerating appliances, i.e., FRDs and FRZs, displays similar performance across different houses, with results obtained for FRZs being better than FRDs, a fact that can be attributed to the higher energy profile of the FRZs, as well as to their more constant current draw when compared to the FRDs that are used more often — opening/closing the door — and therefore energy consumption pattern varies to a greater extent. An exception to the disaggregation of the load pattern of the FRDs exists in house 4, where, despite the *NM* indicating the signal is less noisy than other houses, the disaggregation performance appears to be poor. As mentioned earlier in Section 3.2, in house 4, a second FRD was present that was not monitored with a smart plug, since the device was in a basement and a connection to the router could not be established [4]. By manually inspecting the signal, it was observed that the second FRD, which was not monitored, was also connected to the same phase and had an almost identical signature. Therefore, the algorithm was also detecting the second FRD that was installed in the house, which led to a large decrease in disaggregation accuracy vs submetering of one FRD. It is also worth noting that although FRDs and FRZs present a very similar consumption pattern, the seq2subseq algorithm was able, to a great extent, to discriminate between the two devices and therefore the FPs were minimal. This can be attributed to the small difference in the peak energy levels of the two devices, as well as to the small differences in the duration of the activations of the two appliances.

Table 3.5 presents a summary of the loss in accuracy, both for the aggregate and the appliance phase signal, obtained using Equation 3.5 when using 30-second and 1-minute instead of 10-second data, that is, Acc_j in Equation 3.5 corresponds to the results with 10-sec granularity and Acc_i to either 30-seconds or 1-minute granularity. In general, there is a positive correlation between the loss in accuracy and the sampling frequency. Disaggregation accuracy of refrigerating appliances — i.e., FRDs and FRZs — in all houses, except house 2, when using the appliance phase signal, is almost invariant

w.r.t. granularity levels. This is expected as refrigerating appliances tend to have a constant periodic signal. Furthermore, a significant decrease in disaggregation accuracy of the dryer and the WM in house 1 when using lower frequency data is observed. This deterioration is especially visible when using the aggregate of three phases, as the combination of higher levels of noise and the reduction in the granularity of the data led to about 50% decrease in disaggregation accuracy. Therefore, it can be concluded that, in general, disaggregation of refrigerating appliances can be highly effective when using data with granularities in the range of 1 minute, whereas devices with sparse activations and multi-state load profiles tend to require a higher sampling frequency to achieve the same performance. Also, our experiments demonstrate that another advantage of disaggregating meter readings per-phase is the ability to better disaggregate appliances with sparse activations and variable multi-state current draws with less granular data, as is the case with the dryer in house 1.

Table 3.5: Seq2subseq granularity loss compared to disaggregation accuracy when using 10-second data. Negative values indicate drop in accuracy.

ID	Appliance	Aggregated phases		Appliance phase	
		30 sec	1 min	30 sec	1 min
1	CM	–22.10%	–	–9.78%	–25.02%
	Dryer	–10.46%	–58.70%	–1.53%	–4.77%
	FRZ	–2.03%	–4.74%	–0.90%	–3.04%
	FRD	–0.68%	–2.19%	–2.06%	–3.91%
	WM	–15.72%	–41.34%	–5.36%	–17.94%
2	DW	–31.07%	30.08%	–10.27%	30.71%
	FRZ	–3.51%	–8.73%	–2.62%	–7.31%
	FRD	–3.50%	–10.20%	–2.98%	–8.86%
4	FRZ	–2.14%	–1.47%	–2.00%	–1.57%
	FRD	–4.77%	–6.26%	0.51%	–0.93%
	MW	–6.83%	–10.40%	–0.08%	–6.17%
5	CM	–10.93%	–26.88%	–4.56%	–13.07%
	FRD	–0.75%	–6.52%	–0.12%	–2.56%

On the other hand, as it can be observed from both Tables 3.4 and 3.5, disaggregation accuracy of DW does not follow the same pattern as other appliances. More specifically, accuracy, when using either the phase aggregate signal or only the measured single phase signal, shows the improvement when using 1-minute data w.r.t. 10 10-second data. The refrigerator signal consists of one or more high-energy pulses with lower energy levels around the main pulses. Also, there were some activations with low energy — probably corresponding to rinsing cycles — that, when using a lower sampling frequency, that energy was spread across different samples, and therefore it was more difficult to be distinguished. The seq2subseq algorithm was unable to disaggregate the pattern of the energy consumption outside the main pulses when using data of higher granularity, as well as these low-energy uses. On the contrary, when using 1-minute data, as these low-powered level signals were aggregated in a per-minute manner, the aforementioned algorithm was able to better assign that energy to the DW usage.

Lastly, disaggregation results for both CMs in houses 1 and 5, as well as MW in house 4, indicate that a higher sampling frequency can greatly increase accuracy for appliances with sparse loads (see Table 3.2) and short activations.

3.3.2 Labelling EV Usage

As described in Subsection 3.2.6, PECAN Dataport Dataset [52] was used to train the algorithm, via transfer learning, to label the usage of an EV charger on an unseen three-phase metered dataset without submetering. Load was disaggregated per phase, and using the post-processing technique described in Subsection 3.2.6 FPs were eliminated. Disaggregation results after the post-processing, and using the manually annotated EV activations as ground truth, are presented in Figure 3.1. In Table 3.6 estimated accuracy results as given by Equation 2.14, both for 1-minute and 15-minute data, are presented. The estimation of the EV's charger signal as described in Subsection 3.2.6 was used to juxtapose the signal obtained from the seq2subseq algorithm. This example of labelling a dataset where ground truth data are absent, underlines the importance of using the per-phase signal when compared to the phase-aggregate, as otherwise the elimination of the FP results from similar loads would be impossible.
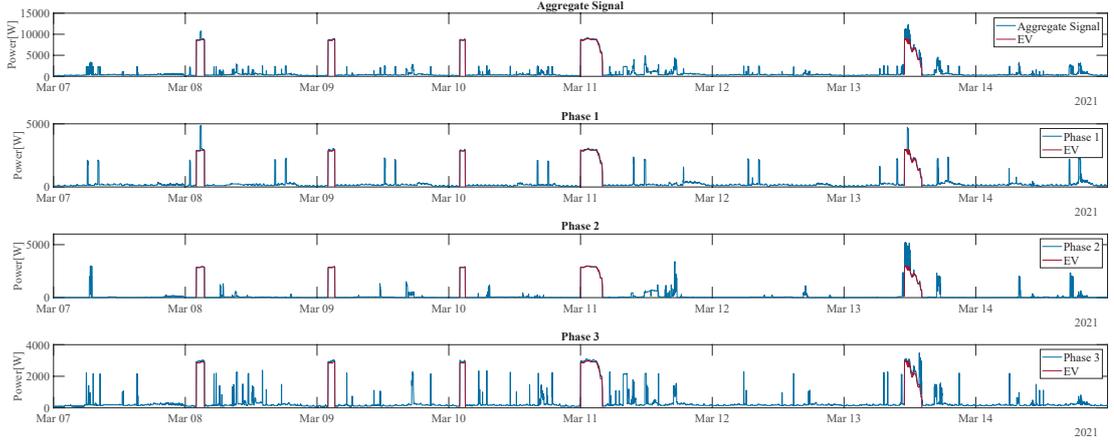
Figure 3.1: Labelling the EV usage: Energy disaggregation results per phase on the unseen household. Phase 1, 2 and 3 show the phase 1, 2, and 3 aggregate measurement, respectively. EV denotes the total EV disaggregated load (top row) or per phase estimate (second, third, and fourth row).

Table 3.6: Estimated accuracy of disaggregated EV load when ground-truth data are absent.

	ϕ_1	ϕ_2	ϕ_3	ϕ_{All}
1 min	90.65%	93.13%	90.98%	91.84%
15 min	86.99%	87.46%	86.74%	87.46%

3.4 Discussion & conclusions

In this chapter, the improvement of load disaggregation per phase in three-phase instal-lations was demonstrated quantitatively, over the traditional approach of disaggregating the sum of three phases. This is especially timely given the update in national smart meter roll-outs to provide for the growing number of households that include high-power loads such as EVs. Appliances that tend to be hard to disaggregate in the literature, due to sparse activations and variable load profiles, are more prone to noise from unknown appliances and therefore disaggregating per phase mitigates the effect of false positives. The appliances that benefit mostly from disaggregating per phase are WMs, TDs, CMs, and MWs. Devices with similar load profiles benefit from being spread across different phases. This demonstrates the importance of carefully picking the phase on which each

appliance is connected during a three-phase installation. Therefore, it is recommended to distribute refrigerating appliances, when more than one exists, in different phases, as well as connecting resistive appliances with similar profiles, such as KETs, CMs and stoves on different phases. However, great attention should be given such that connected appliances are evenly spread across the installations' phases to avoid any possible imbalances in the active and reactive power of the grid. This can be achieved by carefully separating devices based on their profile and avoiding connecting all of the devices that share the same characteristics on the same phase. A mixture between devices that are purely resistive and devices that are consuming a higher amount of reactive power—such as WMs, EV chargers, etc.—should be pursued. The feasibility of accurately labelling an EV charger without the presence of ground truth data using the per-phase signals, as well as knowledge transferred from another one-phase dataset, was also demonstrated, and the annotated dataset was released. Lastly, as smart metering is moving towards less granular data with the main concern being data privacy of the end users, disaggregation using lower sampling frequencies in the area of 15 min is explored in Chapters 4 & 6.

Chapter 4

EVs load disaggregation & LCA

The content related to EV load disaggregation presented in this chapter has been published in Vavouris, A., Garside, B., Stankovic, L., & Stankovic, V. (2022). Low-frequency non-intrusive load monitoring of electric vehicles in houses with solar generation: generalisability and transferability. Energies, 15(6), 2200. The content related to LCA of EVs presented in this chapter has been published in Vavouris, A., Stankovic, L., & Stankovic, V. (2023). Integration of Drivers' Routines into Lifecycle Assessment of Electric Vehicles. Transportation Research Procedia, 70, 322-329.

4.1 Introduction & Background

Decarbonisation of transportation is a major activity worldwide towards "securing global net-zero by mid-century and keeping 1.5 degrees within reach" [70]. Transportation was responsible for 27% of the UK's carbon dioxide (CO_2) emissions in 2019, of which over 90%, or 111 Mega-tonnes (Mt), of CO_2 were a product of road transport vehicles. Cars, including taxis, have played a major role in these emissions as, combined, they produced 68 Mt of CO_2, corresponding to 61% of the road transport vehicle emissions or a staggering 15% of UK's annual emissions [71]. As transportation is an essential part of our daily activities, and therefore cannot easily be reduced, EVs are a promising solution to tackle this challenge. The short- to medium-term aim is to replace vehicles that run with internal combustion engines with electric ones, especially if the electricity

53

is produced by RES.

Driven by global climate change goals and transition to net-zero economies, many governments worldwide have provided attractive incentives to EV users, leading to a tremendous boom in EV purchases for residential and business use. Indeed, according to the International Energy Agency (IEA), at the end of 2023, more than 40 million EVs were in use, with the Net Zero Emissions Scenario expecting more than 300 million EVs to be in circulation by 2050 [58]. In the UK alone, as of April 2025, there are approximately 1.5 million battery electric vehicles (BEVs) and 825,000 plug-in hybrid electric vehicles (PHEVs), a number that is expected to skyrocket in the coming years [72]. The vast increase in the number of EVs has led to installation of charging points both privately and publicly. By April 2025 there were more than 76,500 public charger points installed throughout the UK with over 20% of them rated as "rapid" chargers (50kW or above) [73].

Although the manufacturing and recycling process — powertrain, batteries, and end-of-life — of EVs is more carbon intensive, incentives to adopt EVs is mainly driven from the positive outcomes from the usage of electricity instead of fossil fuels that can compensate for the higher manufacturing and recycling process CO_2 emissions [74], especially when electricity is generated using renewable sources of energy. Although numerous studies looked into LCAs and estimation of total CO_2 emissions during a vehicle's production and recycling stage [75, 76], still factors influencing GHG emission quantification during usage, which contributes to a significant share of emissions [74] remain unexplored. More specifically, EVs' fuel impact on the GHGs is dynamically changing and directly correlated to the generation mixture during the charging period, which is directly influenced by the end-users' charging routines. This increase in penetration of EVs in the market, and the subsequent change in the power demand, leads to power grids facing great challenges regarding the ability to supply, transfer and distribute power. Indeed, the exponential increase of electric car sales — both PHEVs and fuel cell, which are fuelled from the grid — requires major changes in the energy markets and grid infrastructure as electrification of transportation poses numerous challenges for the existing power networks. In particular, modelling shows

that large-scale residential charging of EVs could result in overloading of distribution networks during peak hours if infrastructure upgrades and smart grid management are not implemented [77, 78].

Understanding where and when EVs are charging is important for uptake modelling, supply planning, and grid infrastructure reinforcement [79]. Knowledge of EV charging patterns is also required for smart grid solutions such as DR [80] and vehicle-to-grid (V2G) [81]. Future energy polices and transport planning would also benefit from accurate information on EV charging patterns. In addition, information about household consumption as a result of EV charging could be useful for customers to manage the running costs of EVs similarly to fuel costs for petrol and diesel cars. Awareness of the financial and carbon footprint implications of EV charging at home may incentivise customers to charge at public points and places of work, or opt in to DR-based tariffs and V2G programmes, alleviating the overloading of distribution systems during peak hours as a result and maximising charging from solar feed-in.

There are two options for monitoring the presence of EV charging on power networks. The first is intrusive load monitoring, a popular approach which sub-meters the EV charger at charge point, requiring additional hardware installation and maintenance costs; while this may benefit the end user and car manufacturer who have access to the EV's charging consumption statistics, the data are often not readily available to utilities, grid operators and network operators for infrastructure planning and grid demand management. A preferred alternative is NILM, from which energy consumption and time-of-use of the EV chargers are obtained via advanced signal information processing of aggregate power data, collected at a single point of measurement, e.g., a smart meter. An up-to-date literature review on NILM for EVs is included in Subsection 2.4.3.

The LCA assessment of different models has been the topic of discussion of many research papers throughout the years with research focusing both on the modelling and simulation of different vehicular technologies — internal combustion engine vehicles (ICEVs), BEVs and PHEVs — as well as on studies that aim to estimate the emissions during different stages of a vehicle's life including the production and the recycle stage. LCA is a methodology that is standardised worldwide based on the ISO 14040:2006

Standard [82]. The European platform of LCA [83] — a European Commission's project that aims to provide good practice in LCA use and interpretation — is widely used in business and policymaking for solutions towards sustainable production and consumption. As EVs have only emerged during the last decade, there is insufficient data for full-cycle large-scale deployment—i.e., production, usage and recycling—and therefore, research is focused on methods to compare the actual carbon footprint of the EVs to that of current fossil-fuelled vehicles.

In [84], an extended review of the different LCA models available for the comparison of the different vehicular technologies is presented with a focus on the effects of the batteries on the total carbon footprint of the EVs. Complementary to that, authors in [76] presented a comparison between the different proposed LCA models for EVs and ICEVs as well as between the different life cycle cost models for the same technologies. Authors concluded that although the introduction of EVs would be beneficial to the environment as the GHGs will decrease, the toxicity caused to humans is increased due to the more demanding production of the electric powertrains and high-voltage batteries. The LCA estimation of a vehicle takes into account different parts of the vehicle's life, including the production stage — i.e., raw material extraction, vehicles' components manufacture and vehicles assembly — the usage stage — i.e., the carbon footprint of the production of the fuel used and the vehicles emissions — and lastly the recycling stage — i.e., the carbon recovered during recycling/reusing and the carbon footprint produced due to materials being buried in landfills. The majority of the different LCA models that have been proposed in the literature can be grouped in one of the following categories:

- Cradle to gate (CTGa): assessment of the production stage, i.e. from the raw materials to the production of the vehicle. In [85], the first CTGa model is presented for the assessment of the mass production of a battery used in an EV. For a typical 24kW battery of an EV it was estimated that a total of 3.4 metric tonnes of CO_2eq will be emitted;
- Cradle to cradle (CTC): assessment of the production, usage and recycling stage with materials' recovery for reusage;

- Cradle to grave (CTGr): the same approach as with CTC, but with the difference that the assessment ends right after the recycling stage and therefore does not take into account the repurposing of materials. In [86], a CTGr approach in estimating the LCA of EVs in Poland and the Czech Republic is presented, with the results for the EVs ranging between 172–276 gCO_2eq/km for Poland and 132–214 gCO_2eq/km for the Czech Republic;

- Well to tank (WTT): assessment of the carbon footprint of the supply of different fuels used to run a vehicle. In [87], a WTT LCA assessment model is used with EVs emitting $155gCO_2eq/km$ and petrol cars $300gCO_2eq/km$;

- Tank to wheel (TTW): assessment of the carbon footprint due to the operation of a vehicle, and lastly,

- Well to wheel (WTW): assessment that combines both the WTT and the TTW methods. In [88], a WTW LCA model is introduced, with results ranging from 160 gCO_2eq/km for EVs, 270 gCO_2eq/km for petrol-powered vehicles and 230 gCO_2eq/km for diesel-powered vehicles.

In general, literature tends to tackle the LCA problem by different methods that consider the production, usage and end-of-life withdrawal of the vehicles. Assessment models, as summarised in [76] and in [84], are either performed in a generic way or are focused on a specific country and therefore the carbon footprint of the electricity used to charge the EVs is calculated based on a country average. However, models do not take into account the intraday variability of the carbon footprint of the electricity produced in a network that is correlated with the generation mixture at that specific moment. Considering that, and the fact that end-users are not charging their vehicles uniformly throughout a day, make an assumption that the average carbon footprint of the electricity is representative, is leading to under- and/or over-estimations of the actual impact of the EVs. Therefore, the integration of users' routines within the LCA models would improve their accuracy and, at the same time, improve the fairness when comparing different vehicular technologies. Lastly, as current LCA models assume a national average when estimating the footprint of the electricity used to charge the EV batteries, the variance in generation mixture of different regions in a national level is not

taken into account, a fact that is very timely especially considering the EV initiatives to reduce GHGs emissions that are being deployed in different countries around the world — which by nature have a finite budget — and therefore do not fully optimise the deployment of these initiatives.

This chapter addresses all the aforementioned gaps and proposes a novel, rigorous methodology for EV load disaggregation and evaluation, leveraging upon a DNN-based approach that has already been shown to outperform other learning approaches for NILM regression. A detailed and robust methodology for large-scale evaluation of EV load disaggregation from household smart meter data is presented, leveraging on a prior NILM algorithm that has been shown to have excellent regression performance on standard household appliances (see Chapter 3). Seq2subseq DNN is used to perform a sequence transformation and therefore, is appropriate for identifying electrical load signatures. The seq2subseq network was chosen as a trade-off between the convergence speed of seq2seq and the computational load of seq2point, as the proposed methodology should be both accurate and computationally efficient for scalability. This approach is then compared with a RF classifier as used for EV load classification in [54], and load reconstruction as in [67]. Lastly, the output of the disaggregation methodology is used to inform and augment LCA models by incorporating usage factors (as obtained from load disaggregation) and better quantify vehicles' lifetime carbon footprint.

The main contributions of this chapter can be summarised as:

- Adapting seq2subseq DNN-based NILM from [64] to EV load estimation, providing full details to enable reproducibility of the work, including hyperparameter tuning and post-processing steps;
- Evaluation of performance of the DNN approach when training and testing at the same location on 15 real houses from two geographical regions in the USA from the Dataport dataset, with 1 and 15 min resolution data containing high power interference from AC and different EV load profiles, where the EV charging power, duration of EV charging events, sparsity of charging events, and the relative noise or interference from unknown loads that could negatively affect disaggregation performance, are calculated and reported for each house;

- Rigorous evaluation of the scalability of the above NILM approach, with a focus on creating realistic test scenarios including generalisability on unseen households with EVs with similar EV load profile from Austin, Texas and cross-domain transferability evaluation on unseen houses with a different EV load profile from New York (NY);

- Quantifying generalisability and cross-domain transferability of the proposed methods by adapting metrics of [1, 89];

- Evaluation of meaningfulness of standard and NILM-specific metrics and recommendations for EV load disaggregation for network operators;

- augmenting existing EV LCA models and enhancing their accuracy by including usage factors that impose specific time of charging patterns obtained through a combination of NILM and qualitative data.

For NILM evaluation, the most popular metrics are discussed in Subsection 2.4.1. The choice of the dataset and how challenging it is to disaggregate loads of interest can be measured through the noisiness of the dataset as discussed in Subsection 3.2.2 and Equation 3.2. Additional metrics to calculate the generalisation loss that occurs when testing a NILM model on unseen houses are described in [89]. Given the potential impact of residential EV charging on the smart grid and the benefits of NILM for EV charging consumption and time-of-use for network operators and energy consumers, this chapter presents and discusses results using the above metrics.

The rest of this chapter is organised as follows. In Subsection 4.2, a rigorous approach to evaluate the generalisability and transferability of seq2subseq DNN for consumption estimation of EVs is proposed, which is followed by Subsection 4.2.2 that expands on the methodology followed to integrate the end users' specific routines to current LCA models. In Section 4.3.1, the proposed EV NILM methodology is evaluated using generic regression metrics, as well as NILM-specific consumption metrics, which is followed by Section 4.3.2 where the results of the LCA methodology are presented. This is followed by Section 4.4, where observations are discussed and conclusions are made in relation to EV load estimation to inform grid demand and LCA of EVs.

4.2   Methodology

The methodology for estimating the EV charging load and the approach to improve the LCA models through the integration of EVs' charging patterns are presented in this Section.

4.2.1   EV load disaggregation

4.2.1.1   Experimental data selection and preparation

The data acquisition process for the development of a supervised NILM methodology involves the selection and preparation of a dataset with aggregate and sub-metered appliance power measurements, sampled at low to very-low sampling rates, constituting "labelled data" for algorithm training and testing. As discussed in Subsection 2.4.3, the Dataport [52] dataset has been primarily used for EV NILM evaluation in the past, as it contains many households in different areas of the USA with different appliances, including EVs, that have been monitored continuously for a long period. The data portal was accessed through a University research account, providing free access to 1-second, 1- and 15-minute data collected from 73 houses across Austin, Texas, California, and NY. Of these 73 houses, 8 from Austin, 7 from NY and 0 from California were listed as owning EVs. Available data for Austin houses span over 12 months — from 1 January 2018 until 31 December 2018 — and for NY houses over 6 months — from 1 May 2019 until 31 October 2019.

Table 4.1 represents a summary of the Dataport houses used for the experiments, including metadata on the presence of AC and solar generation, and the total amount of EV charging time per household as well as the sparsity of EV load. The latter information, i.e., the total amount of EV charging time and the sparsity of EV load — that is rarely stated in the literature — but provides an indication of the amount of data available for training and testing. Houses 2335, 3517 and 5058 were omitted from Table 4.1 as either their EV sub-metering was found to be null or contained no EV charging activity. From the remaining houses, 3 were discarded as the data were faulty and/or scarce: House 3000 appeared to have erroneous data as the addition of

mains reading, i.e., the amount of energy consumed in the household, and solar power generation, which is either consumed or fed back into the grid if the production is greater than usage, did not add up to the grid reading, i.e., the total energy equilibrium that is apparent from the connection point of the house to the grid; House 7719 included only 71 hours of charging events throughout a period of 12 months, which was equivalent to only 10 activations of a charge and insufficient to produce meaningful results; House 9053 had noisy sub-metering, i.e., there were unusually long periods where the load pattern did not resemble a typical EV charging signal, probably because there was another load metered on that plug. Lastly, house 4767 appears to have changed the EV charger from 4 kW to 6.6 kW after 3 July 2018 and therefore, in the experimental results later on, results are presented for 4767-1 and 4767-2, to represent House 4767, before and after this change of the charge, respectively.

All 9 selected houses have solar panels, and all but one have high power AC interference, making the NILM task challenging. The whole dataset of each of these 9 houses is used for the experimental tests, so that different energy usage patterns can be observed across different seasons of the year. Austin houses contain data that spread across all seasons in a year, whereas NY houses contain data from late Spring until early Autumn.

Table 4.1: Summary of EV charging in Dataport houses used for the experiments, including *NM* [1] of the considered households in Dataport dataset. EV sparsity is calculated as the charging duration divided by the total monitoring duration.

Area	House	EV charging power [kW]	Charging duration [h]	EV sparsity	AC/solar	$NM^{(T)}$	$NM^{(EV)}$
Austin	661	3.3	781	8.92%	Yes/Yes	86.35%	39.34%
	1642	3.3	982	11.21%	Yes/Yes	82.37%	39.19%
	4373	3.3	1359	15.51%	Yes/Yes	75.91%	44.48%
	4767-1	4.0	485	10.98%	Yes/Yes	86.60%	36.49%
	4767-2	6.6	485	11.16%	Yes/Yes	85.65%	26.12%
	6139	3.3	622	7.10%	Yes/Yes	90.50%	40.31%
	8156	3.3	615	7.02%	Yes/Yes	90.67%	47.34%
NY	27	3.3	338	7.65%	Yes/Yes	74.99%	21.15%
	1222	6.6	139	3.15%	No/Yes	82.86%	25.54%
	5679	6.6	210	4.76%	Yes/Yes	76.76%	30.67%

4.2.1.2 Quantifying interference

NILM is a source separation problem, where any consumption measurement other than the target loads of interest is considered as an interfering signal, or noise. Therefore, the noisier a dataset, i.e., the more unknown or non-submetered loads, the more challenging the classification and disaggregation problem, directly impacting the accuracy. To quantify the difficulty of successfully estimating individual loads from aggregate Equation 3.2 was used. The above measure assumes that there is an equal interest in estimating all M "targeted"/submetered loads. In the case discussed, and in Table 4.1, the NM for $M = 1$ is presented where the only m of interest is the EV, and denotes $y_t^{(1)} = y_t^{(EV)}$. Since this research is interested only in disaggregating EV loads, all other loads contributing to the aggregate would be considered as noise. Therefore, Equation 3.2 is slightly revised such that the noise, i.e., unknown loads, are only considered during EV charging times, to capture better their interfering effect:

$$\% - NM^{(EV)} = \sum_{t=1}^{T} c_t \left| 1 - \frac{y_t^{(EV)}}{y_t} \right|, \tag{4.1}$$

where an indicator $c_t = 1$ if EV charging is ON during time sample t, i.e., $y_t^{(EV)} > 0$, and zero, otherwise.

Table 4.1 includes the noise metrics $\% - NM^{(T)}$ and $\% - NM^{(EV)}$ from Equations 3.2 and 4.1 respectively, for all the households under consideration. The higher the NM, the more interference from unknown loads, and therefore the more challenging to accurately estimate energy consumption. It can be observed that a lower $\% - NM^{(T)}$ metric does not always imply a lower $\% - NM^{(EV)}$ metric. Both metrics will be reviewed in an attempt to explain classification and regression performance.

4.2.1.3 Train – test split

The entire dataset is split into train and test datasets at the pre-processing stage. A rigorous approach to split training and testing data is proposed, where a small number of days are randomly selected from each month to make up the test dataset. The number of days selected from each month is set to obtain a train–test split ratio of around

70:30, resulting in 10 days of each month kept for testing purposes. Days are chosen at random to guarantee a natural distribution of EV charging vs. non-charging windows, and to demonstrate that days have not been "hand-picked". Selecting the same number of days from each month ensures that the method is tested equally across all seasonal variations in solar generation and appliance use, e.g., AC in summer, furnace in winter. Leaving the test windows in-order creates a more realistic simulation of a real-world NILM system and allows for complete EV loads to be reconstructed for visualisation and evaluation of performance in terms of consumption metrics. After the test dataset has been formed, the remaining windows are randomly undersampled — by removing windows with no EV charging — to obtain a balanced train dataset that is randomly ordered to ensure no bias in the training of the classifier. This process is summarised in Figure 4.1 and repeated for each house.
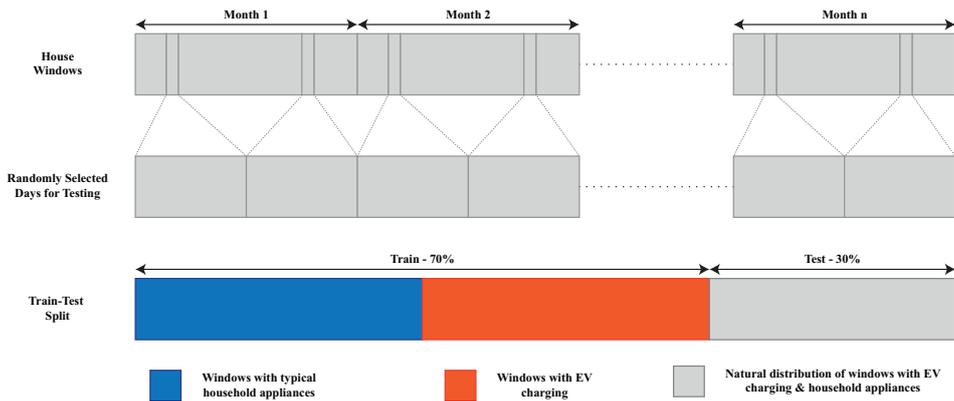
Figure 4.1: Visual demonstration of the proposed method for obtaining a balanced train dataset and a realistic in-order test dataset for each house, where n is the number of months in the data collection period for a given house. Note that in the train dataset, the blue block contains periods without EV charging and the orange block is formed from the periods with EV charging and other household's appliances running in parallel.

4.2.1.4   Regression based on DNN

Seq2subsseq learning is adapted and optimised from Subsection 3.2.3, to disaggregate EV loads. For the disaggregation algorithm to work successfully, the whole EV charging event must be included in the targeted subsequence of the DNN. Thus, the optimal

seq2subseq size ω is set from Equation 3.3, where L [in mins] is the usual length of EV charging period, and S [in mins] is again the resolution of the data samples — 1 or 15 min. The epoch number — the number of times that the learning algorithm worked through the entire training dataset — was also varied, and the number of epochs that maximise performance was used, using an early stopping criterion on the validation set. In addition, other hyperparameters, including the number of generator filters as well as the number of discriminator filters, were tuned using the validation set, and depend on the sampling rate. Lastly, the number of layers was adapted based on the window size. The selected values are reported in Section 4.3.1.

Given the produced sub-sequences of EV load, a simple correction procedure similar to the procedure in Subsection 3.2.3.3 is applied. It was observed that the seq2subseq model produced some FP results, mostly due to interference from other similar loads. All of these FP sub-sequences had one feature in common: their maximum value appeared to be very small, around 5 W. Therefore, these values are simply zeroed. Additionally, negative values that were produced from the DNN were also zeroed, as negative values have no physical significance.

4.2.1.5   Generalisability and transferability evaluation

Since labelled data are scarce, hard to collect, and thus only available for a small portion of households, the proposed methods must be able to produce reliable results on unseen houses in a similar dataset — generalisability — and unseen houses, i.e., houses without any labelled data, from another domain — transferability. To this effect, generalisation loss is proposed in [89] to evaluate the performance of a NILM algorithm for both event detection from classification and load estimation on an unseen house. For event detection, generalisation loss, as a percentage, is given by Equation 4.2 — a comparison between the classification accuracy on unseen (ACC_u) and seen houses (ACC_s), where ACC can be the *Accuracy* (Equation 2.8) or $F-$score (Equation 2.11) metric. Similarly, for regression accuracy, Generalisation Loss, as a percentage, is given by Equation 4.3 — a comparison between the error on the unseen houses (ERR_u) and seen houses (ERR_s) — where ERR can be any of the standard regression metrics such as MAE, SAE, and

RMSE.

$$G_{\text{loss}}^{\text{class}} = \left(1 - \frac{\text{ACC}_{\text{u}}}{\text{ACC}_{\text{s}}}\right)\%, \tag{4.2}$$

$$G_{\text{loss}}^{\text{reg}} = \left(\frac{\text{ERR}_{\text{u}}}{\text{ERR}_{\text{s}}} - 1\right)\%. \tag{4.3}$$

For the experimental evaluation of generalisation loss of regression performance, the use of Equation 4.4 instead of Equation 4.3 is proposed. This approach is preferred as a more accurate measurement of performance of consumption estimation, and is calculated based on the *Acc* metric (Equation 2.14).

$$G_{\text{loss}}^{\text{energy}} = \left(1 - \frac{Acc_u}{Acc_s}\right)\%. \tag{4.4}$$

The same generalisation loss equations can be used to evaluate generalisability and transferability. In the former case, unseen houses would be those from similar datasets. In the case of cross-domain transferability evaluation, unseen houses would be from another domain.

4.2.2   Informing LCA of EVs

Following the methodology for the disaggregation of the EV charging load, this subsection presents the methodology to augment current LCA models in order to increase the precision of the usage estimation of the different technologies and therefore increase the trustworthiness of the solutions and the fairness in the comparison of the different technologies. An EV charging time-of-use informed LCA model is introduced, and a summary of the steps followed to estimate the carbon footprint of each technology is presented in Figure 4.2. As this model focuses on the effects of the EVs' usage compared to ICEVs' under different charging profiles and different regions — both national and international — carbon emissions involved during the battery and vehicle production and end-of-life stages were calculated based on the updated CTGr LCA model of the European Federation for Transport and Environment [74] and therefore only the methodology followed to calculate the carbon footprint of the usage cycle and

how this is affected based on different charging times, different charging speeds and different regions is presented.

Figure 4.2: Methodology Summary.

4.2.2.1 Time of charging

To integrate end-users' routines in the usage component of LCA models, different types of data were collected across different countries, including survey data and actual energy consumption data. Countries that were considered in the research were the UK, Norway and Germany.

In UK, as there are no publicly available datasets of EV usage, time of charging was based on the 2022 smart chargepoint survey [90] commissioned by the department of business energy and industrial strategy (BEIS) where 1,000 EV drivers participated in the research as well as on the department for transport report on EV charging research [91]. According to the responses of this survey, the majority of the respondents have access to a private driveway, garage or other form of off-street parking, with only 6% parking on-street, a fact that is detrimental to the selection of charging technology and charging place. A total of 66% of end-users own a dedicated chargepoint, 26% use a standard 3-pin cable which is directly plugged into the mains socket and 1% use a private access communal dedicated chargepoint. Therefore, a total of 93% of the participants have access to charging at home and select to do so. According to [91], the vast majority of people with dedicated chargepoints prefer charging their vehicle during nighttime, with 78% of the participants reporting charging their EV overnight regularly. Therefore, in this research, two different users' profiles with BEVs are studied:

I. a user with a slow charging 3-pin cable (2kW) that charges directly after normal working hours at 18:00 with an average total duration of 11 hours; and,

II. a user with a dedicated chargepoint (7kW) that charges overnight with an average duration of 3 hours.

As the UK consists of Great Britain (GB) and Northern Ireland (NI) — where a different electricity system operator (ESO) exists — assessment was only carried out for the GB.

In Norway, a field study was performed in [92,93] (see Chapter 5). Based on smart metering data, comprising BEVs and on closed-format questionnaires in a smart-home district in the greater area of Frederikstadt, two charging profiles were created for:

I. homeowners that own a dedicated chargepoint (11kW) with scheduling capabilities — as the energy price in that area varies on an hourly basis user opted to charge during the night hours at midnight with an average duration for a full charge of 3 hours, when the energy is cheaper; and,

II. homeowners that do not own a dedicated chargepoint and therefore charge their cars 3-pin cable (3kW) based on their daily routines — i.e., after returning from work, circa 17:00 — with an average total charge duration of 11 hours.

In Subsection 3.3.2, a household in Germany was monitored and labelled for a period of 1 year — 1 January 2021 to 31 December 2021 — where an EV was present with an installed fast EV charger of 11kW. The aggregate readings, as well as the annotated BEV chargers' activations at 30 min resolution, were used for creating Germany's user profile. The average charging duration was approximately 3 hours. The data can be accessed at [94].

4.2.2.2 Fossil fuels & electricity carbon footprint

Fossil fuels' carbon footprint is correlated with the penetration of renewable fuels that made up 7% of the total road and non-road machinery fuel in the UK in 2022 [95]. Renewable fuels are produced with the use of either crops or wastes known as feedstocks. When comparing the GHGs emissions of fossil fuels to the renewable ones, there is a total saving of 81%. Considering the indirect land-use change (ILUC) — i.e. the unintended consequence of switching land use for the generation of renewable fuels — this percentage is slightly less at 77%. These renewable sources were included into the proposed LCA model.

Compared to ICEVs, whose fuel refill timing is not correlated to their carbon footprint, EVs use energy that is instantaneously produced. Therefore, data regarding

the generation mixture as well as their CO_2 emissions were gathered through the ESOs on a half-hourly interval for 1 year, starting on 1 January 2022 to 31 December 2022. For the GB, the carbon intensity API [96] was used to retrieve the carbon footprint of the produced energy on a regional and national level. For Norway and Germany, the generation mixture obtained through the European network of transmission system operators for electricity (ENTSO-E) [97] was used and then translated to an approximate carbon footprint based on the updated work of [74] for the parameters of [98] to better capture the actual European region instead of global averages, i.e. 997 gCO_2eq/kWh for coal, 434 gCO_2eq/kWh for gas, 34 gCO_2/kWh for solar PVs, 14 gCO_2eq/kWh for offshore winds, 11 gCO_2/kWh for hydro, 12 gCO_2/kWh for onshore wind and lastly 5 gCO_2eq/kWh for nuclear.

4.2.2.3 Usage assessment

A one-year simulation of the actual footprint based on the charging time — as identified through the surveys and the actual electricity consumption of different households — was performed. Results were then extrapolated to a vehicle's full lifetime — i.e., the total expected mileage before withdrawal from circulation. Powertrain parameters for the LCA model of a medium-sized vehicle are presented in Table 4.2 as obtained through [74].

Table 4.2: Powertrain parameters for a medium-sized vehicle.

	Petrol [l/km]	Diesel [l/km]	BEV [kWh/100km]	BEV (capacity) [kWh]
Medium-sized	7.5	6.2	17.5	60

4.3 Results

4.3.1 EV load disaggregation results

This section first describes the evaluation strategy, then presents the experimental results for regression and generalisability/transferability for EV NILM as per Subsection 4.2.1.4. More specifically, regression models are trained and tested on the same house, referred

to as "observed" scenario, to demonstrate the ideal performance of the methodology under realistic test conditions, i.e., randomly selected days of in-order windows that have not been balanced. Secondly, generalisability and transferability are evaluated on unseen houses that are not part of the training set. The metrics that are presented are the product of ten independent executions of the same algorithm to ensure experimental repeatability. Both the mean value and the standard deviation of each metric are presented to further underline the meaningfulness of each metric. A metric with a low standard deviation is more robust to randomness that is introduced due to the random initialisation of algorithms. Therefore, these metrics are considered to produce much more concise results as they present a more accurate performance of the models. Results are then compared with the RF classification and load reconstruction strategy of [67].

4.3.1.1   Evaluation strategy

The steps of the methodology described in Subsection 3.2 for EV disaggregation from a realistic test dataset, with a natural imbalance, in-order windows and randomly selected days were first performed. The training set was split 60% for training, and 10% for cross-validation to determine the best set of parameters, which are then fixed for final results of testing on an unseen 30% of the samples. As explained in Section 4.2.1.3, the testing set used for the presented results below comprised a total of 120 days for each Austin house and 60 days for each NY house when testing and training on the same household. For generalisability and cross-domain transferability tests, all data available from unseen houses — i.e., 12 months for each Austin house and 6 months for each NY house — were used as testing sets. Repeatability of experiments is also performed by repeating testing on the same conditions and data ten times.

Simulations were carried out such that regression models were trained and tested on the same houses individually to observe how accurately EV consumption can be estimated under these ideal conditions. Generalisability tests were then carried out to fully evaluate the NILM approach under conditions required for implementation in a real-world NILM system. This procedure involved training regression models on a selected number of houses and testing on unseen houses belonging to the same

geographic area and similar EV load signatures. Transferability tests were carried out on unseen houses in a different geographical area and with different EV load signatures.

4.3.1.2 Seq2subseq: regression simulation setup

For the seq2subseq with conditional GAN network, the window size is set to 512 and 64 samples, for 1 and 15 min data, respectively, as these values led to the best performance on the validation set. L1 loss was used in both setups, whereas SGD and ADAM optimisers were used for the discriminator and generator filters, respectively. The initial learning rate was 0.001 (for SGD) and 0.0005 (for ADAM). The momentum term of ADAM was equal to 0.5, and the weights on L1 and GAN term for the generator gradient were 100 and 1, respectively. For 1 min data, a total of 7 layers were used and thirty-two generator and discriminator filters in the first convolutional layer. For 15 min data, a total of 5 layers were used and 4 generator and discriminator filters in the first convolutional layer. Using the early stopping criterion on the validation set, it was concluded that the best value for the number of epochs is 120.

4.3.1.3 Performance: observed scenario

Figure 4.3 presents the results obtained by training and testing on the same household for 1-minute resolution for the DNN-based regression. The standard deviation for all metrics over 10 runs of the experiment, in identical conditions, is less than 2%, which indicates that all the experiments are repeatable. However, in Figure 4.3 it is observed that the MAE metric has a significantly larger standard deviation of 10–15%. This is a common observation of DNN-based regression, and therefore why presenting results exclusively with the MAE metric, which is rife in recent NILM literature, can often be misleading, especially if repeatability of experiments is not demonstrated.

Comparing *Acc* and MR metrics between load estimation via regression vs load reconstruction from [67], it can be seen, as expected, that the DNN network is more susceptible than RF to insufficient samples in the training set, as exemplified by house 1222 results, which has the fewest EV charging hours. For the same reason and the fact that house 4767 changed its EV charger, with a significantly different wattage,
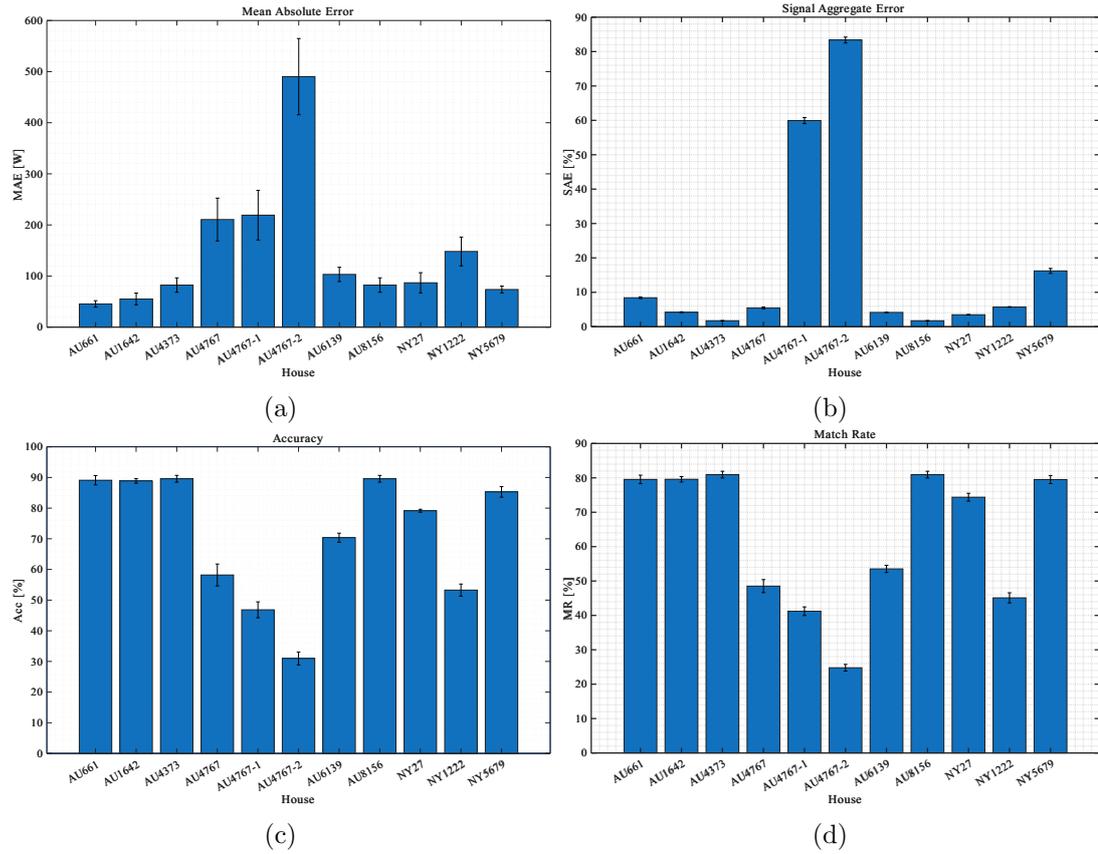
Figure 4.3: Mean values and standard deviations of assessment metrics, training and testing on the same household, using seq2subseq, 1 min data.

in the middle of the year — as discussed in Section 4.2.1.1 — the regression network also performs relatively poorly. Otherwise, the *Acc* and MR performance measures are consistent for EV load estimation from the regression approach. Furthermore, house 6139 has relatively poorer load estimation performance compared to others because it has the highest NM value as shown in Table 4.1. Regression was also performed on 15-min data. Figure 4.4 presents results produced from training and testing on the same household on 15-minute data for the DNN-based, seq2subsesq regression network. Using 15 min granularity data, as the available activation windows were fewer — approx. 7% less activation windows than in the 1-min data — training and testing on the same household produced poorer results. This is a result of the combination of fewer activation windows and lower data granularity. As discussed previously, DNNs are more susceptible to the number of training samples compared to RF classification and load

71

reconstruction approaches.
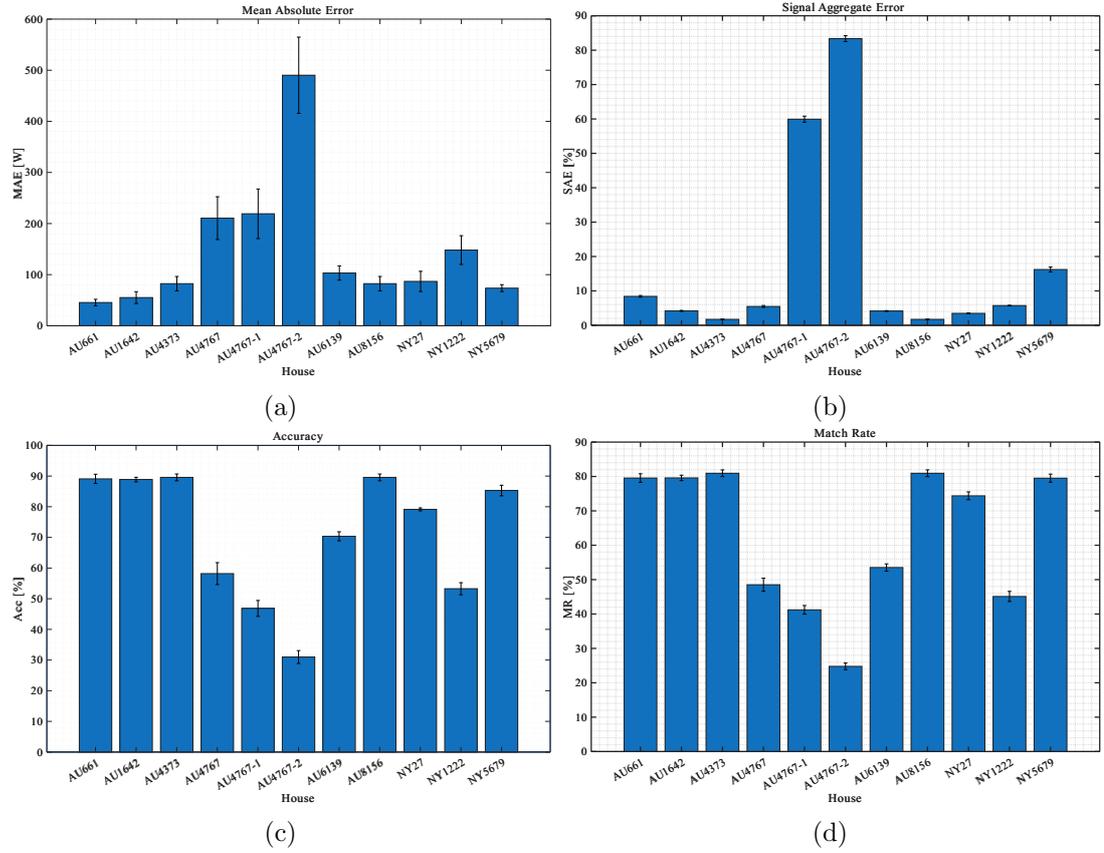
Figure 4.4: Mean values and standard deviations of assessment metrics, training and testing on the same household, using seq2subseq, 15 min data.

Table 4.3 demonstrates the loss incurred in consumption estimation performance, reducing the granularity of meter readings from 1 min to 15 min. The classification and load reconstruction of [67] is included for comparative purposes. Loss for load consumption estimation was calculated as:

$$Loss = \left(1 - \frac{Acc_{15\min}}{Acc_{1\min}} \right) \% \qquad (4.5)$$

As expected, the loss in $F-$score, i.e., accuracy of detecting EV load charging events, is very small except for house 6139, which had a relatively high noise metric — as per Table 4.1 — and house 1222, which had insufficient training. Energy estimation using RF classification and load reconstruction is more robust to lower frequency

data when compared to seq2subseq regression. As it was already stated, this can be the result of multiple factors, mainly due to insufficient activation windows when training the seq2subseq algorithm and also due to the fact that compared to the load reconstruction method using RF classification results, seq2subseq network does not have a priori knowledge of the EV charging level. Additionally, the loss in energy estimation performance for both approaches is correlated with the NMs as per Table 4.1, where houses 6139, 8156 from Austin area and house 1222 from NY area, have relatively higher NMs and therefore greater losses in performance when using lower frequency data.

Table 4.3: Classification and regression granularity loss based on $F-$score and Acc metrics, respectively, using RF classifier and load reconstruction and seq2subseq network.

Area	House	Classification (RF)	Load reconstruction (RF)	Regression (seq2subseq)
Austin	661	6.43%	7.32%	15.29%
	1642	-2.24%	8.52%	22.22%
	4373	6.30%	10.67%	32.77%
	4767	-0.09%	8.43%	29.61%
	6139	25.05%	36.25%	22.43%
	8156	6.93%	25.58%	33.83%
NY	27	-9.60%	-6.48%	4.62%
	1222	15.10%	20.16%	44.47%
	5679	5.26%	10.12%	25.84%

4.3.1.4 Generalisability results

As with any real-world NILM scenario, the proposed solution should be able to transfer knowledge from known houses to unknown ones that belong to the same area or, in this case, use a similar EV charging load. This is essential, as the collection of metadata and/or labelled data for all households is a costly and time-consuming process, and end-users are not always keen on sharing their personal information. Generalisability evaluation performs testing in houses, not included in the training set, that:

- belong to the same geographical area as the houses in the training set; and,
- use the same EV charging level as houses in the training set.

Table 4.4 presents results obtained by testing in the same area as training houses

using the seq2subseq network. Each of the experiments was performed by testing on all houses in an area, apart from one, which was kept for testing purposes. For example, results presented in Table 4.4 for Austin house 661 were obtained by training on the entire period for all Austin houses except house 661, and testing on the entire period of house 661. Results of Table 4.4 are compared against those of Figure 4.3, where the difference in performance when testing on an observed house vs testing on an unseen house in a similar geographic area, namely Austin or NY is evident. This is also captured via the $G_{\text{loss}}^{\text{reg}}$ and $G_{\text{loss}}^{\text{energy}}$ metrics of Equations 4.3 and 4.4, respectively, which indicate the MAE and *Acc* loss in performance. The two loss metrics are generally in agreement in terms of relative performance, except for the NY houses. The $G_{\text{loss}}^{\text{reg}}$ of house 27 is unusually high because it is the only house in NY with a 3.3 kW EV, and the regression network was trained on the other two houses with 6.6 kW EVs, therefore, the energy consumption is overestimated. However, this is less pronounced in the $G_{\text{loss}}^{\text{energy}}$ metric. Overall, the performance loss is negligible across all metrics, except for the marginal drop in performance for Austin houses 661, 4373 and 8156. This is captured by the positive G-loss for these values, which are less than 15%. Houses 1222 and 5679 experience a more significant drop in performance, as captured by both $G_{\text{loss}}^{\text{reg}}$ and $G_{\text{loss}}^{\text{energy}}$, because they are both trained on house 27, which has about 50% more EV load charging events at 3.3 kW and therefore the energy consumption is underestimated for these two houses with 6.6 kW EVs.

Furthermore, as captured by the large negative G-loss values, houses 4767-1 and 4767-2 now have significantly improved EV load estimation performance because the issue of insufficient training data previously encountered has been resolved with training on all other houses.

Table 4.5 present the results of generalisability tests for training on all houses with 3.3 kW loads regardless of geographical area, except house 1642, and testing on unseen house 1642, for both 1 and 15 min resolutions for the seq2subseq DNN approach. Comparing both Figures 4.3 and 4.4 with Table 4.5 for the regression network, and as indicated by $G_{\text{loss}}^{\text{reg}}$ and $G_{\text{loss}}^{\text{energy}}$, it can be seen that while the 1-min results are similar on observed and unseen scenarios, there is a significant improvement in performance for the

Table 4.4: Mean values of assessment metrics and $G-$loss, training on all houses of an area apart from one, and testing on the unseen house from the same area, using seq2subseq algorithm, 1 min data.

Area	House	MAE [W]	SAE	*Acc*	MR	$G_{\text{loss}}^{\text{reg}}$	$G_{\text{loss}}^{\text{energy}}$
Austin	661	73.54	6.29%	81.79%	70.00%	61.8%	8.19%
	1642	52.26	1.02%	89.46%	80.84%	–5.21%	–0.66%
	4373	99.16	4.33%	87.86%	77.92%	20.44%	1.88%
	4767	198.8	29.92%	63.60%	40.05%	5.60%	–9.30%
	4767-1	72.47	9.09%	82.48%	71.29%	–66.92%	–76.09%
	4767-2	278.6	50.90%	58.92%	28.94%	–43.17%	–90.19%
	6139	95.68	6.94%	71.61%	54.56%	–7.29%	–1.78%
	8156	97.08	12.63%	76.31%	63.56%	17.92%	14.80%
NY	27	177.8	4.37%	63.62%	47.49%	104.9%	19.61%
	1222	189.3	58.42%	36.56%	34.14%	27.82%	31.37%
	5679	110.6	2.21%	76.56%	61.69%	50.17%	10.24%

15-min results due the availability of additional training data from multiple houses. It can therefore be concluded that the regression network results are generalisable, without loss of performance, for similar EV charging levels.

Table 4.5: Mean values of assessment metrics and $G-$loss, training on all houses that were charging on 3.3 kW, and testing on unseen house 1642, using seq2subseq algorithm, with 1 min and 15 min granularity data.

Granularity	House	MAE [W]	SAE	*Acc*	MR	$G_{\text{loss}}^{\text{reg}}$	$G_{\text{loss}}^{\text{energy}}$
1 min	1642	50.41	6.70%	89.83%	80.96%	–8.56%	–1.08%
15 min	1642	76.30	14.97%	84.60%	71.47%	–46.64%	–22.40%

4.3.1.5   Transferability results

Evaluation of cross-domain transferability to assess how robust a model is to training and testing on different geographical areas and different EV charging levels is presented. Transferability tests can be summarised as follows:

   I. testing on an unseen house in NY and training on all other houses from Austin,

regardless of EV charging level;

II. testing on an unseen house with an EV charge level of 6.6 kW and training on all houses with EV charge level of 3.3 kW, regardless of geographical area; and,

III. testing on two unseen houses and training on a generic mix of houses from different areas and different EV charging levels.

Table 4.6 shows the outcome for transferability tests I. and II., for both 1 and 15 min resolutions, with the seq2subseq network approach. House 5679 from NY with an EV charging level of 6.6 kW was tested on RF and DNN regression models trained with all Austin houses containing EVs with 3.3 kW charging level. The training set comprised houses 661, 1642, 4373, 6139, and 8156.

Table 4.6: Mean values of performance and generalisation loss metrics for transferability tests I. and II., using seq2subseq algorithm, for 1 min and 15 min data.

Granularity	House	MAE	SAE	*Acc*	MR	$G_{\text{loss}}^{\text{reg}}$	$G_{\text{loss}}^{\text{energy}}$
1 min	5679	183.4 W	24.32%	61.15%	58.67%	149.0%	28.30%
15 min	5679	201.5	29.60%	56.92%	48.83%	18.67%	10.01%

Comparing Figures 4.3 and 4.4 with Table 4.6, and captured by $G_{\text{loss}}^{\text{energy}}$ than $G_{\text{loss}}^{\text{reg}}$, a drop in load estimation performance for both granularities is observed. The drop is more pronounced for 1 min granularity. Interestingly, from the regression network's output load reconstruction plots, it is observed that while the seq2subseq algorithm is correctly detecting EV charging events, it underestimates the EV load charging level since the network was trained on lower EV charge loads.

Similarly to Table 4.3, Table 4.7 demonstrates the loss introduced by using data with granularity of 15 min compared to 1 min. *Acc* metric was used the regression problem, whereas $F-$score values from the RF classification and load reconstruction approach in [67] are included for comparative purposes. As expected, the granularity loss is more pronounced during transferability than in the observed scenario — see house 5679 in Table 4.3. The regression network is less affected by reduced granularity when directly compared to RF classification and load reconstruction, as observed in Table 4.7.

Finally, a practical approach, as per transferability test III., was taken whereby

Table 4.7: Classification and regression granularity loss based on $F-$score and Acc metrics, respectively, using RF classifier and load reconstruction and seq2subseq for transferability tests I. and II.

Area	House	Classification (RF)	Load Reconstruction (RF)	Regression (seq2subseq)
NY	5679	14.96%	28.69%	6.92%

generic learning models were trained using a mix of houses across different geographic areas and containing different EV charging loads. The training set comprises houses 661, 4373, 4767, 6139, 8156, 27, and 1222. Testing was performed on the unseen Austin house 1642 and NY house 5679, with EV charge loads of 3.3 kW and 6.6 kW, respectively. Results are presented in Table 4.8. Comparing Figures 4.3 and 4.4 with Table 4.8, a significant drop in performance is observed when testing on house 5679. This could be the result of more houses that are charging at 3.3 kW in the training set compared to 6.6 kW. On the other hand, results for 15 min data on house 1642 are significantly improved, which is a result of more data available to the network during the training process. The $G_{\text{loss}}^{\text{energy}}$ than $G_{\text{loss}}^{\text{reg}}$ loss metrics for this transferability test III. — as shown in Table 4.8 — compared to transferability tests I. and II., as shown in Table 4.6, are relatively unchanged for 1 min granularity, but there is less loss for 15 min granularity. This shows that the seq2subseq regression model performs equally well on all transferability tests.

Table 4.8: Mean values of performance and generalisation loss metrics, for transferability test III., using seq2subseq algorithm, for 1 min and 15 min data.

Granularity	House	MAE	SAE	Acc	MR	$G_{\text{loss}}^{\text{reg}}$	$G_{\text{loss}}^{\text{energy}}$
1 min	1642	50.49 W	4.54%	89.81%	81.12%	–8.42%	–1.06%
	5679	178.4 W	26.94%	62.21%	49.21%	142.23%	27.06%
15 min	1642	78.66 W	11.69%	84.13%	71.15%	–44.99%	–21.72%
	5679	183.6 W	33.00%	60.75%	46.04%	8.13%	3.95%

This experiment demonstrates that if an adequate number of houses of a certain wattage level are included in the training set, then when testing on an unseen house that uses a same power level charger, the model is agnostic to the other EV power levels that are presented in the training set, and produces an accurate result. Lastly, Table 4.9

demonstrates the loss introduced by using data with granularity of 15 min compared to 1 min and compares the DNN-regression approach to the RF classification approach of [67]. As it can be observed the DNN regression approach is much less susceptible to reduce data granularity when compared to an RF classification and load reconstruction approach.

Table 4.9: Classification and regression granularity loss based on $F-$score and Acc metrics, respectively, using RF classifier and load reconstruction and seq2subseq for transferability test III..

Area	House	Classification (RF)	Load Reconstruction (RF)	Regression (seq2subseq)
AU	1642	-1.24%	6.72%	6.32%
NY	5679	15.22%	28.52%	2.35%

4.3.2 LCA results

In this Subsection, the results obtained by following the steps discussed in Subsection 4.2.2 are presented. Results are both on the national level for the GB, Norway and Germany, as well as on a regional level for the GB.

4.3.2.1 National level

Figure 4.5 illustrates a comparison of the GHG emissions between the ICEV and EV model of [74] compared to the proposed model, where user-charging-routine information is integrated into the models. A medium-sized vehicle — as presented in Table 4.2 — with an estimated usage of 225,000 km[1] in its lifetime is assessed in three different households. In the GB household, based on the two scenarios — i.e., a user with a 3-pin system and a user with a dedicated chargepoint — a 5.8% increase (i.e. an increase of $\sim$0.8 tCO$_2$ in the vehicle's lifetime emissions) and a — 13.3% decrease (i.e. a decrease of $\sim$1.9 tCO$_2$ in the vehicle's lifetime emissions) is observed for Scenarios I. and II., respectively. In the 3-pin scenario, the increase in the carbon footprint is expected due to the peaking power plants that are introduced to the grid to meet the increased

---

[1]Please note that only the usage component of the selected LCA model [74] has been adapted. All other components remain unchanged. The battery component for EVs incorporates the requirement for a battery replacement after a specific mileage.

demand that is usually exhibited during the evening hours. In the German household, based on the smart metering data, the BEV is charged mainly during night — when the grid is less stretched and therefore base power plants can handle the load. A reduction of 12.9% (i.e., a decrease of $\sim$2.7 tCO$_2$ in the vehicle's lifetime emissions) is observed. Lastly, for Norway's scenarios, according to smart metering and interview data, in the 3-pin Scenario (Scenario I.) a slight increase is observed of 1.5% (i.e. an increase of $\sim$0.2 tCO$_2$ in the vehicle's lifetime emissions) whereas in Scenario II., in which a similar behaviour with German household is exhibited — i.e., charging during night hours — a negligible reduction of 0.6% (i.e. a decrease of $\sim$0.1 tCO$_2$ in the vehicle's lifetime emissions) is observed. This is due to the particular nature of Norway's generation mixture, which consists almost exclusively of hydro generation and therefore peak demand is not covered with the use of carbon-intensive peaker plants.
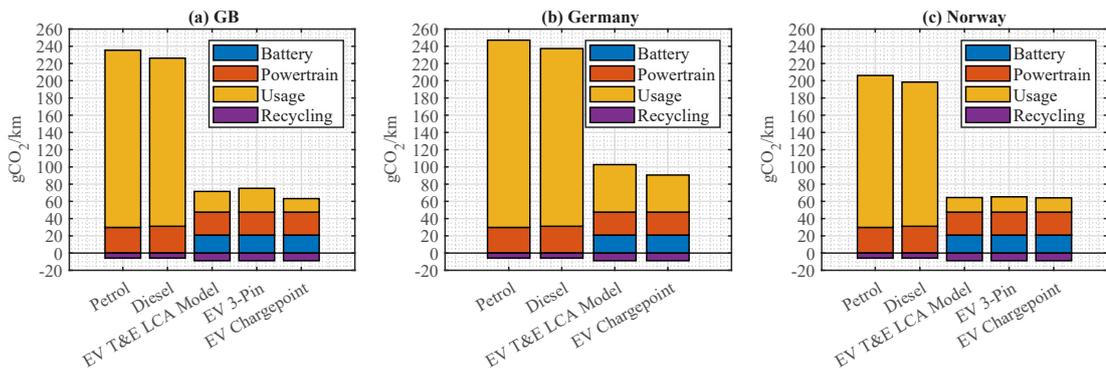
Figure 4.5: LCA assessment comparison for: (a) GB, (b) Germany and (c) Norway.

4.3.2.2   GB regional level

In Table 4.10, a summary of the results obtained for the two different scenarios for the GB is presented. In general, it can be observed that a dedicated chargepoint, with the ability to charge during the night hours at a faster pace, can greatly reduce the emissions when compared with the common 3-pin charger. In addition, England exhibits almost the same level of carbon footprint as the GB average, Scotland exhibits $\sim$16% and $\sim$27% reduced emissions in the first and second scenario, respectively, and lastly Wales exhibit an increased carbon footprint of $\sim$15% and $\sim$13% increased emissions

in the two scenarios. The difference between the regions of GB can be attributed to the different generation mixture, as well as to the different levels of energy imports. In Figure 4.6, the increase or decrease in a vehicle's lifetime emissions when compared with the GB average for the two scenarios is presented. This was calculated based on the carbon footprint per km as presented in Table 4.10. The divergence of the carbon footprint from the GB average is given as:

$$DV(reg, i) = (E_{reg,i} - E_{GB,i}) \times R \times 10^{-6} [\text{tCO}_2],$$ (4.6)

where $E_{reg,i}$ is the CO_2 emissions on region reg for scenario i in grams, $E_{GB,i}$ is the average CO_2 emissions of GB for scenario i in grams, and R is the range in km.

In GB, Scotland is the only country that exhibits a better-than-average carbon footprint. England falls slightly above average, whereas Wales exhibits the highest carbon footprint due to the increased usage of fossil fuels in electricity generation. In general, areas of Northern Britain — i.e. North-East England, North-West England, South Scotland, North Scotland and North Wales and Merseyside — exhibit higher levels of CO_2 savings due to higher penetration of RES, including wind and solar. On the other hand, areas of the South and South-East Britain — i.e., South Wales, East Midlands, South England, South-West England and South-East England — exhibit the worst performance.

4.4 Discussion & conclusions

During data processing and algorithm tuning, it was observed that, in the presence of houses with solar panels, it was better to extract EV load charge events without solar generation. EV load signatures have a distinctly high power level, and therefore, the drop in amplitude caused by solar generation is insufficient to completely obfuscate the EV signal. In the Dataport [52] houses considered in the study, EVs were connected to the grid in the evening and night hours, when solar generation is either very low or non-existent. This pattern agrees with the daily routines, as people tend to use their vehicles to commute to work during morning and afternoon — when solar generation

Table 4.10: Carbon footprint estimation of the usage parameter for the regions of the GB.

Region	3-pin charger (18:00 − 05:00) [gCO$_2$/km]	Dedicated chargepoint (01:00 − 04:00) [gCO$_2$/km]
GB	66.279	54.327
England	67.944	54.969
Scotland	48.464	45.400
Wales	75.946	61.181
South-East England	71.947	58.423
London	69.434	55.887
South England	80.018	62.526
South-West England	74.265	57.924
East England	65.217	52.866
East Midlands	82.851	62.321
West Midlands	67.118	53.526
South Wales	86.849	67.373
North Wales & Merseyside	56.865	49.500
Yorkshire	66.555	54.438
North-East England	43.657	41.874
North-West England	49.470	44.300
South Scotland	46.315	44.467
North Scotland	49.679	47.402

is at its peak. It is therefore worth exploring the possibility of storage of energy produced during the daytime and using that energy later to charge EVs and help reduce grid peaks that usually occur in the late afternoon/evening. Complementary metrics for measuring the accuracy in estimating the load consumption of the EV charging events are *Acc* and MR, whilst MAE and SAE can explain the performance of regression networks. Similarly, generalisation loss as a metric based on *Acc* and MAE, provide a good representation of performance loss of these measures due changes in granularity of the meter readings, as well as due to generalisability to unseen houses in a similar geographic area and with similar EV charging loads, and transferability to unseen houses in different geographic areas and with different EV charging loads. Although the DNN-regression approach successfully disaggregated the EV load under different scenarios, the ensemble classification models were more robust to insufficient EV charging events for training. That is, the seq2subseq DNN is especially sensitive to the number of training samples, which takes precedence over the noisiness level of
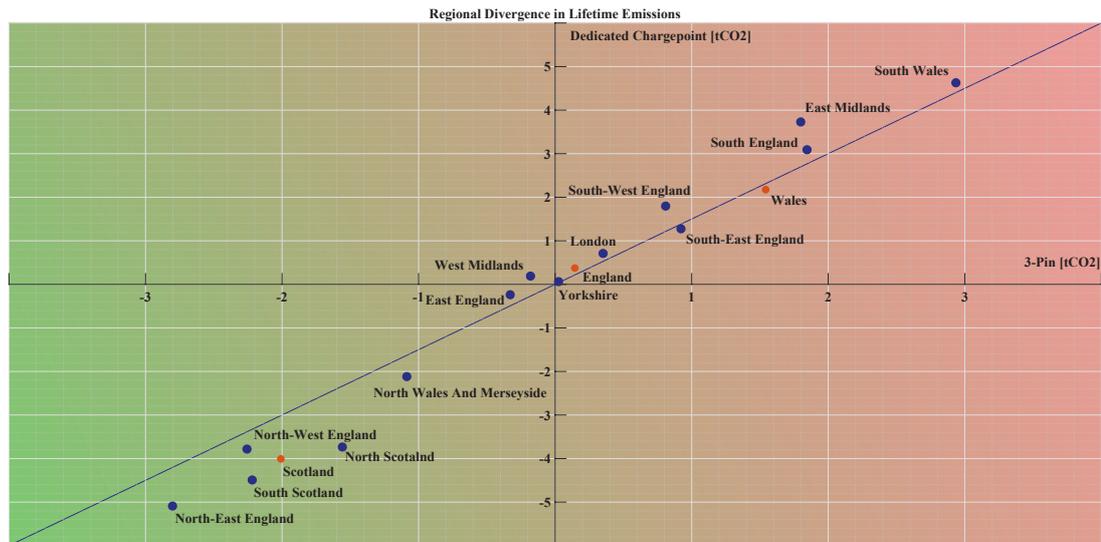
Figure 4.6: Regional Divergence in Lifetime Emissions When Compared with GB Average, Scenario I & II

the house due to interfering loads. Otherwise, it was observed that regression models accurately performed EV load estimation on the same house the models were trained on, as well as showed excellent generalisability performance when tested on unseen houses for similar EV charging levels in different geographic areas. During generalisability and transferability experiments, it was observed that the regression network is less affected by lower granularity readings than the RF classification and load reconstruction approach. The proposed final recommendation for EV charging event detection, as well as accurate energy consumption for each charging event, is therefore a seq2subseq DNN, when plenty of training data are available in a mixed mode approach, with data from different geographic areas, and especially with a balanced number of EV charging load levels to avoid bias towards a particular EV charging level.

This study demonstrated the feasibility of accurately estimating EV load charging consumption at scale by energy providers, using only smart meter measurements at resolutions of 1 minute and 15 minutes. Specifically, a regression network approach based on DL seq2subseq architecture with a conditional GAN was evaluated. Evaluation was carried out for three scenarios:

I training and testing on different portions of an observed house — observed scenario;

II generalisability across houses with similar geographic area and EV load charge; and,

III cross-domain transferability in unseen houses from different geographic areas and different EV load charge levels.

The merits of typically used regression and NILM-specific energy consumption metrics were presented for all experiments and discussed, in conjunction with generalisation loss metrics and NMs, which are an indicator of unknown loads interfering with the EV load in the aggregate meter readings.

Further to that, the application of the EV load disaggregation in conjunction with survey data has been demonstrated as an approach to increase the accuracy of the LCA models as well as the fairness of the comparison of different vehicular technologies. Although time of charge information can be retrieved from smart vehicle chargers, the approach taken relies only on smart metering data and thus reducing intrusiveness and the requirement to connect to individual endpoints/APIs, and increasing the scalability of the approach. Results, obtained for different charging routines as well as different users' locations, were presented. Particular attention was given to quantifying the effect of different geographical areas, both on a regional and national level, on the actual carbon footprint of each technology, a factor that can greatly affect the actual carbon footprint of each technology. In contrast with ICEVs, where refuelling timing does not affect the carbon emissions, EVs' charging routines can greatly affect the actual GHGs emissions during their lifetime. From the results presented in Section 4.3.2, lifecycle emissions per vehicle type can vary from -12.9% up to +3.8% considering the different users' charging routines. For GB specifically, as illustrated in 4.7, the usage component of LCA models for EVs vastly varies across the different regions of GB, which is a direct result of the different generation mixture present in each area and the time of charging. In fact, the usage component can vary by up to 100% for 3-pin chargers when comparing North-East England and South Wales (see Table 4.10.

A significant deviation between the carbon footprint of 3-pin chargers and dedicated chargepoints has been observe,d which is related to the time of charging of each technology, with dedicated chargepoints resulting in a reduced footprint for EVs. Although

Figure 4.7: GB – EV usage, regional carbon footprint, from dark green (lower) to dark red (higher).

there is an uptake of fast chargers, there is still scepticism and barriers towards the installation of dedicated chargepoints. More specifically, according to the latest EV smart chargepoint survey in the UK [90], the main barriers were: (i) too expensive to install a dedicated chargepoint at home (44%); (ii) a 3-pin cable is fast/easy enough to meet my charging needs (39%); (iii) not enough space/too complex to install a dedicated chargepoint at home (16%); and (iv) I don't have permission to install a dedicated chargepoint at my home (13%). Therefore, modelling of the lifecycle emissions of EVs should consider the existence of different charging technologies.

This is especially timely given the introduction of load-shifting initiatives throughout the world to reduce CO_2 emissions, as well as the introduction of smart chargers that can be programmed to charge during specific periods. In addition, in different parts of the world, initiatives that support the purchase of EVs through subsidies and withdrawals of ICEVs are being rolled out to reduce the countries' GHGs. Therefore, it is essential to first target specific areas that demonstrate the lowest carbon footprint per kWh of electricity, as in this cas,e the reduction of the GHGs will be faster as well as the compensation of the increased carbon unleashed during the production of an EV. Further

research should be carried out to expand in more regions around the world as well as in different end-users' profiles. In addition, as the EV market share is expected to rapidly increase with a plethora of different vehicles available to consumers, it is crucial to introduce novel recommender system solutions that will be able to identify the best candidate vehicle based on the bespoke requirements of an end-user.

Chapter 5

A complex mixed-methods data-driven energy-centric evaluation of net-positive energy households

5.1   Introduction & background

As many countries worldwide commit to net-zero goals, different approaches are being implemented to reduce carbon emissions, including the introduction of greener means

of transportation such as EVs, switching to RES and the establishment of carbon-neutral communities. Smart districts and local energy communities deploy housing that attempts to accommodate residents' needs while minimising the carbon footprint of living spaces. Besides good thermal insulation in the building design, energy-intensive routines of residents need to be considered when estimating their carbon footprint. Net-zero and even net-positive — i.e., the total energy production exceeds total energy demand annually — neighbourhoods that benefit from increased penetration of RES at the end-user level, together with digital smart home technologies (SHT) that can help implement energy conservation practices [99], are being implemented in different parts of Europe and UK. In this chapter, the focus is not on the design of net-positive energy residential communities but rather on the energy-centric evaluation of how truly net-positive a building is when considering the energy practices of its residents and how these are affected by SHTs, LCTs and a dynamic electricity pricing system.

A range of energy efficiency solutions and policy incentives, tailored towards energy conservation and mitigating the effects of climate change and reducing the economic cost to end-users, have been intensified following the Paris Agreement in response to the ever-increasing emissions of GHGs in combination with the turmoil in the energy markets. The impact of these solutions in the European Union (EU) can be seen in the report of the IEA [100] where an annual decrease of 3.5%, equivalent to 94.9 terawatt hours, of energy consumption in the EU was observed in 2022, leading to a reduction of 202 megatonnes of carbon emissions, compared to the global average increase of 1.9% in total energy consumption equivalent to 168 megatonnes of carbon emissions. In addition to the introduction of energy-efficient appliances and incentives, according to the 2023 consumer conditions scoreboard published by the European Commission [101], 72% of respondents believe that they need to personally do more to tackle climate change, and 57% are considering their environmental impact when purchasing goods and services. Furthermore, in the aforementioned report, it was shown that 71% of the EU population, including Norway and Iceland, changed their habits to save energy in line with the soaring energy prices, with Norway being one of the countries exceeding the average of the EU. A similar observation about the engagement of people in climate

policy actions is echoed in [102], where it was concluded that although there was a rise in negative sentiment following popular policy events such as the Paris Agreement, positive sentiment was more prominent in social media.

The reduction of CO_2 emissions through the rapid electrification of future urban buildings has been highlighted as an area of paramount importance for future study in [103] and of equal importance to the decarbonisation of the power sector, with net-zero and even net-positive energy buildings being introduced throughout Europe. However, although these houses are designed to reduce energy usage and carbon emissions, the actual energy consumption of the households is often higher than designed. The increased demand for energy services, such as high indoor temperatures, is the direct result of energy efficiency measures, such as better insulation. Consequently, people can afford to have higher temperatures due to the efficiency of their living spaces or because they become less attentive to savings, as they are aware that their household is more efficient. In some more rare cases, the increase in energy demand can also be said to relate to users' interaction with technology, such as the user interface and sociotechnical mismatch effects. The first occurs when households replace their appliance with a smarter one and do not know how to use it, adjusting the device poorly and consuming more energy. The second occurs when the technologies work efficiently only when they are operated as designed; however, they do not fit with households' everyday lives. Indeed, the energy performance gap [104], between actual energy requirements of lived-in buildings compared to expected energy consumption — according to standards such as ISO 16343:2013 [105] — has been attributed to different factors, including unrealistic occupants' behavioural assumptions and unpredictable usage habits [106].

Therefore, it is imperative to understand and quantify the deviation between actual and predicted energy consumption and to explore energy efficiency approaches that take into account the practices and routines of the end user. Such approaches include more accurate predictions of expected energy consumption and lead to solutions that can help end-users reduce their energy bills and carbon footprint through flexibility in their routines. This can take the simple form of shifting flexible loads to maximise RES generation and decongest the grid at peak demand, which in turn reduces wholesale prices

and dependency on non-renewable generators to meet demand. Flexible load shifting curtails the peak demand, avoiding the use of fossil fuels to supplement renewables; for example, climate change has forced UK emergency coal power plants to be used with hot [107] and cold [108] temperature, mostly led by the extravagance of using heating and cooling appliances even if the temperature is not extreme [109].

While the understanding and prediction of energy consumption in households has been the subject of numerous studies [110], these generally focus on qualitative [111] or quantitative [112] data analysis. In a review of different approaches for energy research design methods [113], the importance of bridging qualitative analysis — which can offer great detail and high explanation but with limited capabilities in scaling — with quantitative analysis — which can easily scale up but may lack in explanatory power — is highlighted. Following a critical review [8] of how building energy efficiency is affected by occupant behavioural patterns (considering occupational behaviour, energy efficiency, conservation, and consumption analysis), it was concluded that in most research, holistic approaches are not employed but tend to be focused on a singular area of interest such as ventilation and heating. Similarly, a review of over 200 articles, of which about 83.48% focused on quantitative data with predominant usage of basic statistical approaches on energy behaviour of households [114], highlighted the need for mixed-methods research on building energy consumption to provide insights not only on "what" is being consumed but also "how" and "why". These review articles make the case that energy-related mixed-methods approaches are needed but still in their infancy, with no specific framework in place to better analyse occupant lifestyles that can lead to a better understanding of user profiles and routines, and hence improved energy efficiency recommendations.

Mixed-methods research, that is, combining quantitative and qualitative data collection and analysis in one study, was introduced as a means to reduce bias — as a result of only quantitative or qualitative research — and improve the robustness and depth of research findings by neutralising the weaknesses of each type of data [115], and can be categorised as:

- Exploratory sequential mixed-methods, where the research first focuses on the

qualitative analysis and the quantitative data are used in order to provide more detailed explanations;

- explanatory sequential mixed-methods, where the research first focuses on the quantitative analysis and the qualitative research is used to provide more detailed explanations;

- convergent mixed-methods, where quantitative and qualitative data — that are collected approximately at the same time — are merged as a way to analyse a problem; and,

- complex designs with embedded core designs, where a primarily quantitative or qualitative design can be intersected with a secondary method, or a mixed-methods design can be intersected within another methodology or within a theoretical framework.

In the context of solely quantitative energy disaggregation and load scheduling research, bias can be introduced during the development of the disaggregation models through the selection of the training dataset or through assumptions made regarding the availability of appliances and the level of accepted flexibility from the end user. Therefore, in order to limit bias, a complex mixed-methods approach to propose a framework to jointly understand the "what", "how" and "why" of energy consumption in net-positive dwellings is adopted. A brief review of mixed-methods approaches in the literature and identified gaps that are addressed via the proposed framework is presented next.

A mixed-methods clustering approach for energy data using quantitative survey data — variables related to energy and socioeconomics — and qualitative codes associated with transcripts from interview data was proposed in [116], whereby a two-step process was followed. First, quantitative and qualitative data were clustered separately, and secondly, links between the clusters were identified. Clear links were identified that can unlock findings that would not have been possible analysing only quantitative or qualitative data, such as households that exhibit the same energy consumption but have completely different socio-economic characteristics and different levels of awareness about clean energy. In [117], via case studies in Spain and in the Benelux, a mixed-methods design process was proposed, integrating occupant behaviour and attitudes towards energy use

and indoor conditions. Although quantitative parameters such as temperature, relative humidity, CO_2 levels, and parameters such as sound, light, and movement were used, actual energy consumption was not analysed. Based on the practices of the occupants, profiles were generated and compared with the average profiles used in simulations and energy regulations using an embedded design in order to explain and validate quantitative analysis through qualitative data. The importance of occupant comfort and "convenience and time" was highlighted as a major parameter that affects actual energy use in a household.

Different approaches have been proposed in providing activity load consumption and feedback to end users, with the majority of them jointly analysing qualitative smart meter/submetered data and qualitative sociodemographic data to produce more meaningful feedback through in-home displays (IHDs) or mobile apps. In [118], a mixed-methods convergence approach, using qualitative electrical energy measurements from sub-metering devices and smart meters together with demographic data, was proposed to quantify the energy intensity and temporal routines of occupant activities, leveraging on quantitative NILM research and qualitative practice theory research. In [119], different methodological approaches including analysis of large databases, surveys, qualitative interviews, indoor measurements and electricity readings, combined with surveys and qualitative interviews, showed that people's intentions are not mandated by the amount of energy they consume, but by the domestic activities they engage in, such as regulating indoor climate, cooking and laundry. An exploratory mixed-methods approach was implemented to understand energy consumption after IHDs installation in [120], with quantitative analysis performed first with the objective of quantifying the change in energy consumption before and after IHD installation. Qualitative analysis was then performed to understand the reasoning behind the reduced energy consumption identified through the energy data and, therefore, to explain why energy consumption was statistically significantly lower than before IHD introduction.

Though previous work reviewed above has demonstrated the value of mixed-methods approaches to reduce bias in findings of pure qualitative or quantitative research for understanding energy demand, there is still a gap in the literature in explaining energy

Chapter 5. Energy-centric evaluation of net-positive energy households

consumption patterns in homes to using this understanding to improve energy efficiency measures. Indeed, most prior work reports occupant energy use patterns, occupant-building interactions, and uncovering relationships between behaviour and influencing factors, without relating to explain the "why" and "how". Therefore, in this chapter, the hypothesis that mixed-methods analysis would provide the tools to explain from the "what" to the "how" and "why" of end-user energy consumption to directly inform energy efficiency initiatives is made. To this end, the main contribution of the chapter is a complex mixed-methods methodology intersecting quantitative load disaggregation methods from granular smart meter data, quantitative cost reduction analysis from dynamic pricing profiles and qualitative analysis of interviews and questionnaire data from state-of-the-art net-positive/plus buildings. Specially, the methodology answers the following:

- "what" is the energy gap between energy consumption and RES production of plus-home living spaces and "why" this gap arises, exploring through granular smart metering data the extent of the energy gap and through qualitative data the reasoning behind it;

- "what" is the deviation between actual energy consumption and net-positive energy balance, and "how" this can be explained through the lens of household routines, intersecting interviews and time-of-use questionnaires to improve activity-level disaggregation;

- "what" is the deviation between actual and expected energy bills, "why" net-positive houses exposed to dynamic electricity pricing do not always have a zero bill, and "how" this can be explained through time of use tariffs in relation to their energy-intensive activities, explaining quantitatively the cost deviation through the lens of qualitative findings related to practices and motivations of end-users; and,

- "what" are the insights gained on user-centric load shifting potential, "why" they are suited to the user based on their routines, and "how" load shifting is actionable when aligning with dynamic energy pricing, as a means to reduce

CO_2 emissions, estimating quantitatively the load shifting potential through the lens of disaggregated activities and exploring qualitatively the motivations and acceptance of flexibility by the end-users.

In order to answer the above questions and to perform a systematic analysis of energy prosumption in these new emerging urban areas to evaluate their social and technical impacts, a combination of qualitative and quantitative data from a PEN were required. Although a number of load consumption datasets exist, these are usually collected from consumer-only households that are geographically distributed throughout a country without forming a single neighbourhood (the result of a recent review of load consumption datasets is presented in Table 1 of [121]). Therefore, in this chapter, apart from the methodology to answer the "what", "why", and "how" of energy prosumption, the data generated from the case study of a Norwegian PEN that enables such systematic analysis by bringing quantitative and qualitative data together is included. NorPEN dataset, stands apart from other available datasets and enables this sort of research, as:

- It is the first dataset of households located in a designed PEN;

- besides energy consumption and disaggregated load, estimated energy production, voltage, and current data are also provided in a sampling rate of 10 seconds;

- the utility billing power consumption vector and the variable tariff vector are provided in a sampling rate of 60 minutes;

- weather parameters that can affect consumption and RES production, including temperature, humidity, and solar insolation; and,

- qualitative data, including in-depth semi-structured interviews and time-of-use surveys with the homeowners, are included in the dataset.

As the design and construction of net-positive buildings and neighbourhoods require a multidisciplinary approach, involving architects, engineers, energy experts, sustainability design professionals, and social scientists, it is expected that this dataset will be valuable to stakeholders involved in the design process of net-zero and net-positive buildings and

districts, by facilitating quantitative and qualitative energy analysis, perceived vs actual energy efficiency of designed positive energy districts (PED), as well as to stakeholders involved in the study and deployment of smart microgrids with RES penetration.

The rest of the chapter is organised as follows. In Section 5.2, the complex mixed-methods approach is explained, comprising a quantitative and qualitative data collection process, the estimation of renewable energy production, the disaggregation of the activities and the exploitation of the energy price information. This is followed by Section 5.3 where the mixed-methods evaluation approach and the key findings are presented as per the above four questions. Lastly, in Section 5.4, the key conclusions, including limitations of the study and future work directions, are discussed.

5.2 Methodology
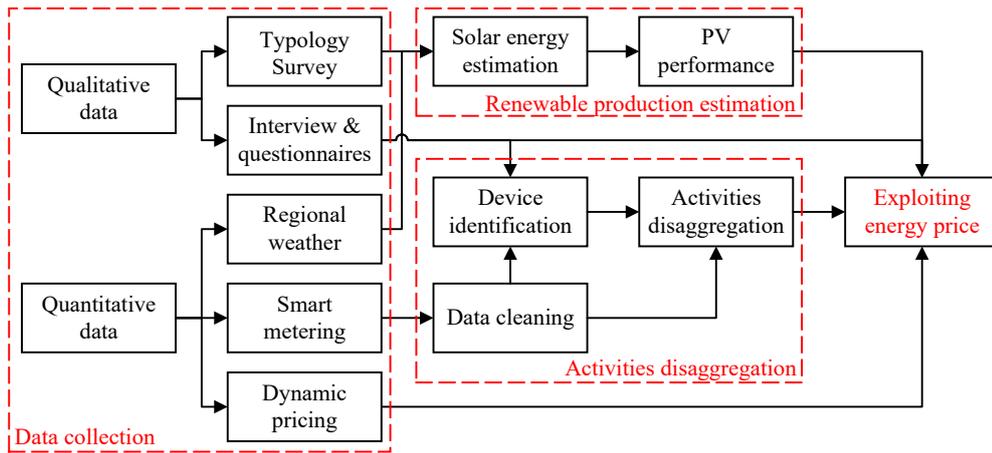
Figure 5.1: Mixed-methods approach flowchart showing building blocks of the overall methodology adopted
.

In order to quantify the energy gap between energy consumption and production in a net-positive dwelling, explain the deviation through the lens of disaggregated activities and deviation between actual and expected energy bills, the overall methodology of Figure 5.1 is followed, where each of the blocks is described below.

5.2.1   NorPEN dataset collection methods

The households participating in this study were recruited in a neighbourhood in Eastern Norway (Østlandet), which is within the general concept of the PED/PEN, which houses approximately 70 middle-income families. Quantitative and qualitative data (more information regarding the validation and quality of data are available in Appendix A) were simultaneously collected from the pilot project. The new urban area consists of several housing zones that are not yet fully developed. In this study, a zone built between 2018 and 2019 with buildings having a range of different typologies (including detached, semi-detached and flat-apartments) and different sizes (ranging from under 100 sqm to approx. 200 sqm) was targeted. The total area of the neighbourhood of the aforementioned households is approximately 130 metres wide by 325 metres long. The houses were designed to meet all their energy demands through electricity and in an environmentally friendly manner, meeting passive house standards, equipped with solar panels, ground source heat pumps for space heating and domestic hot water (DHW), and smart home technology, including a smart energy management system. In this smart district, in contrast to standard practice, the solar panel installer buys the energy surplus without deducting the network tax that is being paid to supply the grid with power. Thus, each homeowner has their consumption settled against their share of the production, and, therefore, are getting paid the actual amount of money that their panels produce. Further to that, all households have an EV or a PHEV. Some households have a dedicated EV fast charger, while others rely on generic 3-pin chargers due to the additional costs of installing a dedicated chargepoint. A ground-source heat pump system was installed at the community level. In addition, passive house standards were also taken into account during the design of the houses.

A door-to-door canvassing recruitment process [93] was conducted throughout 10 days. During April 2022, 9 in-depth, semi-structured, face-to-face interviews on energy practices assisted by SHTs were conducted with one or more householders — in one specific case including the presence of teenagers during the interview. More information regarding the structure of the interviews is presented in [93]. The full

interviews are included in the NorPEN dataset [122]. The 9 households were selected for this study based on two criteria: different demographic characteristics (i.e., age, sex, educational background, and occupation) and housing typology. All selected households were equipped with three-phase installations, either fully BEVs or PHEVs, with some households also having a fast charge point installed. Interviews were recorded, transcribed and analysed through traditional coding and content analysis techniques [123]. As described in [93], which uses the same data as this study to explore social practices with respect to energy use, data saturation was swiftly achieved for three main reasons. First, semi-structured interviews enable the exploration of the same questions with all participants. Second, the homogeneity of the sample in terms of housing characteristics, appliance type, make and availability, access to smart technologies and EVs, as well as prosumers scheme. Third, the qualitative and quantitative data triangulation strengthen the reliability and validity of the study. Adhering to the general data protection regulation (GDPR) guidelines, written consent declarations were obtained to collect, process, and publish data after anonymisation for a period of 2 months, stretching from mid-February to mid-April 2022. Out of the 9 households, due to connectivity issues, data were collected from 6 households, more specifically houses 1, 3, 5, 6, 7 and 9. Table 5.1 presents a summary of the selected households, their typology, and their characteristics.

The monitoring period was selected as it is during spring and spans almost evenly before and after the northward equinox, with daylight ranging from approximately 8 to 14 hours for the whole monitoring period. As the district of Eastern Norway is located north of the Tropic of Cancer and in close proximity to the Arctic Circle, it exhibits very short days during the winter period — as low as 2 hours per day — and on the other hand, extremely lengthy days during summer — exceeding 18 hours.

Although it is challenging to accurately estimate the consumption and production profile of a household from a two-month sample, the period was selected to minimise the intrusiveness to the house occupiers and maximise the extracted information. More specifically, in the monitoring period, the solar irradiation and the temperature are neither extremely low, as would have been the case around the southern solstice, nor

extremely high, as would have been the case around the northern solstice. As can be observed in Figure 5.2, the monitored period spans the linear area of optimal solar production, and the monitored period average deviates by approx. 5% of the yearly average. Further to that, the temperature during the two-month period varied between -8°C and 19°C, with the yearly variation being between -15°C and 30°C and the average temperatures between -4°C and 17°C for the whole year. Therefore, the temperature range during the monitored period is close to the yearly average. Since heating and hot water demand are correlated with outdoor temperature, a monitoring period close to the yearly average provides a realistic estimate of the yearly average. In addition, the monitored period contains both periods of normal working days and a week of school holidays — during February — which usually affects the energy consumption as people tend to travel during breaks. Therefore, the monitored period can be considered a representative sample both for production and consumption.
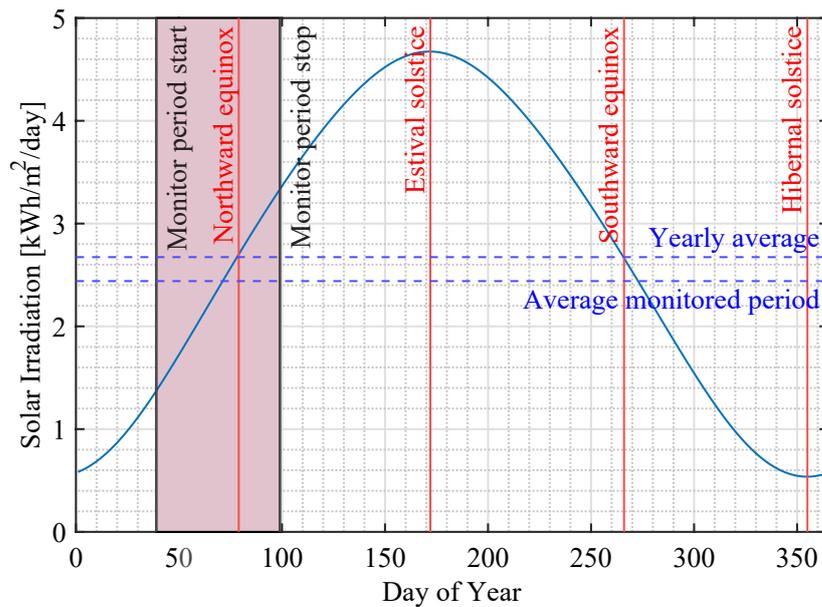
Figure 5.2: Average maximum solar irradiation on the PV panels located on the roofs of the buildings of the neighbourhood under study, calculated as described in Section 5.2.3. Longitude and latitude considered are limited to satellite data granularity (0.5 × 0.5 km) with the coverage of the neighbourhood (0.130 × 0.325 km). PV panels are installed in a circular pattern with different tilts per house relative to the sun's position.

A summary of the houses involved in the study is presented in Table 5.1. The

97

description of the occupant profiles is based on the way that the homeowners self-identified during the aforementioned in-person interviews.

Table 5.1: Summary of households.

ID	Type	Area	Floors	Rooms	PV [kW]	Azimuth (tilt)	Occupancy (age)
1	Detached	$148m^2$	3	4	4.8	$65^o/-115^o$ (15°)	2 adults ($>$60) & 1 dog
3	Detached	$193m^2$	3	4	6.4	$35^o/-145^o$ (15°)	2 adults ($\approx$34) & 1 child
5	Semi-detached	$90m^2$	2	2/3	14	40° (30°)	2 adults ($\approx$30) & 2 children
6	Detached	$148m^2$	3	4	4.8	$85^o/-95^o$ (15°)	2 adults ($>$60) & 1 teenager
7	Semi-detached	$90m^2$	2	2/3	14	40° (30°)	2 adults ($>$38) & 2 teenagers
9	Semi-detached	$131m^2$	2	3	9.6	$-30^o/150^o$ (15°)	2 adults ($\approx$32) & 1 child

5.2.1.1 Household aggregate readings

The household aggregate readings were collected via the energy provider through smart meters installed within the households. The data were securely transmitted from the households to the utility provider and then, through Azure, transferred to the Server located in Glasgow, Scotland. Smart metering data were collected from 6 households for a period of two months (2022-02-09 23:00:00 – 2022-04-09 22:00:00 (UTC)). Three-phase power supply is installed in all households, and therefore the smart metering data — sampled at 10-second intervals — contain information about the total active and reactive power as well as the voltage and the current consumption on a per-phase basis. Due to technical issues, it was only possible to collect smart meter data from 6 households. The aggregated active/reactive import/export power and the current and voltage of each phase were collected. The voltage readings correspond to the potential difference between each phase and the neutral line. Readings that failed to be transmitted were discarded from the smart meter, and therefore, subsequent readings do not contain information about the non-transmitted readings. Although all households have PV panels installed, these are wired in a separate circuit, and therefore, there is no solar interference in the collected data.

5.2.1.2 Utility billing vector & hourly tariff

Apart from low-frequency readings (10 seconds), the utility provider collected an hourly sample for billing purposes. The hourly sample included the cumulative active and reactive import and export energy, with the active energy samples measured in Watt-hours [Wh] and the reactive energy samples in Volt-Ampères-Reactive-hours [VARh]. These variables are used by the utility provider to bill the end-users. The transmission and collection of these readings are more robust as these are required to meet the utility requirements. Therefore, these readings can be used to estimate and interpolate missing values in the dataset. The hourly price vector of electricity in the region where the households are located is also included in the dataset. The price vector in Norske Krone (NOK) reflects the price of the energy consumed in the past one hour without the inclusion of the Value-Added Tax (VAT). Figure 5.3 shows the hourly price vector for the monitored period. At the time of data being collected, the VAT rate was 25%,
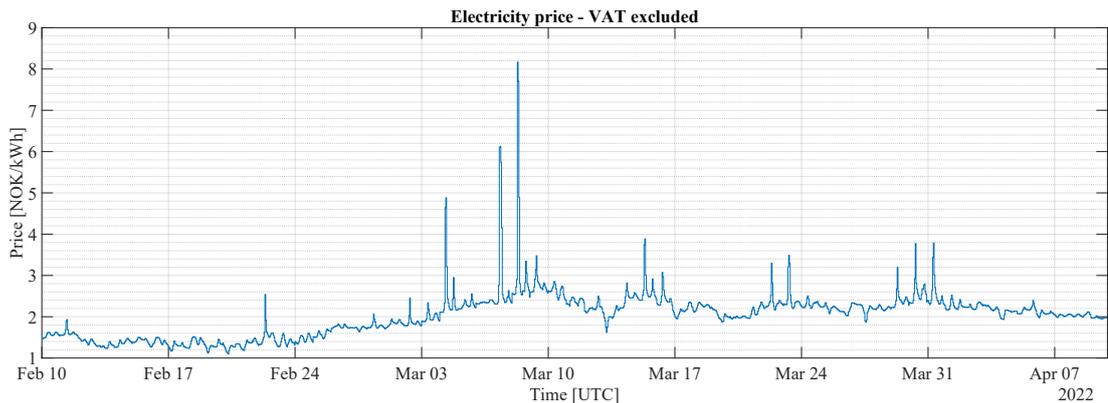
Figure 5.3: Hourly electricity price for the monitored period, VAT excluded. ©2024 IEEE

the grid fees for importing energy from the grid were approx. 0.4 NOK/KWh and the compensation for providing energy to the grid when exporting was approx. 0.1 NOK/kWh. Note that VAT is charged only when importing energy from the grid and not when energy is exported to the grid.

5.2.1.3 Cleaning of smart meter readings

Collected meter readings occasionally suffer from gaps, which need to be filled. To facilitate the interpolation process and estimate the quality of the data, a Quality Index was calculated based on the length of the gaps. The Quality Index (QI) is given by:

$$QI(l) = \frac{S_{expected} - S_{missing}(l)}{S_{expected}}, \tag{5.1}$$

where $S_{expected}$ are the total number of samples expected (10-sec samples for a period of 2 months, i.e., 509,400 samples) in the dataset and $S_{missing}$ are the total number of consecutive samples that are missing with a duration less than l, where l is the length of the gap. Figure 5.4 represents the QI (Equation 5.1) for the 6 households with smart metering data. For each household, the number of missing samples that exceeded a certain duration was calculated. This step is considered necessary as the quality of the activity disaggregation results is related to the quality of the submetered data. Gaps in the data that spanned less than 1 hour were replicated using the nearest-neighbour interpolation method under the constraint that the total consumption during that hour should be equal to the difference between the two billing measurements, i.e., the total energy consumed during that hour. Gaps that spanned for more than one hour were filled based on average historical data, i.e., the average of the consumption on the same day of the previous weeks — using again the constraint that the total energy consumption per hour should be equal to the billing energy power.

5.2.1.4 Weather data

The weather data were collected from The Norwegian Meteorological Institute [124], including the following variables: air temperature [^{o}C], relative humidity [%], surface pressure [hPa], precipitation [in mm/h], wind speed [m/s] and wind direction [deg] sampled at 5-min intervals.

The dew point T_d was calculated through the vapour pressure and saturation vapour
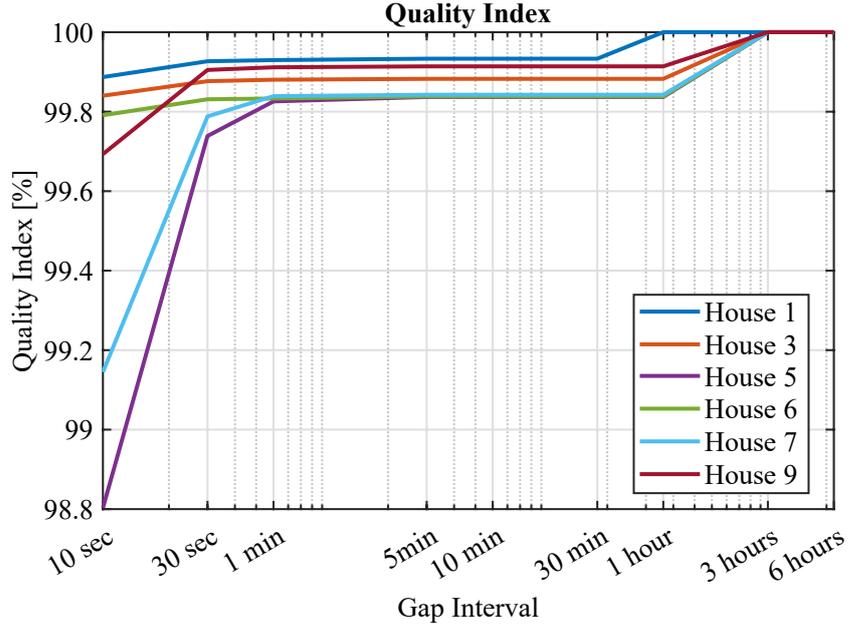
Figure 5.4: Quality index of continuous smart meter data samples for each of the 6 households included in the study, highlighting gap intervals that needed to be filled before calculating energy consumption.

expression of the relative humidity as:

$$RH = 100\% \times \frac{E}{E_s}, \tag{5.2}$$

where, based on the Clausius-Clapeyron [125] relation, the vapor pressure is given by:

$$E = E_0 \times e^{((L/R_v) \times (1/T_0 - 1/T_d))}, \tag{5.3}$$

and the saturation pressure by:

$$E_s = E_0 \times e^{((L/R_v) \times (1/T_0 - 1/T))}, \tag{5.4}$$

with the saturation vapour pressure $E_0 = 0.611$ kPa, the latent heat of vaporisation $L = 2.453 \times 10^6$ J/kg, the gas constant for moist air $R_v = 461$ J/(kg×K), $L/R_v = 5423$ K, $T_0 = 273.15$ K and T being the air temperature. By solving for the dew point, T_d, it

101

will be given in Kelvin by:

$$T_d = \frac{1}{\frac{1}{T} - \frac{L}{R_V} \times ln\left(\frac{RH}{100\%}\right)}.$$ (5.5)

The dew point was converted to Celsius by subtracting the constant 273.15 from the Kelvin temperature. Temperature, dew point, and relative humidity for the entire monitored period are presented in Figure 5.5.
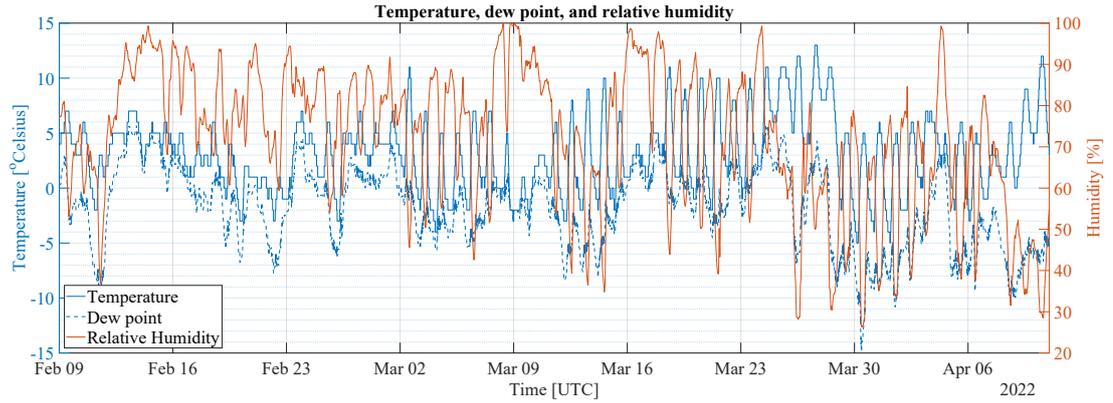
Figure 5.5: Weather data for the monitored period. ©2024 IEEE

The solar data were generated using Copernicus climate change service information 2024 [126] in 1-min intervals[1]. The collected data were the global (GHI), the beam (direct) (BHI), and the diffuse (DHI) solar irradiance, the beam (direct) normal irradiance (BNI), the cloud coverage, the cloud type and the albedo. A sample of the collected data are presented in Figure 5.6. The zenith angle of the solar disc was calculated as:

$$\theta_Z = cos^{-1}\left(cos(\phi)cos(\delta)cos(\omega) + sin(\delta)sin(\phi)\right),$$ (5.6)

where ϕ is the latitude, δ is the declination of the Sun and ω is the hour angle. The declination [127] of the Sun (δ), with a range $-23.5^o \leq \delta \leq 23.5^o$, is given by:

$$\delta = \Phi \times cos\left(\frac{C(d - d_r)}{d_y}\right),$$ (5.7)

---

[1]Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains.

where Φ is the tilt angle and equal to 23.5^o, $C = 360^o$, d is the Julian day, d_r is the Julian day for summer solstice (equal to 172 for non-leap years), d_y is the number of days per calendar year (i.e., 365 days or 366 days for leap years). The hour angle, ω, is given by:

$$\omega = 15^o \times (t - 12), \tag{5.8}$$

where t is given by:

$$t = hours + minutes/60 + seconds/3600. \tag{5.9}$$
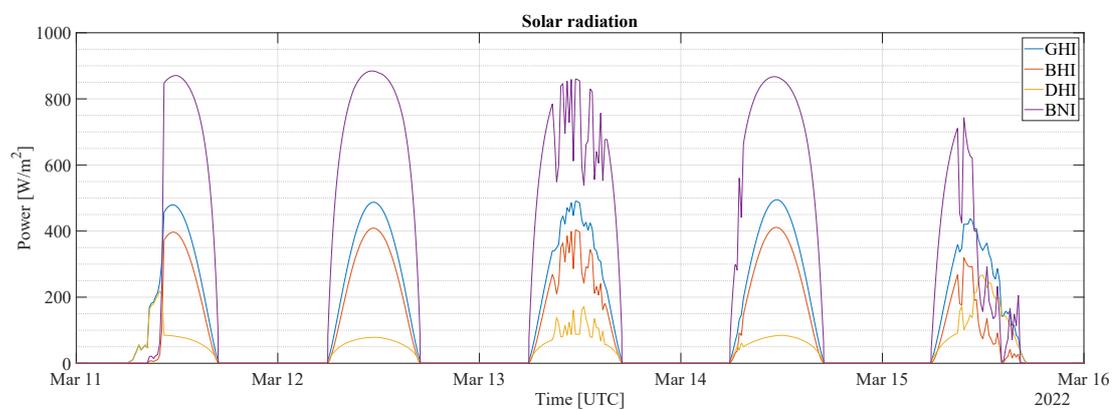
Figure 5.6: Solar radiation data sample. ©2024 IEEE

5.2.1.5 Qualitative data

The qualitative data consists of in-depth semi-structured interviews on household energy practices and a time-of-use of electric appliances survey. Interviews were conducted face-to-face over 10 days in April 2022, simultaneously with the recruitment process. They lasted an average of 67.5 minutes and were audio-recorded and transcribed *ad verbum*. Minor language editing was performed after the transcriptions, considering that the interviews were conducted in English, though neither the researcher nor the participants were native English speakers. Due to the semi-structured character of the interviews, the questions were arranged into six themes:

- "walking through" the smart homes and smart apps to understand how technologies

mediate energy practices;

- motivations for buying a smart home in the new neighbourhood, and,

- motivation for buying an EV, both to uncover meanings and ways of engagement;

- understanding changes in energy practices due to new materialities — old house versus new house;

- heating and cooling practices; and,

- sociodemographics.

The themes of the interview guideline were drawn from previous studies on household energy practices within the theoretical framework of social practice theories. In this sense, the interviews aimed to go beyond the traditional occupants' behaviour and lifestyle approaches and focus on variations of energy practices (individual energy-consuming habits and routines) that are rooted in collective socio-material structures [119]. Interviews were conducted with all 9 households; however, only 4 interviews are included in NorPEN dataset, namely households 1, 3, 5 and 9. As interviews with households 6 and 7 included extensive sensitive information, anonymisation of the interviews so that they could be understood was not feasible. In the period following the interviews, a survey on the time of use of electric appliances was sent by phone message or email to two households, selected due to extreme cases of the ratio of production to consumption. A time-of-use survey was developed based on [128] and consisted of:

- type of appliances;
- frequency of use;
- time-of-use during weekdays;
- time-of-use during weekends; and,
- appliances for long-term illnesses.

This detailed information on households' energy-consuming routines and habits in relation to appliances contributed to mapping hourly usage patterns and validating load disaggregation results on these households. Based on the validated load disaggregation results of these two households, the knowledge of the electric signatures and of the household routines was then transferred to the remaining households under study.

5.2.2 Disaggregation of activities from smart meter readings through transfer learning

Sub-metering devices used to measure energy consumption at the appliance level were not installed in the monitored households. Therefore, energy consumption on a per-appliance basis is estimated based on the total energy consumption and validated through soft labels from the qualitative data analysis — i.e., interviews and surveys — as well as through the quantitative data. For example, Sofie (house 9) discussed her vehicle's charging patterns:

> I guess it would usually probably be around late afternoon evening is when we would be charging it. When we're going out for the day. (Sofie, 32-years old, house 9)

a fact that was cross-validated from the questionnaire and the actual load data. Different ML models have been used in the literature for the load disaggregation problem (see Section 2.4). A seq2subseq model (see Sections 3 & 4) and a WaveNet model [129] were shown to effectively perform the load disaggregation task by transfer learning from publicly available datasets, and are used for the disaggregation of appliances of the households under study.

As the aforementioned models are based on supervised learning, training data are required to develop the models. Therefore, publicly available data sets were used to train load disaggregation models. Based on the interview data and questionnaires, the installed appliances were identified and the most adequate datasets, which contain similar appliances, were selected (see Subsections 2.4.2 & 2.4.3). More specifically, ECO [4], REFIT [56] and PECAN [52] datasets as well as the EV consumption dataset in [94] were used. ECO dataset contains three-phase residential smart meter data as well as sub-metering of 6 households for a period of 6 months with a sampling frequency of 1 Hz. The ECO dataset was considered adequate, as it contains similar installations — i.e., three-phased ones — and similar appliances to the ones targeted in the research. A summary of data availability in the ECO dataset is presented in Table 3.1. The REFIT dataset [46] contains smart meter data as well as sub-metering of 20 households for a

period of 21 months with a sampling frequency of 1/8 Hz. As with the ECO dataset, the REFIT dataset was considered adequate as it contained a variety of different households with several different appliances that were similar to the ones targeted. PECAN dataset includes EV loads from several households in Texas and NY area with a sampling rate of 1 Hz (see Table 4.1). Finally, the EV dataset presented in Subsection 3.3.2 that contains data from one year of a household in Germany, where a high-power EV charger — i.e. 11kW — is installed with a sampling rate of 1/60 Hz, which coincides with the presence of similar EV chargepoints in the smart neighbourhood that is being studied was used.

Publicly available datasets were resampled at the same sampling rate as collected data. As the targeted households had a sampling rate of 1/10 Hz, the other datasets used were down-sampled or up-sampled to the same rate. As ECO [4] and PECAN [52] were sampled in 1-second intervals, downsampling was performed by aggregating the energy consumed during each 10-second period. REFIT [46] dataset, which had a sampling rate of 1/8 Hz, could not be directly resampled as the data are required to be down-sampled by a non-integer. Therefore, the data were resampled at the new lower rate by interpolating the values. Finally, the EV dataset in [94], which has a sampling rate of 1/60 Hz, was up-sampled by assuming the same power level throughout the 60-sec period. Table 5.2 contains a summary of the households and appliances used for training from the open-accessed datasets.

Based on the interview data, households were split into two categories, the ones that had a high power EV charger — i.e., a dedicated charger with a nominal power of 11kW — and the others that used a portable EV charger (3kW) that plugs into a standard residential socket (esp. for PHEV). More specifically, regarding their charging routines, Brian (house 1) stated that a dedicated charger capable of being programmed is installed in his household:

> Yes! I have programmed my charger to start at 1 o'clock at night because it's when the energy is cheaper. So, I always charge my car at night. (Brian, 61-years old, house 1)

On the other hand, Sofie (house 9) stated that they have a PHEV:

Table 5.2: Households used for training of the models. ©2024 IEEE

Target loads	REFIT [46]	ECO [4]
Heating	1, 9, 16	–
WM	1, 6, 8, 9, 18	1
Tumble dryer (TD)	1	1
Washer-dryer (WD)	9, 18	–
DW	1, 6, 8, 9, 18	2
Electric hobs (HBs)	–	2
Electric oven (OV)	–	2
CM	–	1, 3, 5, 6
KET	6, 8, 9	1, 2, 3, 5, 6
MW	6, 9, 18	4, 5
FRD	8, 18	–
FRZ	6, 8, 18	1, 2, 3
Fridge-Freezer (REF)	18	1, 2, 3
EV	**PECAN** [52]	**Dataset** [94]
Low-power (AU)	661, 1642, 4373, 6139, 8156	–
Low-power (NY)	27	–
High-power	–	1

We have plugin hybrid [. . .][2] 50 Kilometers. Mm-hmm. And then after
that it goes on to gas. But we don't really use gas that much cause we don't
go very far. (Sofie, 32-years old, house 9)

with a standard 3-pin socket system installed in their property due to the cost of getting
a dedicated charger:

I was looking into that one. The prices were starting to get expensive,
then they became expensive all the time instead. . . (Sofie, 32-years old, house
9)

Two different models were used for these two groups. The same procedure as

---

[2]Note that [. . .] is the ellipsis symbol that denotes an intentional omission of a word, sentence, or
whole section from a quotation from interview data without altering its original meaning.

in Chapter 4 was followed for the training of the EV models. The model used for the disaggregation of high-power EV charger loads was trained on the household from Subsection 3.3.2 as it showed a similar load profile. On the other hand, the disaggregation of the rest of the EVs was based on a model trained on a selection of households from the PECAN dataset [52] that exhibited a similar low-power charging level, i.e., houses 661, 1642, 4373, 6139, 8156 from Austin and house 27 from NY (see Table 5.2).

As the rest of the household appliances exhibit a more complex signal, compared to the relatively high power and long duration EV charging, a WaveNet network was used to estimate their load. The training data set consisted of a mixture of different households from the REFIT [46] and ECO [4] data sets that contained the same appliances (see Table 5.2). The targeted appliances were the most commonly used high consumers — as identified through the questionnaire: HT, WM, TD, WD, DW, HB, OV, CM, KET, MW, FRD, FRZ, and, REF.

More specifically, from the REFIT dataset the following houses were used for training the models:

- house 1 (WM, TD, DW, HET);
- house 6 (FRZ, MW, KET, WM, DW);
- house 8 (FRD, FRZ, KET, WM, DW);
- house 9 (MW, KET, WM, WD, DW, HET);
- house 16 (HET); and,
- house 18 (FRD, FRZ, REF, MW, WM, WD, DW),

From the ECO dataset the following houses were used:

- house 1 (CM, TD, REF, FRZ, KET, WM);
- house 2 (DW, REF, FRZ, KET, HB, OV);
- house 3 (CM, REF, FRZ, KET);
- house 4 (MW);
- house 5 (CM, KET, MW); and,
- house 6 (CM, KET).

Based on the interviews collected in households, as well as the appliance availability

and time-of-use survey, the appliances were grouped into different household routines, taking into account different activation times. The routines were based on known relationships between activities and the appliances used in those activities to connect quantifiable data on appliances with the range of activities that define daily life at home [118]. Energy-intensive activities were taken into account and grouped into the following categories: breakfast, lunch, dinner, laundry, cleaning, heating, refrigeration, and vehicle charging. Breakfast, lunch, and dinner were further grouped into the cooking practices activity, and laundry and cleaning were also grouped into a single category. The identified routines, with the corresponding time windows and appliances, were:

- EV charging (EV): all-day;
- heating (HT): all-day;
- refrigeration (FRD, FRZ, REF): all-day;
- laundry/cleaning (WM, TD, WD, DW): all-day;
- breakfast (KET, CM): 05:00–10:00;
- lunch (HB, OV, MW): 10:00–15:00; and,
- dinner (HB, OV, MW): 15:00–21:00.

Appliances that can be used during different activities were grouped based on time-of-use. The amount of energy consumed in a household that was not a result of the aforementioned appliances/activities is considered as a non-disaggregated load and presented as a separate activity, namely "Other".

Through the combination of quantitative data analysis and interviews, energy consumption on a per-activity basis was further explained. For example, the heating practices of the homeowners were explored, with Brian (house 1), compared to Sofie & Arthur (house 9), who discussed his high thermal comfort expectations:

> In these rooms, the daily living rooms we prefer to have it around 22/23 degrees, ah, in the winter (emphasis), and in the bathroom we prefer around 24/25 (Brian, 61-year-old, house 1).

During the activity-level disaggregation process, electrical heating load was observed only at some households, as others were able to cover all of their heating needs through

109

the ground source heat-pump system. The methodology was validated through the soft labels on houses 1 and 9, and rolled out across all other houses in the study. As time-of-use surveys were not available for houses 3, 5, 6 & 7, validation of the disaggregated loads was performed through manual inspection of the electricity load profiles by an energy expert. Figure 5.7 illustrates a sample of the results of the disaggregation of activities as well as a sample of the aggregated active power signal for house 1.
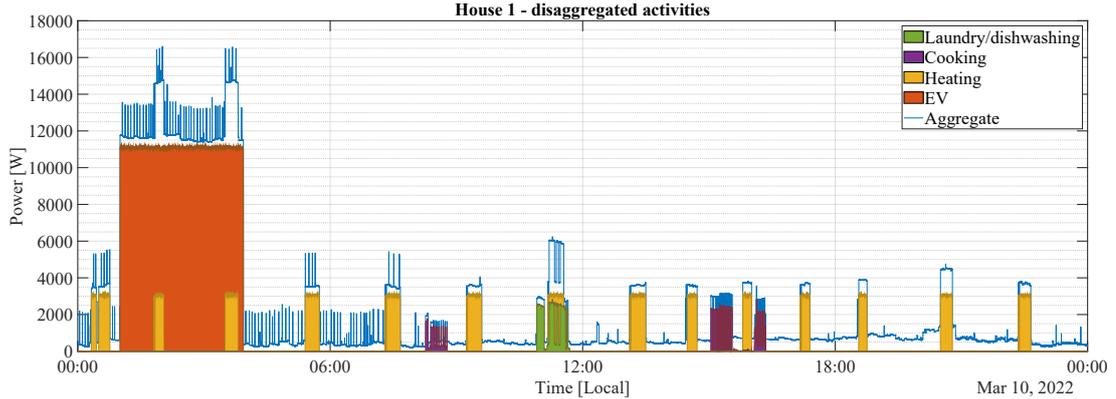
Figure 5.7: House 1: aggregated power and activities breakdown. ©2024 IEEE

5.2.3 Estimation of renewable production per house

As solar production was not monitored, PV production was estimated based on the installed solar capacity, roof tilt and azimuth angle (see Table 5.1) and the local weather and solar data (see Subsection 5.2.1.4). PV panels are either installed in a fixed tilted rooftop (30°) or on flat rooftops with dual-tilt system (15°). The azimuth angle is measured from South with positive values towards the West and negative values towards the East. Data were collected on an hourly basis as market clearance occurs once every hour. Based on the installed capacity of PVs, as well as on the orientation and tilt of the solar panels installed on the rooftop — assuming a fixed azimuth and tilt angle — an hourly estimate of the energy produced through the PVs was calculated based on the widely used and cited global solar energy estimator (GSEE) simulation model

of [130]. The direct plane irradiance is then given by (as in [130], Eq. (2)):

$$I_{dir,p} = \frac{I_{dir,h} \times cos(\alpha)}{cos\left(\pi/2 - \alpha_s\right)}, \tag{5.10}$$

where α is the plane incident angle given by (as in [130], Eq. (1)):

$$\alpha = arcocos\left[sin(h) \times cos(\alpha_t) + cos(h) \times sin(\alpha_t) + cos(\alpha_p - \alpha_s)\right], \tag{5.11}$$

and the diffuse plane irradiance by (as in [130], Eq. (3)):

$$I_{dif,p} = I_{dif,h} \times \frac{1 + cos(\alpha_t)}{2} + a \times (I_{dir,h} + I_{dif,h}) \times \frac{1 - cos(\alpha_t)}{2}, \tag{5.12}$$

with $I_{dir,h}$ and $I_{dif,h}$ being the global direct and diffuse irradiance respectively, a being the albedo, h being the angular elevation of the center of the solar disk above the horizontal plane, α_p being the solar panel azimuth, α_t being the solar panel tilt and lastly α_s being the solar azimuth, i.e., the angle between the projection of Sun's centre onto the horizontal plane and due south direction. Lastly, based on the work in [130] and the PV performance model presented in [131], panel efficiency was calculated based on temperature-dependent parameters. An average temperature-dependent efficiency of 93% [131] based on the latitude and temperature of the neighbourhood (in general, the annual relative efficiency decreases as we move towards the equator due to the increase in temperature), a panel efficiency of 20% and an inverter efficiency of 90% were used.

5.2.4   Exploiting energy price information

In Norway, the energy market is cleared on an hourly basis. The hourly balance of import minus export is calculated, and then the customer is either debited or credited the equivalent amount. The hourly energy price per kWh — import and export — is communicated to the customer one day in advance. The import cost of energy per kWh — denoted as $b(t)$ [NOK/kWh] — is the sum of the price per kWh — denoted as $p(t)$ [NOK/kWh], the VAT — denoted as VAT and currently 25% —, and the grid fees — denoted as g [NOK/kWh] — which were approx. 0.4NOK/kWh for the monitored

period. Therefore, the hourly import cost in [kWh] is given by:

$$b(t) = p(t) \times (1 + VAT) + g. \tag{5.13}$$

On the other hand, the export gain per kWh — denoted by $s(t)$ [NOK/kWh] — is the sum of the price per kWh[3] (without the addition of VAT) plus a small compensation for supplying the grid — denoted as c [NOK/kWh] — which is approx. 0.1NOK/kWh. Therefore, the hourly export gain per kWh is given by:

$$s(t) = p(t) + c. \tag{5.14}$$

The energy balance, i.e., the energy exported subtracted from the energy imported per time slot and denoted as $E_{bal}(t)$ can be expressed as:

$$E_{bal}(t) = E_c(t) - E_p(t), \tag{5.15}$$

with $E_c(t)$ and $E_p(t)$ being the energy consumed from the appliances and the energy produced (from the solar panels) at time t, respectively. $E_p(t)$ was estimated through the solar insolation data and the installed capacity as described in Section 5.2.3 through the methodology provided in [130]. The appliances' energy consumption, $E_c(t)$, can be expressed as:

$$E_c(t) = \sum_{i=1}^{n} \alpha_i(t) \times E_i \tag{5.16}$$

where α_i is the state of the $i-th$ appliance out of a total of n appliances and E_i is the energy vector of the $i-th$ appliance. Therefore, the energy cost per time-slot can be expressed as:

$$C(t) = E_{bal}(t) \times w(t) \tag{5.17}$$

---

[3]In Norway, during the monitoring period, as already mentioned in Section 5.2.1, the energy produced is sold at the same price as the energy imported from the grid (without including VAT).

where $w(t)$ is set to $b(t)$ or $s(t)$ if energy is imported or exported, respectively. By combining the above equations, the energy cost per time-slot can be written as:

$$C(t) = \begin{cases} \left(\sum_{i=1}^{n} \alpha_i(t) \times E_i - E_p(t)\right) \times (p(t) \times (1 + VAT) + g) & E_{bal}(t) \geq 0 \\ \left(\sum_{i=1}^{n} \alpha_i(t) \times E_i - E_p(t)\right) \times (p(t) + c) & E_{bal}(t) < 0 \end{cases} \tag{5.18}$$

and the total energy bill as:

$$B = \sum_{t=1}^{T} C(t) \tag{5.19}$$

where T is the total monitoring period.

The financial gain obtained through load shifting is capped by the maximum amount of flexibility that each user is willing to accept on a per-activity basis. Therefore, the maximum financial gain will be obtained when B is minimum, under the constraints that a continuous event cannot be split, i.e., an appliance activation cannot be intermitted and split into sub-activations, that certain appliance activation are bounded by the activation of another appliance, i.e., certain appliances' loads are dependent on previous appliances loads — e.g., the TD and the WM — and that activation constraints are imposed by the requirements of the end-users.

As inferred from the empirical study and validated through the smart meter data, several users selected to export their solar energy (instead of self-consuming) during the solar production hours, as the energy price was higher and then import energy from the grid during cheaper energy hours. The partial average arbitrage gain through this strategy can be obtained by combining Equations 5.13 & 5.14 and can be expressed for each household as:

$$G_{arb} = E_{shifted} \times \left(\bar{p}_{high} \times (1 + VAT) - \bar{p}_{low} + g - c\right), \tag{5.20}$$

where $E_{shifted}$ is the amount of energy that is not self-consumed but exported to the grid during higher energy price periods and later re-imported during lower energy price periods, $\bar{p}_{high}$ is the average electricity price during the exporting period and $\bar{p}_{low}$ is the average electricity price during the lower tariff hours.

5.3   Mixed-methods evaluation approach and key findings

Energy plus-home neighbourhoods are expected to exhibit an energy net-positive balance, i.e., the total energy produced should exceed the total energy consumed. Following the qualitative methodology (empirical study) of the households in this study, it was concluded that those who moved into this energy-plus neighbourhood had expectations of close to zero/negative energy bills. However, after about two years of living in their new homes, the residents agreed during a community meeting with the real estate and energy supply companies that their energy bills were much higher than they had anticipated. Therefore, initial enthusiasm from being able to reduce the bills and achieve net-positive energy balance was replaced with anger and disappointment in the new builds. This motivated the study to determine a systematic methodology for evaluating net-positive and net-zero buildings in terms of energy consumption, taking into account occupant behaviour such that they are meaningful to the building occupants and therefore actionable through flexibilities in their domestic routines. The methodology is demonstrated through a case study on a net-positive community in Norway from six participating households, summarised in Table 5.1, all equipped with a smart meter.

5.3.1   Explaining the energy gap between energy consumption and production in net-positive dwellings

The ratio of estimated solar PV energy production (see Section 5.2.3) is first determined to measured energy consumption from smart meter data. A ratio of total production to total consumption greater than 1 indicates a true net-positive, and the smaller than 1 ratio indicates higher consumption with respect to production. This is shown for the case study, monitored over a period of two months, in the third row of Table 5.3. Only house 9 is net-positive, followed closely by houses 5 and 7 with a close to 1 ratio.

This can be visualised in Figure 5.8, which shows the total energy consumption and production of each household. As can be observed in Table 5.3 and Figure 5.8, houses 5, 7 and 9 have a ratio close to 1, with energy consumption almost matching production. However, houses 1, 3 and 6 have over twice as much consumption as production, with

(a) House 1 (b) House 3 (c) House 5

(d) House 6 (e) House 7 (f) House 9
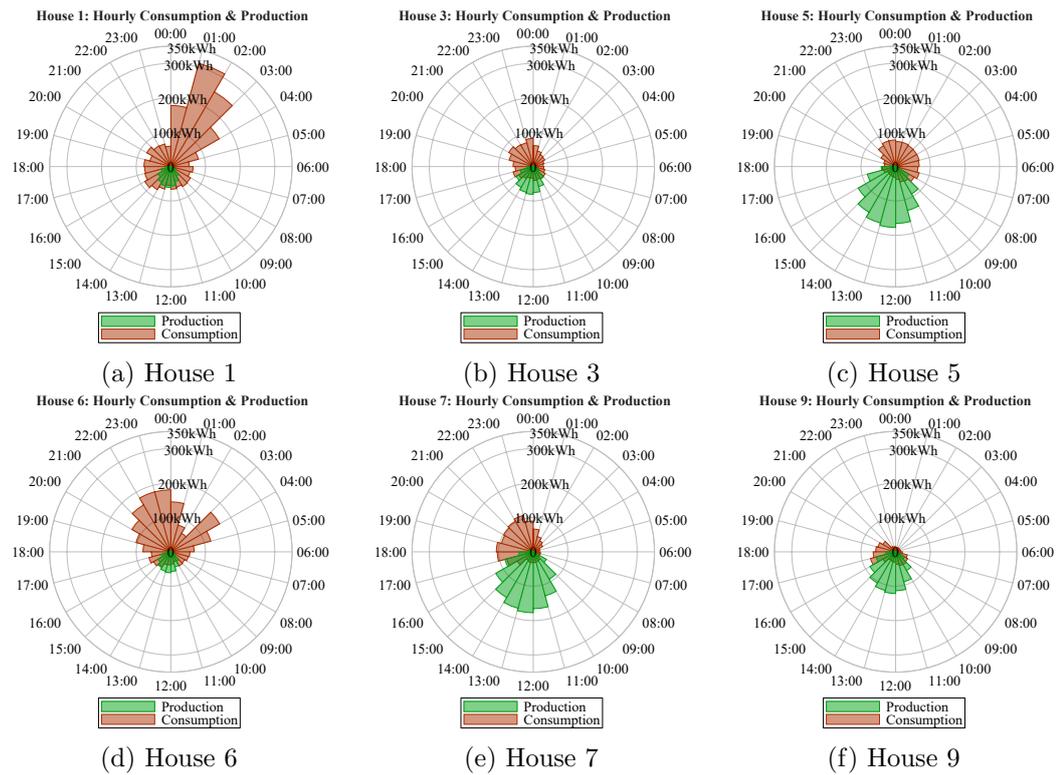
Figure 5.8: Differing levels of hourly discrepancy between energy consumption and production, totalled over the monitoring period, for each of the 6 households.

ratios much less than 0.5, with houses 1 and 6 consuming approximately five times the energy produced. Houses 1, 3 and 6 are completely detached houses, with larger living areas and comparably less production capacity — less space for solar panels (see Table 5.1) on the rooftop due to a roof patio. On the other hand, houses 5 and 7 are semi-detached/terraced houses with a smaller living area and thus lower energy consumption, which is almost compensated by the higher PV production capacity — larger number of solar panels installed on the rooftops (see Table 5.1). Therefore, the actual topology of a building and the limitations that this may introduce in terms of installation capacity of renewables, greatly affect the net balance of future home living spaces and need to be taken into consideration at the design stage. However, in order to do so, it is important to accurately quantify the consumption needs of the inhabitants of these dwellings, which can only be done through the lens of household routines and activities, as discussed next.

5.3.2 Explaining the deviation through the lens of disaggregated activities

As shown in [118], understanding households' consumption through the lens of occupant activities or daily routines offers better actionable insights than aggregate-level smart meter consumption. Following the proposed quantitative methodology of load disaggregation together with qualitative empirical research described in Section 5.2.2 for the same two-month period, the actual consumption of essential energy-intensive routines of heating, cooking, laundry/cleaning, EV charging together with refrigeration consumption are determined and shown in the fourth to eighth rows of Table 5.3. Over 50% of the consumption for all households in the study can be explained. From the empirical study, these "Other" loads can be attributed to smart devices that are running all day, including automation for ventilation/purification of the household, auto blinds and robot vacuums that are charging all day.

Table 5.3: Energy breakdown and estimated electricity cost balance.

	House 1	House 3	House 5	House 6	House 7	House 9
Consumption [kWh]	2276	1181	1289	2045	1300	762
Production [kWh]	414	552	1208	414	1209	826
Ratio	0.18	0.47	0.93	0.20	0.93	1.09
Heating [kWh]	630 (28%)	149 (13%)	0 (0%)	318 (16%)	134 (10%)	0 (0%)
Cooking [kWh]	90 (4%)	77 (6%)	46 (4%)	53 (3%)	64 (5%)	48 (6%)
Laun./clean. [kWh]	74 (3%)	174 (15%)	81 (6%)	131 (6%)	178 (14%)	186 (24%)
EV [kWh]	718 (32%)	258 (22%)	799 (62%)	1061 (52%)	627 (48%)	95 (12%)
Refrigeration [kWh]	138 (6%)	159 (13%)	94 (7%)	144 (7%)	91 (7%)	92 (12%)
Other [kWh]	626 (27%)	364 (31%)	267 (21%)	338 (16%)	206 (16%)	341 (43%)
Bill (B) [NOK]	4429	1649	770	3951	785	185

Heating energy consumption corresponds to the additional energy consumed for space heating when the ground source heat pumps cannot meet the demand. All detached households (houses 1, 3 and 6) and only one of the semi-detached households (house 7) do not meet their heating requirements solely through the ground source heat pumps but need additional energy to achieve their thermal comfort levels, a fact that can be attributed to the higher than expected heating expectations as highlighted by

the empirical study. Cooking activities across all houses are responsible for the same percentage of the total bill (in the range of 3% – 6%), whereas laundry and cleaning activities greatly vary across the participating households. From the empirical study and occupation as per Table 5.1, as expected, households with more occupants (house 7) and households with young children (houses 3 & 9) tend to consume more energy for their laundry/cleaning practices due to the increased demand laundry, tumble drying and dishwashing. An exception to this pattern is house 5, which, although occupied by two adults and 2 children, has a lower laundry/cleaning consumption due to the reduced usage of the TD, concluded from load disaggregation methodology (see Subsection 5.2.2). EV charging greatly varied across the households due to the transportation requirements of the homeowners. As the data correspond to the post-COVID period, from interview data, households 3 & 9 mostly work from home and therefore their transportation needs are lower. On the other hand, households 1, 5, 6 & 7 commute on a daily basis, charging every single day, resulting in their EV charging consumption contributing to almost 60% of their total energy consumption. Lastly, refrigeration also varied across the different households, with detached houses 1, 3 and 6 having higher consumption than semi-detached houses 5, 7 and 9. Indeed, refrigeration of house 9 consumes 1.5 times more than that of house 1. All houses were already furnished with A-rated white goods when sold — semi-detached house 9 had a REF whilst detached house 1 had two refrigerating appliances.

Finally, all essential cooking, laundry and refrigeration-related loads for all houses are covered by solar PV production. As discussed previously, additional heating was not expected due to the communal ground source heating provision, which explains the deviation from net-positive. Although EV charging provision in terms of infrastructure was planned, expected charging patterns and consumption are much lower than actual, especially for houses 1, 3 and 6, whose PV capacity can not meet EV charging together with essential cooking, laundry and refrigeration. This has serious implications for the electrification of transportation as residential charging is growing, and planning net-positive dwellings must take this into account with better informed models.

5.3.3   Explaining deviation between actual and expected energy bills

The estimated electricity balance for each household shown in Table 5.3 was calculated based on Equation 5.19, taking into account the hourly energy consumption and production and the hourly pricing vector. A key observation from the last row in Table 5.3 is that, despite being close to or net-positive, houses 5, 7 and 9 do not have a zero bill, although production should be meeting consumption costs. This deviation is explained next through the energy pricing strategy in Norway, with similar approaches being followed by the majority of countries participating in the Nord Pool [132], where energy prices vary hourly and consumers/prosumers are directly exposed to the price variability for both energy import and export, with billing tied to the day-ahead market price.

As observed in Figure 5.9a, although the electricity price was relatively stable before the end of February 2022, from that point on the price exhibits high variability due to the turmoil in the energy market as a result of the embargo of Russian fuels in several parts of the world following Russian invasion of parts of Ukraine [100]. As can also be seen through the trend line in Figure 5.9a, the price of electricity appears to be increasing throughout March 2022, with a small decline during April due to better weather conditions, decreased energy demand, and stabilisation of the energy market. In Figure 5.9b, the high variance of hourly electricity prices can be observed (on a logarithmic scale), especially during the peak morning hours. Outliers during the period from 07:00–09:00 reached 7 NOK/kWh an almost 4-fold increase from the average. Energy end-users were directly impacted by the hourly variance of the energy price vector, with the households under study commenting on their unexpectedly high energy bills.

> [. . .] but we have an extremely expensive energy in Norway this year. ...
> we are used to pay under 50 øre [∼0.047 euro] for a kWh, and this year we
> have paid 4-5 krone [∼0.45 euro] for a kWh, so it is extremely. So, many
> people in Norway are broke, and the government is going to take some of
> the bill for us. (Brian, 61-years old, house 1)
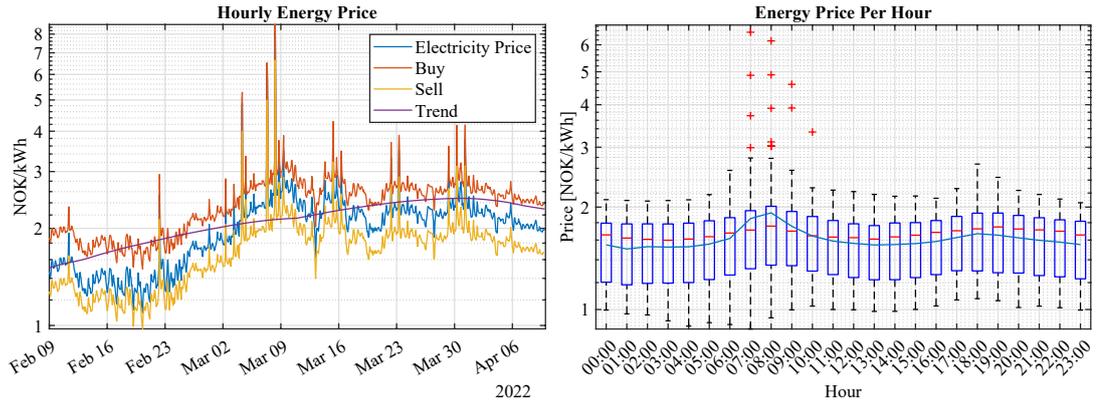
(a) Energy price hourly vector.

(b) Hourly energy price variance over the monitoring period.

Figure 5.9: High energy price fluctuation during the monitored period with evident spikes after the start of the energy crisis.

Combining Figure 5.8 and 5.9b, it can be observed that energy production mostly occurs during 10:00 – 15:00 when the electricity prices exhibit a local minimum, whereas the energy consumption occurs mostly during the early morning hours and the late afternoon/early evening hours when the average hourly electricity price exhibit two local maxima. This partially explains the deviation from zero bills for houses 5, 7 and 9, which, although close to or net-positive, experience a significant bill. The bills can partly be compensated by arbitraging — through load shifting and solar energy exports — due to the energy price model: the local minimum during the midday, when the majority of the solar production takes place, has a median export tariff obtained through Equation 5.14 of 1.58 NOK, which is higher than the global minimum during the night hours, with a median import tariff obtained through Equation 5.13 of 1.53 NOK. On the other hand, the import tariff during early morning and early evening hours is 1.92 NOK and 1.67 NOK, respectively. Therefore, by applying Equation 5.20, for all households, a small gain in the range of approx. 4%–8% is achieved. All houses, except house 1, partially consume what they are producing, exporting the majority to the grid, as observed in Figure 5.8. House 1, although importing the majority of its energy during the night hours when the tariffs are cheaper, due to exceptionally high imports (as observed in brown in Figure 5.8a) relative to exports, incurs the largest bill. Houses

3, 6, 7 & 9 import a significant part of their consumption during the evening when, in general, the electricity prices exhibit a local maximum. House 6, like house 1, has a disproportionally higher consumption than production, with the majority of the energy consumed being concentrated between the early morning hours and the late evening hours when the energy price exhibits maxima. House 3, although partly self-consuming, exports a significant amount of energy to the grid, which is later re-imported between late-afternoon and late-evening, when again the energy prices are higher. Net-positive house 9 has a non-zero bill because it is consuming the majority of the electricity during the two local maxima (morning and early evening) when the energy prices are highest and energy production is low. Similarly, house 7, which is close to net-zero, consumes the majority of energy during the evening when the energy prices are higher. House 5, which although following an arbitrage strategy (see Equation 5.20) by exporting almost all the produced energy and importing back from the grid during the night hours, still import a significant amount of energy consumption during late evening and early morning hours when there is no solar production and the energy prices are higher.

From the empirical study it was concluded that although the energy price was communicated to the end-users in advance, households did not engage with the daily fluctuating energy prices (see Figure 5.9a) but rather assumed approximate periods when the energy price was cheaper or more expensive based on their past experience and therefore the actual incurred costs were higher than expected. This is evident for house 1, especially for EV charging, where the household incorrectly assumed it was cheaper to always charge at 01:00 and is further explored in Subsection 5.3.4 in relation to flexibility along the energy price model to reduce the energy bill.

5.3.4 Load shifting potential demonstrated by a case study

From the previous findings in Subsections 5.3.2 and 5.3.3 and the empirical study, it is clear, householders do not fully benefit from different energy feedback apps and automation systems present in their smart homes due to the non-optimal scheduling of the load consumption, mandated by flexible and non-flexible energy consuming practices, as well as due to the inherent complexity of following and scheduling their daily activities

based on the live fluctuating energy prices. In order to further analyse the energy cost on a per-activity basis using the local energy price, stacked plots of the total hourly cost, broken down on a per-activity basis, were created to inform household demand flexibilities, taking both time of use and the local fluctuating energy price into account. Refrigeration, as is the case for the other always-on loads, is considered non-flexible since it cannot be shifted. Furthermore, based on the interview data, routines that are mandated due to external factors, such as the sequence of certain events, e.g., the usage of the TD after the WM, or transport patterns, e.g., the requirement to have the EV charged by a certain time in the morning, and DW followed by cooking, were constraints considered in the rest of the analysis.

The approach for the load shifting potential of a household has been demonstrated using house 1 as a case study since it had the smallest production to consumption ratio explained by activities such as heating and EV charging not covered by production, and it had the highest energy bill, as discussed in Subsections 5.3.2 and 5.3.3, despite actively trying to shift loads to cheaper tariff times:

> [. . .] we charge the car at night, we don't do the DW in the morning
> or when we are making food for dinner, because it is when we have a high
> price, so we usually turn on the DW when we go to bed because it's when
> the energy is cheaper. (Brian, 61-years old, house 1)

While Subsection 5.3.2 quantified (in Table 5.3) and discussed the activity-level energy consumption, it did not show the temporal dimension of when these activities occur in order to analyse flexibilities. Figure 5.10 visualises the relative energy consumption of activities at different times for house 1. Cooking is a non-flexible activity, as stated by the household. On the other hand, EV charging occurs between 01:00 and 04:00, and laundry and dishwashing occur during morning and evening hours — as per the empirical study, these activities are intentionally carried out to coincide with cheaper energy tariffs and are also flexible.

For comparison purposes, net-positive house 9 is also considered. House 9, while open to doing their bit for the environment, they are not convinced that load shifting will make a difference, as per their interview when asked about load shifting:

(a) Energy Breakdown   (b) Heating   (c) Cooking

(d) Laundry   (e) EV   (f) Refrigeration
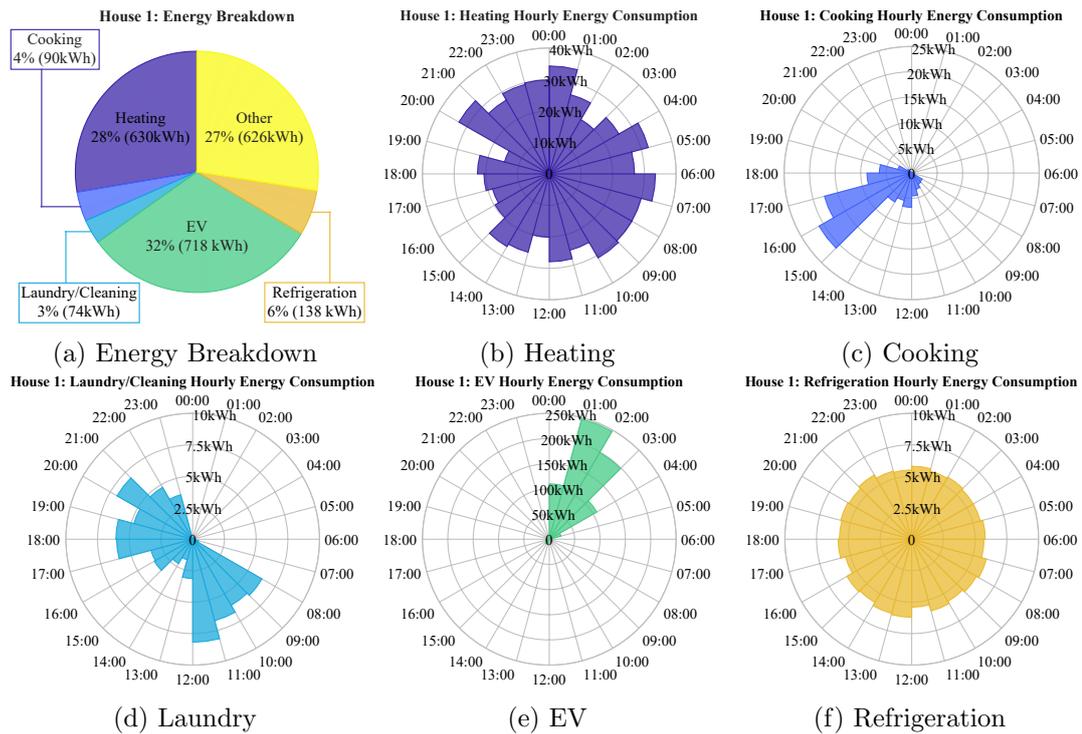
Figure 5.10: House 1: total energy consumption breakdown of heating, cooking, laundry/cleaning, EV charging and refrigeration over the monitoring period.

No, and like I know that a lot of people, or I think some people will maybe wait to do laundry or something, but to be honest, I don't want to do that. [. . .] I don't wanna change what, any daily activities according to energy prices or energy usage because, well, I mean, these houses are great with energy, with the solar panels and everything, but I guess with home and my comforts, I don't wanna change anything because I just wanna be comfortable so, and maybe it's selfish. [. . .] So if I could do some things to save energy and, you know, every, you hear every 10 minutes of how global warming in the environment we need to do our part and to, and things like that. but I don't think, not doing laundry at six in the evening is going to really make a major change with anything. (Sofie, 32-years old, house 9)

As can be observed through Figures 5.11b, 5.11c and 5.11d and from the empirical study, energy-intensive activities occur primarily during evening hours, after work for house 9. As expected, house 9 with an infant, has higher laundry and dishwashing needs,

with over twice the energy consumption compared to house 1, and contributes to 24% of their consumption, as observed in Figure 5.11a and 5.11c. Qualitative analysis of the interview data indicated that house 9 uses their WM more often than DW, tends to do laundry both in the morning and evening, but dishwashing is mostly in the evening after dinner.

(a) Energy Breakdown   (b) Cooking   (c) Laundry

(d) EV   (e) Refrigeration
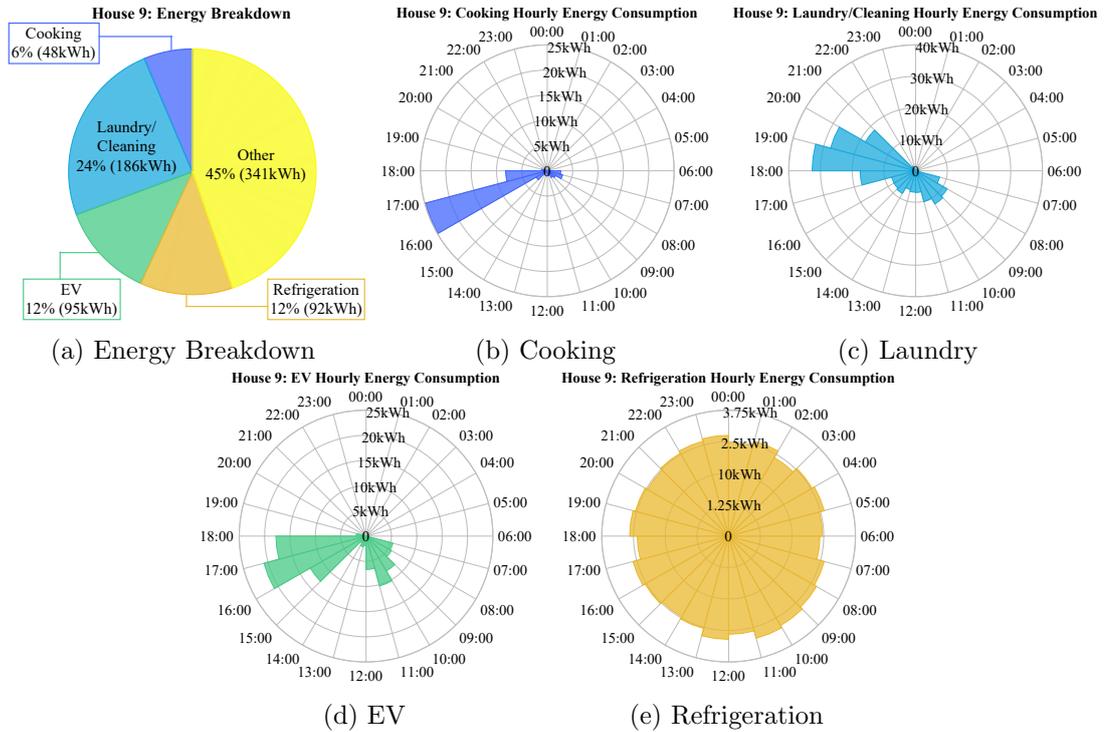
Figure 5.11: House 9: total energy consumption breakdown of cooking, laundry/cleaning, EV charging and refrigeration over the monitoring period.

Figures 5.12a & 5.12b present the total hourly costs on a per-activity basis for houses 1 & 9, respectively. This is in agreement with the previous observation that EV charging is the main contributor to energy bills, followed by heating. Similarly, in house 9, the main contributor is laundry activity and EV charging to a lesser extent at relatively expensive import tariff periods.

Potential for load shifting was estimated per activity, the results of which are presented in terms of total cost reduction and savings per activity given a certain level of maximum accepted flexibility under the constraints imposed either by end-users' practices or intangible loads. For each activity, the total duration, the disaggregated

(a) House 1: hourly cost

(b) House 9: hourly cost

(c) House 1: cost reduction

(d) House 9: cost reduction

(e) House 1: savings

(f) House 9: savings
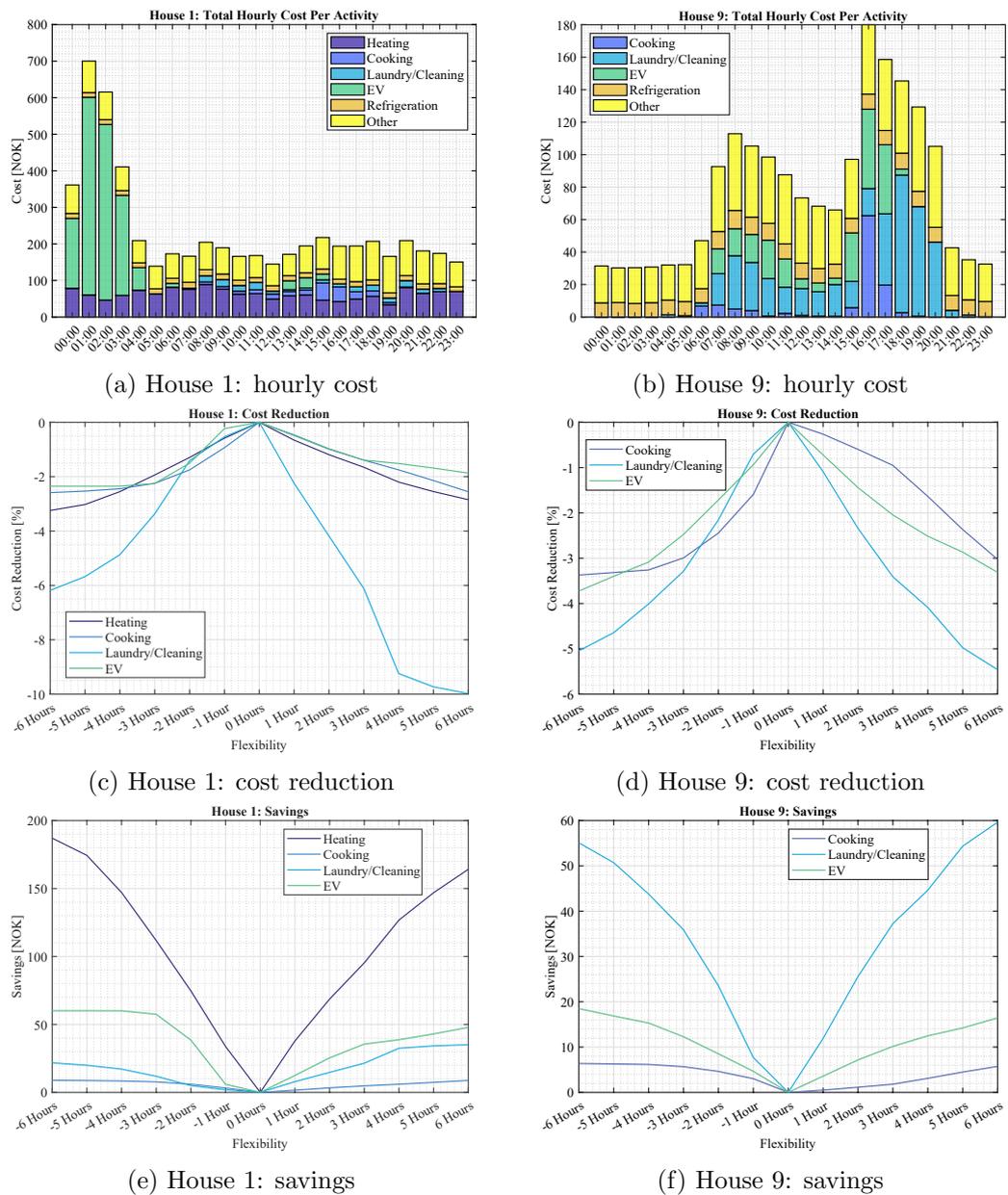
Figure 5.12: Total actual hourly cost and potential cost reduction and savings over the monitoring period, per activity, given different levels of demand flexibility.

energy profile, and the dynamically varying tariff were used to estimate the flexibility potential as discussed in Section 5.2.4. A graph that correlates the accepted flexibility by end-users and the resulting reduction of the cost on a per-activity basis is presented. In addition to the cost reduction graph, a separate graph is produced that enables

end-users to understand their per-activity savings when accepting a certain level of flexibility.

Flexibility analysis was performed in house 1 for heating, though not specified as flexible by the occupant, as the inherent inertia of the building materials can compensate the temperature drop that would occur by moving a heating load. Figures 5.12c & 5.12e depict the maximum possible cost reduction, and therefore savings, per activity for house 1. As can be seen, laundry has the highest potential for cost reduction (up to 10%) in terms of percentage compared to the rest of the activities. However, since laundry routines are not responsible for a high share of the total energy bill (see Figure 5.12b), the total savings of laundry are marginal, that is, up to 30 NOK. Cooking activities also demonstrate a very low capability for cost reduction and savings, mainly due to their low participation in the total bill and their non-flexible nature. On the other hand, EV and heating, which are the most consuming loads (see Figure 5.10a), have a high load-shifting potential. Although Brian (house 1) can monitor the energy price through the energy price app and subsequently schedule his vehicle charger, he does not use it as he does not find that convenient and because he believes that he has already understood — more or less — the price fluctuation. According to Figure 5.12c, although the maximum cost reduction achievable by following the energy prices is approx. 2%, due to the fact that the EV is responsible for a considerable amount of the bill, this reduction can be translated into savings of more than 60 NOK. Lastly, taking into account the inertia of the building materials, heating can be shifted out of the main peak hours — i.e., 07:00 – 09:00 and 17:00 – 19:00 and therefore achieve the maximum possible savings — up to 150 NOK — without sacrificing comfort levels.

Although house 9 did not state that their laundry practices are flexible, flexibility analysis was performed in order to investigate the potential savings. Results for house 9, differ from those of house 1, mainly due to the lower amount of energy used, as well as due to the fact that Sofie and Arthur (house 9) are not actively monitoring and shifting their activities based on energy prices. Unlike house 1, where EV shifting has a very low percentage of improvement, in house 9, EV charging scheduling can lead to a reduction of up to 3.9% of the total costs of the EV. Furthermore, cooking

activities can also greatly benefit from load shifting, even for low levels of accepted flexibility with a maximum possible reduction of up to 3.4%. Lastly, laundry practices, which represent a considerable amount of the total energy used in house 9 as shown in Figure 5.11, can greatly benefit from load shifting with cost reduction of up to 5.5%, i.e., approx 60 NOK. Although house 9 does not expect to make any difference by shifting their activities throughout the day, the flexibility analysis combining Table 5.3 and Figure 5.12f, demonstrated that a reduction of more than 50% of the total bill (and reduced pressure on the grid, and overall more eco-friendly) can be achieved by shifting the daily activities and therefore almost achieving a net-zero utility bill balance.

5.4   Discussion & conclusions

The proposed approach to evaluating the net-positive eenrgy lived-in housing stock is especially timely given the construction of several, designed, net-zero and even net-positive developments throughout the world to reduce the carbon footprint. The built environment is being developed to comply with regulations and not necessarily for actual performance. Jointly considering qualitative data and methods in relation to end-users' routines, as well as dynamic energy pricing and measured consumption and renewable production during design and modelling of the housing stock to inform policy and regulation should be prioritised, as assumptions made during the construction of a building do not always represent the reality. As a consequence, designed "plus" homes, during their usage, fail to achieve their goal. This was demonstrated in this study through the evaluation of a smart neighbourhood in Norway, where, although all houses were designed based on current net-positive standards, they actually failed to achieve that goal. Furthermore, as highlighted through the actual data gathered, in dwellings where the end-user has little understanding of energy production from on-site renewables and dynamic pricing models, end-users who are actively flexible with their energy consumption or expect zero bills are disappointed. A mixed-methods approach-based evaluation of the housing stock that helps pinpoint where assumptions do not meet reality, taking into account household routines and dynamic energy pricing, is proposed. These insights can lead to additional PV panel installation as well as the

load-shifting potential of households to achieve net-zero.

The proposed mixed-methods approach bridges the gap between social science qualitative analyses — which can offer great detail and high explanation but with limited scope in scaling and high cost — with engineering quantitative analyses — which can scale up but can lack explanatory power through abstraction and generalisation of traditional energy data analysis design methods. Although the proposed mixed-methods methodology is shown to more accurately evaluate and explain energy demand of net-positive dwellings by incorporating the diversity of occupants and their practices, the reliance on qualitative data — that could lack accuracy — and the subsequent errors in load disaggregation that embed this qualitative data could affect the accuracy of the overall methodology. Therefore, the main key limitations of the study would lie in the scalability due to the reliance on qualitative data and the accuracy of the methodology due to occupants not providing, intentionally (due to privacy concerns) or not (they can genuinely forget some aspects of their energy-intensive activities), accurate responses in home surveys and interviews. The latter is mitigated in the study through the triangulation and the cross-validation of the qualitative and quantitative data as proposed in Section 5.2.

The proposed methodology can be directly applied to other net-positive dwellings where required quantitative and qualitative data can be collected (smart meter data, PV size and orientation, tariff information, participation in the interviews). Absence of some data used in this study could limit the accuracy and type of findings. Different mixed-methods approaches can be compared by using a different method for one of more of the building blocks of the overall proposed methodology, shown in Figure 5.1. For example, these could be different NILM approaches for the estimation of the load consumption of individual activities, different PV and solar models for the calculation of the energy production and different models for estimating energy cost based on user feedback or appliance sub-metering.

Chapter 6

A Non-intrusive load monitoring-enabled framework for load scheduling in the dairy industry

6.1 Introduction

To meet the demands of the ever-increasing world population [133] and the global increase in calorie intake, it is expected that in a decade the global consumption of food commodities will increase by approximately 15% [134]. With the requirement of increasing agricultural production and reducing carbon emissions at the same time,

there is growing pressure on the agricultural sector to increase its resource and energy efficiency. According to the IEA, in 2020, the agricultural sector emitted $0.4GtCO_2eq$, with expected increase in emissions by $0.1GtCO_2eq$, by 2050 under the stated policies' scenario, which is much higher than the sustainable development scenario — with an expected decrease of CO_2 emissions by $0.3GtCO_2eq$ — and the net-zero emissions by 2050 scenario — with an expected decrease of the CO_2 emissions by $0.5GtCO_2eq$ [135].

Current literature on reducing GHGs of the agri-sector is mostly limited to models enabling the estimation of agricultural emissions [136], which focus mainly on methane and nitrogen emissions from cows, soil, and fertilisers. This has paid off since emissions of agricultural GHGs, specifically methane and nitrous oxides, have been decreasing over the past 30 years in the UK, according to the latest agriclimate report. However, CO_2 emissions from the agricultural sector in the UK amounted to $5.5MtCO_2eq.$, equivalent to almost 2% of total UK carbon emissions [137]. This corresponds to an increase of 22% since 1990, mainly due to the increasing use of energy-intensive agricultural technologies.

According to the latest farm practices survey [138], in England, despite the fact that almost two-thirds of farmers consider emissions when making decisions related to farming processes, the main obstacle that farmers are facing is that there is no clear understanding and information on how to reduce and quantify emissions. Similarly, in Scotland, where approximately 80% of the country's land mass is used for agricultural production and more than 1 in 10 Scottish jobs directly or indirectly dependent on agriculture [139], farmers are struggling with increasing electricity costs since 80% of total energy use on a dairy farm is accounted for by milk cooling, water heating and vacuum pumping [140], activities necessary for the delivery of the final product. Indeed, potential savings of up to 12% in electricity bills could be achieved through the monitoring and analysis of energy usage [140].

To provide comprehensive energy feedback, and support investment, retrofit decisions and benefit from demand flexibility services, it is necessary to understand energy demand on the farm, including consumption of individual high loads. However, since physical monitoring of all equipment on farms is resource-intensive and impractical, recent

attempts to leverage NILM in the agricultural and dairy sector have emerged [26]. NILM was implemented via a multilayer seq2seq DNN, namely a one-directional convolution layer-bidirectional gated recurrent unit (GRU) recurrent neural network (RNN) model, for state detection of milk cooling and vacuum pump equipment of dairy farms in [141]. The model was compared with an LSTM model, with the former performing better in the classification problem, compared to the LSTM network. NILM regression was proposed in [129] to quantify the energy consumption of milking robots, pumps, compressors, and cleaning equipment across 3 farms via the WaveNet seq2point network, where it was concluded that excellent performance can be achieved when models are trained and tested on the same farm, but transferability to other farms failed due to the complexity of metering practices and variation of load profiles/signatures of agricultural equipment.

Recent review papers focusing on energy consumption on dairy farms [9] and livestock systems in the EU [142], highlighted the need to estimate energy consumption to improve energy efficiency, but current studies do not always manage to include on-site consumption data of equipment and facilities due to the unavailability of data, but rather focus on national or regional level averages. The fragmented analyses and the considerable data gaps in existing literature have also been highlighted in [142], with the need for developing a standardised methodology to measure the energy consumption in these systems as an enabler for the reduction of fossil fuel usage and mitigator for the consequences of climate change. Alternative dairy practices and mitigation strategies aimed at guiding the dairy industry toward achieving net-zero carbon emissions are suggested in [143], emphasising the role of supply chain energy modelling as a strategy to reduce energy consumption in dairy settings. The factors that drive energy efficiency in the Dutch dairy sector were studied in [144], where, based on 25,000 dairy-farm records, it was concluded that solar energy is the best method to increase energy efficiency by reducing the dependency on non-RES. The effects of increased production on energy efficiency were also discussed, with the study highlighting the reduced energy efficiency of automated milking systems such as voluntary milking robots compared to traditional milking techniques. The trajectory of Swedish dairy farms to meet net-zero was approached in [145] by assessing the eco-efficiency of the farms through the inclusion

of a sustainability target in the eco-efficiency scores, demonstrating that the vision of a fossil-free economy could be achieved by 2045. Lastly, LCA studies have been performed in both horticultural and livestock farming environments, with LCA studies of vertical farming in the UK highlighting lighting, HVAC as the major energy consumers [146], and LCA methodology aimed in the dairy sector of the United States [147] concluding that no "one solution fits all" approach is possible, with bespoke mitigation strategies adapted to each individual setting being the way forward. Although the energy consumption and LCA of agriculture and more specifically of dairy systems have been studied in the literature, carbon estimation is still based on countrywide long-term averages rather than using actual generation data.

The current literature, reviewed above, does not take into account stakeholders but rather focuses on a system/function-centred approach, which has limited efficacy with users [148]. According to empirical studies and recent research, co-creation and co-design can accelerate the design process and create more innovative concepts and ideas [149,150], while at the same time, increasing user empowerment and democratisation [151]. Co-design between system designers and farmers in the dairy sector in New Zealand was explored in [152], concluding that although the participation of farmers in the design process is of paramount importance, this should be performed in a way that avoids highly structured and process-heavy methodologies, while at the same time allocating sufficient time for initial engagement with farmers to understand their needs and establish realistic expectations of the process.

Further to that, in the agricultural sector, publicly available energy use datasets are available only for a poultry feed facility in Brazil [153] — comprising active, reactive, apparent, voltage and current measured for 111 days at 1-sec granularity — and 2 sites for dairy farms. An hourly synthetic aggregate electrical consumption and estimated PV production data for a year was released in [154] without any measurements for a small-medium dairy farm in Finland to explore integration of renewables in a microgrid. More recently, aggregate energy consumption measurements of three monitored dairy farms in Germany with nine submetered points related to milking robots sampled at 10 seconds for a year were released in [11]. Though this dataset gave insights into the load

profiles of voluntary milking systems, there are more complex loads in the dairy sector that remain undocumented.

In summary, although different approaches have been followed to assess the energy efficiency of dairy farms, these either neglect the emissions related to day-to-day energy intensive activities, or focus only on high-level energy efficiency, without the usage of granular energy datasets (due to the lack of open access datasets) and are not co-created with end-users, hence potentially not being adopted in practice. In this chapter, a co-created NILM-enabled data-driven system approach, through a complex mixed-methods design, aimed at reducing the carbon footprint and the utility bills of dairy farms, is presented, answering the following research questions:

- is it possible to accurately disaggregate, non-intrusively, complex and non-standardised energy-intensive processes across many dairy farms with agritech?

- how can we accurately estimate the carbon footprint of individual processes in dairy farms?

- is load scheduling, minimising carbon footprint and energy bills, even possible on agritech-enabled farms, given tight constraints on timing of energy-intensive processes?

- how do different levels of intrusiveness and granularity approaches in energy monitoring affect load scheduling optimisation and carbon footprint of the energy-intensive dairy activities?

In order to answer the above research questions, the following solutions are proposed and described in this chapter:

- a co-created NILM-enabled methodology to quantify consumption of energy-intensive activities in the dairy sector;

- a scalable, very-low frequency three-phase NILM approach based on DNNs for the energy-intensive dairy sector, demonstrated in three different scale farms in the UK for a period of a year;

- an estimation of load flexibility potential, followed by a co-created load scheduling approach based on the actual identified activity flexibilities that reduces utility costs and carbon footprint;

- a sensitivity analysis of the load scheduler under different disaggregation accuracy and data granularity scenarios; and

- analysis and evaluation of the proposed methodology in 3 farms ranging from small- to large-scale dairy settings.

Further to the aforementioned contributions, this chapter also introduces and describes the curation and release of the comprehensive electrical loads measurement FIELD dataset for a diverse range of typical energy-intensive activities, including detailed labelling and load characteristics information that improves the understanding of the diverse dairy farming activities. The dataset contains granular 1-second active power, aggregated and sub-metered, three-phase readings from 30 dairy farms for a period spanning over 1 year (from the 1st of February 2020 to the 5th of March 2021) that enables seasonal variation analyses in addition to activity recognition, energy consumption analysis of individual energy intensive activities, automated load shifting and DR, renewables and energy storage integration. At the time of publication, this dataset is the largest and only electricity dataset providing aggregate load consumption readings from 30 monitored farms together with a range of submetered readings for a diverse range of energy intensive dairy equipment, spanning voluntary milking systems (milking robots), traditional milking parlours and their submetered components, diverse feeding equipment, cleaning, ventilating, lighting, heat exchanging, and other farming technologies. FIELD dataset has been used as an enabler, by training the DL models, for the proposed methodology.

The rest of the chapter is organised as follows: in Section 6.2, the methodology of the curation and release of the FIELD dataset, as well as the co-creation approach, is presented, followed by the results in Section 6.3. Lastly, a brief discussion and conclusions are presented in Section 6.4.

6.2   Methods

The system diagram for the proposed co-created methodology for the modular optimised load scheduling, based on a complex mixed-methods approach akin to the prior work [92],

is shown in Figure 6.1. The proposed system consists of four main building blocks, where each sub-system can be replaced by different methodologies. The focus is not on optimising each individual sub-system, but rather on a holistic approach that integrates all the sub-systems. Each subsystem is analysed, one by one, in the following subsections.
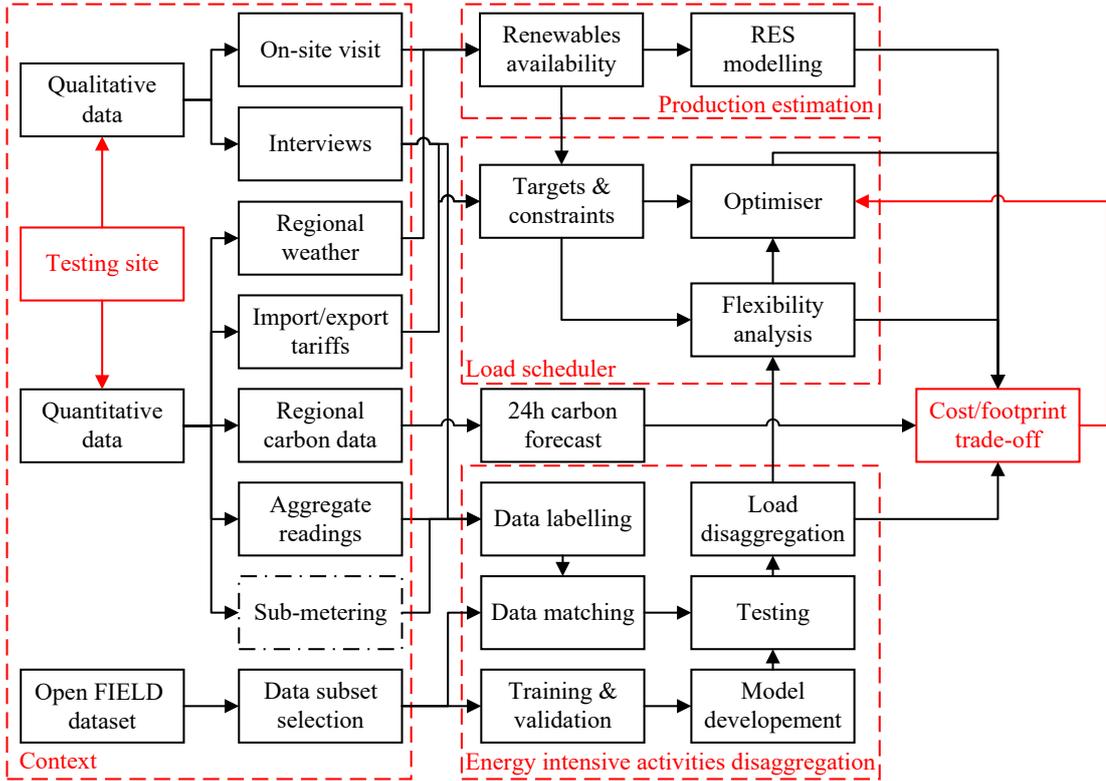
Figure 6.1: System summary.

6.2.1   Context setting

In this subsection, data and constraints that will influence the following data-driven analysis sub-systems to make accurate predictions are defined. The overall system problem is to minimise a farm's carbon footprint and energy export losses by improving energy management through maximisation of the local PV production, storage of excess energy, and load shifting of the energy-intensive farming activities, given time of use and physical constraints of particular activities.

6.2.1.1 FIELD dataset collection methods

Dairy farms, included in the FIELD dataset, were sampled across Germany with the smart metering and sub-metering infrastructure installed through the utility provider. A total of 31 dairy farms were monitored, containing a wide variety of novel dairy technologies, including voluntary milking robots, automated scrapers and climate-control barn ventilation. Out of these 31 farms, due to connectivity and data transmission issues, aggregate data could not be collected from the farms with the following identification (id) numbers: 8, 13, 14, and 19, while submetered data were collected from all the farms, with the number of monitored dairy equipment ranging from 1 to 4 metering points per dairy farm. Table 6.1 presents a summary of the data available, including the monitored duration and the availability of aggregate and/or submetering readings.

6.2.1.1.1 Aggregate readings

Each of the three phases, with a 380V potential difference between any two phases, is monitored individually per farm. Three active power level readings were collected and transmitted every 1 second, corresponding to each of the three installation phases. In case of transmission failure within a given period, all three-phase readings during that period are lost.

6.2.1.1.2 Submetered equipment

On each dairy farm, up to 4 different points were selected for power monitoring that corresponded to different energy-intensive activities. The monitored dairy equipment can be attributed to the following activities:

- milking: including milking robots, breast, and vacuum pumps;
- feeding: including grist mills, feeding augers and feeding cabinets;
- cleaning: including manure removers, heavy-duty cleaners, pipe flushers, dunging and pressure water cleaners;
- ventilation: including cowshed fans, barn fans and ventilation compartments;
- lighting: including indoor and outdoor lighting and infrared lighting;

Table 6.1: Training dataset collection. Overview of measurements available per farm. The submetering column shows the number of submetered dairy equipment. *No activity was identified in the measurements from farm 15, and therefore, the electrical measurements are not included in the dataset.

ID	From	To	Period	Aggregate	Submetering
1	14/03/2020 09:26:21	05/03/2021 00:00:00	11mo 18d 14h 33m 39s	✓	✓(4)
2	13/02/2020 09:09:55	09/07/2020 12:54:48	4mo 26d 3h 44m 53s	✓	✓(4)
3	01/02/2020 00:00:00	05/03/2021 00:00:00	1y 1mo 4d	✓	✓(3)
4	05/02/2020 09:14:26	05/03/2021 00:00:00	1y 27d 14h 45m 34s	✓	✓(4)
5	01/02/2020 00:00:00	05/03/2021 00:00:00	1y 1mo 4d	✓	✓(4)
6	03/02/2020 13:37:20	05/03/2021 00:00:00	1y 1mo 1d 10h 22m 40s	✓	✓(1)
7	03/02/2020 12:58:33	05/03/2021 00:00:00	1y 1mo 1d 11h 1m 27s	✓	✓(1)
8	03/02/2020 12:58:51	05/03/2021 00:00:00	1y 1mo 1d 11h 1m 9s	✗	✓(1)
9	13/02/2020 09:06:09	03/09/2020 14:55:35	6mo 21d 5h 49m 26s	✓	✓(4)
10	09/03/2020 12:27:26	05/03/2021 00:00:00	11mo 23d 11h 32m 34s	✓	✓(3)
11	26/02/2020 11:26:25	27/08/2020 08:35:16	6mo 21h 8m 51s	✓	✓(4)
12	26/02/2020 11:26:27	03/09/2020 14:55:35	6mo 8d 3h 29m 8s	✓	✓(3)
13	11/03/2020 06:59:40	05/03/2021 00:00:00	11mo 21d 17h 0m 20s	✗	✓(1)
14	11/03/2020 06:07:47	05/03/2021 00:00:00	11mo 21d 17h 52m 13s	✗	✓(1)
15	11/03/2020 06:21:30	02/03/2021 13:41:44	11mo 19d 7h 20m 14s	✗*	✗*
16	11/03/2020 06:17:03	05/03/2021 00:00:00	11mo 21d 17h 42m 57s	✓	✓(2)
17	14/03/2020 09:47:38	05/03/2021 00:00:00	11mo 18d 14h 12m 22s	✓	✓(2)
18	14/03/2020 09:35:21	05/03/2021 00:00:00	11mo 18d 14h 24m 39s	✓	✓(4)
19	14/03/2020 09:36:38	05/03/2021 00:00:00	11mo 18d 14h 23m 22s	✗	✓(1)
20	16/04/2020 10:31:24	05/03/2021 00:00:00	10mo 16d 13h 28m 36s	✓	✓(1)
21	16/04/2020 10:44:02	01/02/2021 09:10:23	9mo 15d 22h 26m 21s	✓	✓(4)
22	13/02/2020 09:09:54	03/09/2020 14:55:36	6mo 21d 5h 45m 42s	✓	✓(1)
23	16/04/2020 10:53:50	05/03/2021 00:00:00	10mo 16d 13h 6m 10s	✓	✓(4)
24	11/03/2020 06:19:54	05/03/2021 00:00:00	11mo 21d 17h 40m 6s	✓	✓(4)
25	10/02/2020 13:36:26	20/10/2020 10:02:27	8mo 9d 20h 26m 1s	✓	✓(1)
26	18/02/2020 12:51:12	05/03/2021 00:00:00	1y 14d 11h 8m 48s	✓	✓(1)
27	10/02/2020 13:35:29	05/03/2021 00:00:00	1y 22d 10h 24m 31s	✓	✓(3)
28	10/02/2020 13:31:46	14/08/2020 01:26:57	6mo 3d 11h 55m 11s	✓	✓(2)
29	18/02/2020 12:35:29	05/03/2021 00:00:00	1y 14d 11h 24m 31s	✓	✓(2)
30	18/02/2020 12:52:53	05/03/2021 00:00:00	1y 14d 11h 7m 7s	✓	✓(1)
31	18/02/2020 12:53:19	05/03/2021 00:00:00	1y 14d 11h 6m 41s	✓	✓(1)
	01/02/2020 00:00:00	**05/03/2021 00:00:00**	**1y 1mo 4d**	**26**	**72**

- heating & cooling: including heating plates, warming cabinets and compressors; and,

- miscellaneous: including straw barns for the feeding of the animals, circulation

pumps, industrial washing/ironing machines required for the cleaning/disinfection of the farmers' clothes, gas cannons to scare pests and equipment available in farms' kitchens such as salamanders and extraction hoods.

Specifically, the submetering points (SPs) and the respective equipment that were monitored on each farm (note that all monitored equipment is three-phase powered unless stated otherwise) are presented in Table 6.2.

Table 6.2: Summary of SPs and the maximum power levels of each submetered equipment per farm in the training set. Equipment that is not connected to all three phases but either one or two is marked with ϕ_i, where $i = 1, 2, 3$ are the connected phases.

ID	SP_1	SP_2	SP_3	SP_4
1	Straw barn (3,933 W)	Grist mill 1 (19 W)	Grist mill 2 (2,507 W)	Manure removal (4,509 W)
2	Unassigned (4,945 W)	Light path 1 (9,268 W)	Light path 2 ($\phi_{1,2}$: 2,681 W)	Gang farrowing (29,253 W)
3	Heavy duty cleaner (14,679 W)	Lighting (5,251 W)	Night/outdoor lights (1,290 W)	-
4	Unassigned (16 W)	Lighting (2,443 W)	Fan cowshed (13,510 W)	Lights stable (ϕ_1: 7,808 W) office (ϕ_2: 3,580 W) lights hall (ϕ_3: 5,700 W)

ID	SP_1	SP_2	SP_3	SP_4
5	Unassigned (15 W)	Climate computer circulation pump (119 W)	Distribution (ϕ_1: 70 W) Alarm (ϕ_2: 24 W) Ventilation (ϕ_3: 1,409 W)	Ventilation (7,823 W)
6	Lights (8,358 W)	-	-	-
7	Climate (9,173 W)	-	-	-
8	Infrared lights (8,293 W)	-	-	-
9	Heating plates 1 (8,088 W)	Heating plates 2 (6,852 W)	Infrared lights 1 ($\phi2,3$: 4,738 W)	Infrared lights 2 ($\phi1$: 2,198 W)
10	32A power socket (5,426 W)	Feeding (17,838 W)	Slurry (22,330 W)	-
11	Fan 1 (3,365 W)	Fan 2 (3,425 W)	Fan 3 (3,433 W)	Fan 4 (3,530 W)
12	Feed auger 1 (2,196 W)	Feed auger 2 (2,222 W)	Metal halide lamps (6,545 W)	-
13	Sauna (18,530 W)	-	-	-
14	Industrial WM (17,242 W)	-	-	-

ID	SP_1	SP_2	SP_3	SP_4
16	Freezer cellar (ϕ_1: 8,377 W) Qarming cabinet (ϕ_2: 2,305 W) Thermal bridge (ϕ_3: 1,264 W)	Lights (ϕ_1: 1,416 W) Sliding car door (ϕ_2: 249 W) Extractor hood (ϕ_3: 7,221 W)	-	-
17	Slider front (571 W)	Slider rear ($\phi_{2,3}$: 2,252 W)	-	-
18	Pressure water cleaning (3,362 W)	Vacuum pump (10,310 W)	Compressor (4,536 W)	Milking robot (3,966 W)
19	Feeding auger (4,846 W)	-	-	-
20	Milking robot (6,541 W)	-	-	-
21	Water treatment (7,658 W)	Pipe cooler pump (5,571 W)	Water pump (4,282 W)	Dunging ($\phi_{1,2}$: 1,378 W)
22	Feeding cabinet (24,967 W)	-	-	-
23	Breast pump (2,146 W)	Vacuum pump (12,037 W)	Pipe flusher (29,391 W)	Fan barn (4,724 W)
24	Fryer 1 (23,370 W)	Fryer 2 (6,267 W)	Ironing machine (12,225 W)	Salamander (202 W)
25	Power socket (19,650 W)	-	-	-

ID	SP_1	SP_2	SP_3	SP_4
26	Lights & feeding (28,492 W)	-	-	-
27	Three-phase power socket (10,346 W)	Ventilation 1 ($\phi_{1,2}$: 478 W)	Ventilation 2 & unassigned (1,283 W)	-
28	Ventilation 1 (4,052 W)	Ventilation 2 ($\phi_{1,2}$: 2,947 W) Gas cannon (ϕ_3: 736 W)	-	-
29	Lighting (21,774 W)	Feeding (12,058 W)	-	-
30	Infrared lighting (16,202 W)	-	-	-
31	Lighting (2,618 W)	-	-	-

A selection of the load profiles of different monitored equipment is included in Figure 6.2, where the per-phase consumption of the monitored equipment is presented. Through the per phase presentation of the load consumption, the level of load balancing of each individual equipment can be observed. The selected equipment include a wide variety of farming activities such as milking (see Figures 6.2h & 6.2i), feeding (see Figures 6.2a & 6.2e), cleaning (see Figure 6.2b), ventilation (see Figure 6.2c), heating & cooling (see Figures 6.2d & 6.2g) and miscellaneous activities (see Figure 6.2f). A 24-hour period is presented for all the monitored equipment, apart from equipment that have sparse activations — i.e., industrial WM, see Figure 6.2f — where a single activation is presented, and equipment with dense repetitive consumption patterns — i.e., heating plates (see Figure 6.2d), compressor (see Figure 6.2g), and milking robot

vacuum pump (see Figure 6.2h) — where for illustration purposes, load profiles ranging from several minutes to a few hours is presented.

(a) Farm 1: Grist mill

(b) Farm 1: manure removal

(c) Farm 5: ventilation

(d) Farm 9: heating plates

(e) Farm 10: feeding

(f) Farm 14: WM

(g) Farm 18: compressor

(h) Farm 18: vacuum pump (robot)

(i) Farm 23: vacuum pump (parlour)
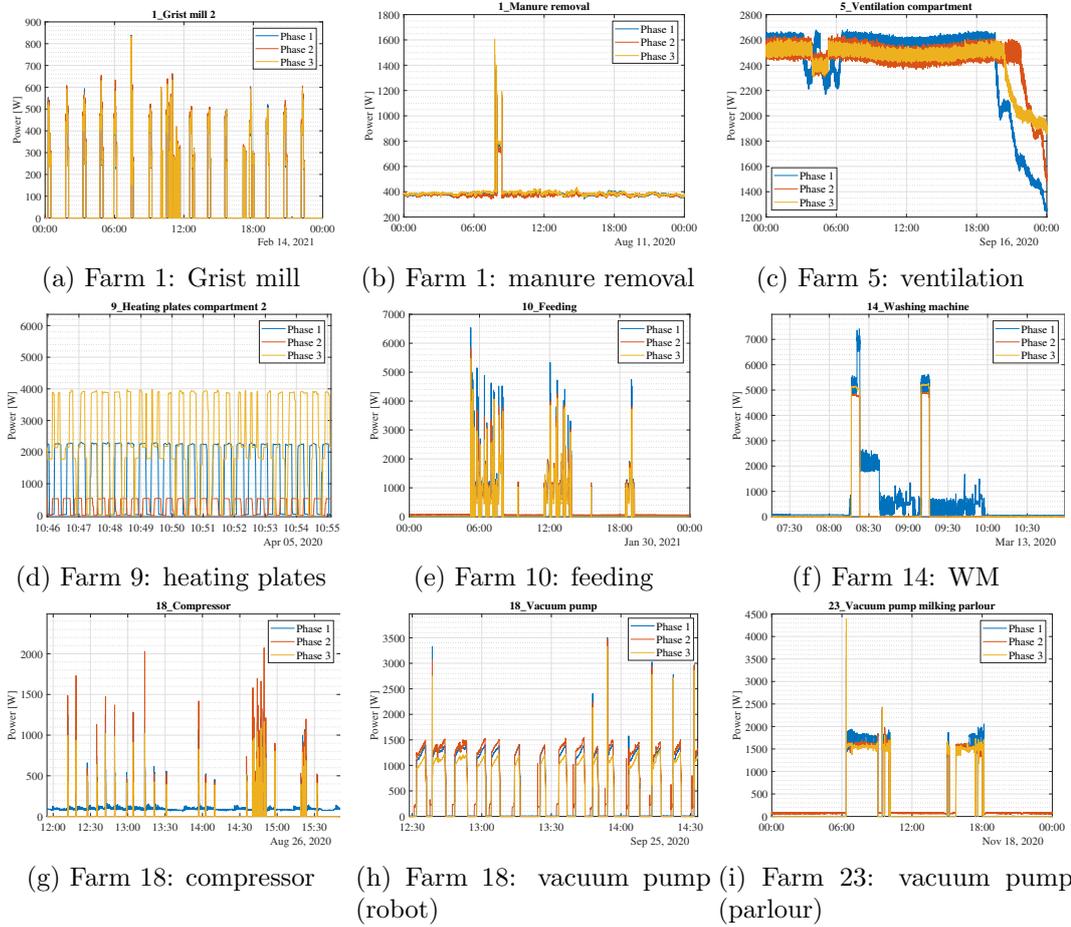
Figure 6.2: A range of selected submetered equipment.

As this dataset aims to enable the comprehensive study of different farming equipment, understanding of the underlying farming processes is essential for the effective application of this dataset to other dairy farms via transfer learning, for example. It is important to note the fundamental difference between the two key milking approaches: milking using a parlour (either a traditional herringbone or a rotary) and milking using voluntary robot milking machines. In this dataset, both types of milking processes have been monitored, with Farms 18, 20 and 21 having voluntary milking robots installed, whereas Farm 23 having a parlour. On the farms with milking robots, the energy consumption of this equipment has been monitored both at an aggregate level and

per sub-process of the milking robot, including water treatment, pipe cooler pump, water pump and dunging (see Table 6.2). The difference between the two milking approaches is evident by comparing Figures 6.2h & 6.2i where in the farm with the voluntary milking robot, the vacuum pump is being used intermittently for a short period every time a cow approaches a milking robot to get milked, whereas in the case of a parlour, where multiple cows are milked simultaneously, there are two distinctive milking rounds (one in the morning and one in the evening) with the milking equipment running constantly during that time. Therefore, although the two equipment may perform a similar task, their consumption pattern vary significantly depending on the underlying farming method. Lastly, due to the variety in sizes across different farms, the same equipment may differ between the two farms (both in terms of power level and activation duration) due to the different number of animals. It is worth noting, though, that the load consumption signal shape is similar, but the magnitude differs. This is also evident in the case of the industrial WM (see Figure 6.2f) where, although the maximum power level exceeds 17,000 W (see Table 6.2), the pattern of the signal resembles those of residential WMs as present in various residential electrical datasets.

6.2.1.1.3   Per-phase signal recovery

Both aggregate and submetered readings in the raw dataset were in the following format: $phase_2 = phase_2 + phase_1$ and $phase_3 = phase_3 + phase_2 + phase_1$, where $phase_i$ is the power level on phase i for $i = 1, 2, 3$. Thus, in order to recover per-phase power readings at each sample point, Algorithm 1 was adopted. Algorithm 1 recovers the original per phase aggregate reading from the aggregate collected per-phase vectors of each farm $\vec{P}_{id,1}, \vec{P}_{id,2}, \vec{P}_{id,3}$ with $id = 1, 2, ..., 31$, and 1, 2, and 3, correspond to Phase 1, 2 and 3, respectively.

   As the submetered points in each farm were organised by distribution boards, with the submetered points on each distribution board numbered in ascending order, each submetered point time-series data contained the sum of the active power readings from the current point and all the submetered points that are connected to the same distribution point with an id smaller than the current one. Therefore, in order to recover

---

Algorithm 1 Aggregate signal recovery

Input: $\vec{P}_{id,1}, \vec{P}_{id,2}, \vec{P}_{id,3}$ ▷ The aggregate collected per-phase vectors, $id = 1, 2, ..., 31$

Output: $\vec{P}_{id,1}, \vec{P}_{id,2}, \vec{P}_{id,3}$ ▷ The aggregate post-processed per-phase vectors,
    $id = 1, 2, ..., 31$

  1: **function** AGGREGATE($\vec{P}_{id,1}, \vec{P}_{id,2}, \vec{P}_{id,3}$)

  2:     $\vec{P}_{id,3} \leftarrow \vec{P}_{id,3} - \vec{P}_{id,2}$

  3:     $\vec{P}_{id,2} \leftarrow \vec{P}_{id,2} - \vec{P}_{id,1}$

  4:     **return** $\vec{P}_{id,1}, \vec{P}_{id,2}, \vec{P}_{id,3}$

  5: **end function**

---

the equipment-level per-phase signal, Algorithm 2 was adopted. Let $\vec{P}_{id,j,i,phase}$ be the collected per-phase power vector of the submetered point i of a distribution board j. The input to Algorithm 2 is the set of collected per-phase vectors of the submetered points (sp) on a distribution board (db), given by:

$$\vec{P}_{id,db,1} = \begin{cases} \vec{P}_{id,db,1,1} \\ \vec{P}_{id,db,2,1} \\ ... \\ \vec{P}_{id,db,sp,1} \end{cases}, \vec{P}_{id,db,2} = \begin{cases} \vec{P}_{id,db,1,2} \\ \vec{P}_{id,db,2,2} \\ ... \\ \vec{P}_{id,db,sp,2} \end{cases}, \text{ and } \vec{P}_{id,db,3} = \begin{cases} \vec{P}_{id,db,1,3} \\ \vec{P}_{id,db,2,3} \\ ... \\ \vec{P}_{id,db,sp,3} \end{cases}$$

where sp is the number of the submetered points on the specific db, and the outputs are the recovered per-phase readings for each submetered point per distribution board.

6.2.1.2   Co-creation sites' dataset collection methods

This Subsection describes the qualitative and quantitative data collection procedure at the three farm sites that participated in the development of the co-created methodology.

6.2.1.2.1   Qualitative data collection

Semi-structured in-depth interviews with a tour of the premises of 3 participating farms were conducted to uncover the different levels of demand flexibilities on a per-activity basis, as well as practical physical constraints. An overview of the farm visits is included in Figure 6.3. The semi-structured interviews initially focused on questions related to the installed electrical equipment of the farms, the electrical connection of the individual equipment, the operational power levels, as well as the import and export tariffs. Further

---

Algorithm 2 Submetered signal recovery

Input: $\vec{P}_{id,db,1}, \vec{P}_{id,db,2}, \vec{P}_{id,db,3}$ ▷ The set of the collected per-phase
   vectors of the submetered points (sp) of a distribution board (db) , $id = 1, 2, ..., 31$,
   $db = 1, 2, ..., n$, where n is the number of dbs.
Output: $\vec{P}_{id,db,1}, \vec{P}_{id,db,2}, \vec{P}_{id,db,3}$ ▷ The set of the post-processed per-phase vectors of
   the submetered points (sp) of a distribution board (db) , $id = 1, 2, ..., 31$
 1: **function** SUBMETERED($\vec{P}_{id,db,1}, \vec{P}_{id,db,2}, \vec{P}_{id,db,3}$)
 2:     **for** $i \leftarrow 1, length(sp)$ **do** ▷ where sp is the submetered point of the distribution
   board (db)
 3:         $\vec{P}_{id,db,i,3} \leftarrow \vec{P}_{id,db,i,3} - \vec{P}_{id,db,i,2}$
 4:         $\vec{P}_{id,db,i,2} \leftarrow \vec{P}_{id,db,i,2} - \vec{P}_{id,db,i,1}$
 5:     **end for**
 6:     **for** $i \leftarrow 2, length(sp)$ **do**
 7:         **for** $j \leftarrow 1, i$ **do**
 8:             $\vec{P}_{id,db,i,1} \leftarrow \vec{P}_{id,db,i,1} - \vec{P}_{id,db,j,1} - \vec{P}_{id,db,j,2} - \vec{P}_{id,db,j,3}$
 9:         **end for**
10:     **end for**
11:     **return** $\vec{P}_{id,db,1}, \vec{P}_{id,db,2}, \vec{P}_{id,db,3}$
12: **end function**

---

to the questions related to the dairy equipment, the interviews focused on the day-to-day
practices and how these are correlated to the energy-intensive equipment. Constraints
that are imposed due to causal processes in the farm (e.g., the requirement to have the
water heated at $90^{o}C$ by the end of the milking round or the legal requirement for the
milk to reach a specific temperature after the milking cycle) were uncovered, as well as
the flexibility level of each of these activities.

The data analysis team and stakeholders (farmers, electricity engineers for energy
monitoring of the farm, mechanical engineers custom-fitting agritech for the farm)
collectively formulated the main objectives of the study to maximise the impact of the
outputs, data requirements and agreed on the methodology steps with regular online
touchpoints to ensure timely feedback for algorithm development and tuning. One of
the key outcomes of the co-creation process was visibility of the heterogeneity of the
equipment across the farms for the same key activities of milking, cooling and heating.
This informed the co-design of the installation of the energy metering equipment to
capture the largest energy-consuming agritech, imports and exports, while at the same
time minimising interference in farm activities.

Figure 6.3: Site visits to the three farms where semi-structured interviews were conducted.

6.2.1.2.2   Quantitative energy demand and production data collection

Aggregate smart meter data and submetering data were collected for the period September 2023 – September 2024, i.e., spanning over 1 year. The period was split into two parts, the first period which corresponds to the development and validation of the methodology (September 2023 – October 2023) at Farm I, and the second period which corresponds to the testing of the methodology (October 2023 – September 2024) across all three farms (Farm I, II, and III).

The installed metering equipment collected voltage, current and power readings approximately every second on an aggregated and sub-metered level, as follows.

- **Farm I:** A small-scale dairy organic farm (grass-based block spring calving herd; circa 250 cows) with traditional herringbone parlour (two milking rounds), and

biomass facilities, with the following monitored equipment:

    I. compressors: ice builder equipment required for milk cooling (7kW, variable);

    II. water heaters: providers of hot water for the disinfection of the parlour after milking (12kW, steady-state);

    III. vacuum pumps: milking pumps for the milk extraction and circulation (11kW, variable);

    IV. biomass facility fans: fans required for the circulation of hot air to dry out organic matter (22kW, variable);

    V. biomass facility boilers: providers of hot water for the water-to-air heat exchanger for the biomass facility (13kW, steady-state); and,

    VI. solar generation: 54kWp with 50kW inverter power[1].

- **Farm II:** a medium-scale dairy farm (grass-based block spring calving herd; circa 400 cows) with a calf house and a rotary parlour that operates only during Spring – Autumn, with the following monitored equipment:

    I. chiller: milk cooling equipment (10kW, variable);

    II. water heaters: providers of hot water for the disinfection of the parlour after milking (30kW, steady-state);

    III. vacuum pumps: milking pumps for the milk extraction and circulation (9kW, variable); and,

    IV. solar generation: 50kWp arrays with 50kW inverters.

- **Farm III:** a large-scale dairy farm (grass-based block all-year-round calving herd; circa 650 cows) with a rotary parlour, with the following monitored equipment:

    I. compressors: milk cooling equipment (20kW, variable);

    II. water heaters: providers of hot water for the disinfection of the parlour after milking (29kW, steady-state);

---

[1]There are three different PV arrays (22kWp, 22kWp, and 10kWp), with the metering equipment monitoring only the two 22kWp arrays as these are placed on the same shed, whereas the third array was not monitored as it is located in a different shed. The installed capacity of the solar panels in the first two arrays is purposely oversized (compared to the inverters) due to the fact that the farm is located in an area with increased likelihood of cloudiness, and therefore the solar panels almost never reach their peak capacity.

III. vacuum pumps: milking pumps for the milk extraction and circulation (9kW, variable);

IV. calf house ventilation: fans required for the circulation of air in the calf house (15kW, variable);

V. water pumps: equipment used to pump water from a watercourse crossing the farm (18kW, variable); and,

VI. solar generation: 100kWp arrays with 100kW inverters.

PV arrays across the 3 farms have no line of sight obstacles and are located on sheds with identical orientations and inclinations (per farm). All monitored equipment is three-phase, with power clamps and voltage meters being installed in only one of the three phases of each equipment, as these loads are symmetrical. The metering equipment remotely provides a sub-second data stream through an MQTT broker. A server was configured to subscribe to the MQTT stream and collect twelve different readings from the farm. To increase the reliability and fault tolerance of the data collection process, a Raspberry Pi located in different premises — to avoid issues arising from internet disconnection and/or power loss — was also subscribed to the same topic of the MQTT broker and collected a backup of the readings from the farm. The two data-streams (server & Raspberry Pi) were synchronised and resampled to 1 sec, resulting in 34 time-series — 9 aggregate phases, 3 PV production, 13 submetered equipment and 9 voltage readings. The resampled time series were then stored in an SQL database. The electricity supply in all three farms is based on a dual tariff system with a day tariff (06:00 – 01:00) of $\sim$ 28p/kWh and a night tariff (01:00 – 06:00) of $\sim$ 17p/kWh. The energy produced by the farm that is not consumed is exported to the grid with no compensation provided by the supplier/grid.

6.2.1.2.3   Qualitative and quantitative data findings that set context

The energy consumption and production data of Farm I for the period from the 15th of September 2023 until the 9th of October 2023 (development period) on a half-hourly basis are presented in Figure 6.4. Milking activities that were identified during the qualitative data collection were matched and mapped to the quantitative data. More
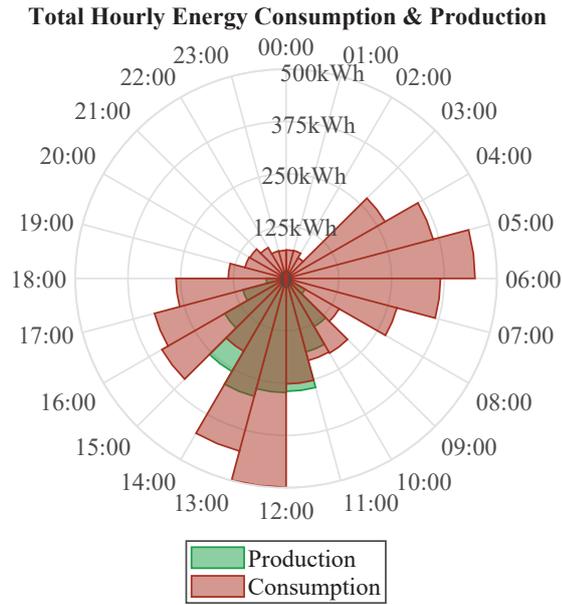
Figure 6.4: Farm I: half-hourly total energy consumption and production of the first monitored (development) period.

specifically, milking activities were identified during the early morning hours between 03:00 and 07:30 with water heaters being deployed in the early morning (03:00 – 05:00) and early afternoon (12:00 – 14:00) whereas milking pumps were deployed in the morning (04:30 – 07:30) and late afternoon (15:00 – 18:00). These farming activities can be identified in Figure 6.4 where energy peaks are observed during the early morning hours (03:00 – 08:00), early afternoon (12:00 – 14:00) and in a lesser extent during the late afternoon (15:00 – 17:00) periods.

During this period, Farm I imported 2868kWh ($\sim$ £803) and 1275kWh ($\sim$ £217) during the day and night tariff period, respectively. The total energy production from the solar panels was 1801kWh with the farm self-consuming 1182kWh (66%) with the rest 619kWh (34%) exported to the grid. As the energy exported to the grid is not compensated, the exploitation of the total solar production of the farm could reduce the total utility bill by 17%; this is estimated based on the amount of energy currently being exported (and hence not compensated) and the equivalent savings if this energy were self-consumed. Further to that, by combining the usage of the total energy production and the dual tariff system, the farm could further reduce the energy bills to a maximum

of 41% by self-consuming the total produced energy and also moving the remaining loads from the peak to the off-peak period.

(a) Energy consumption: 26th of September 2023.

(b) Energy consumption: 29th of September 2023.
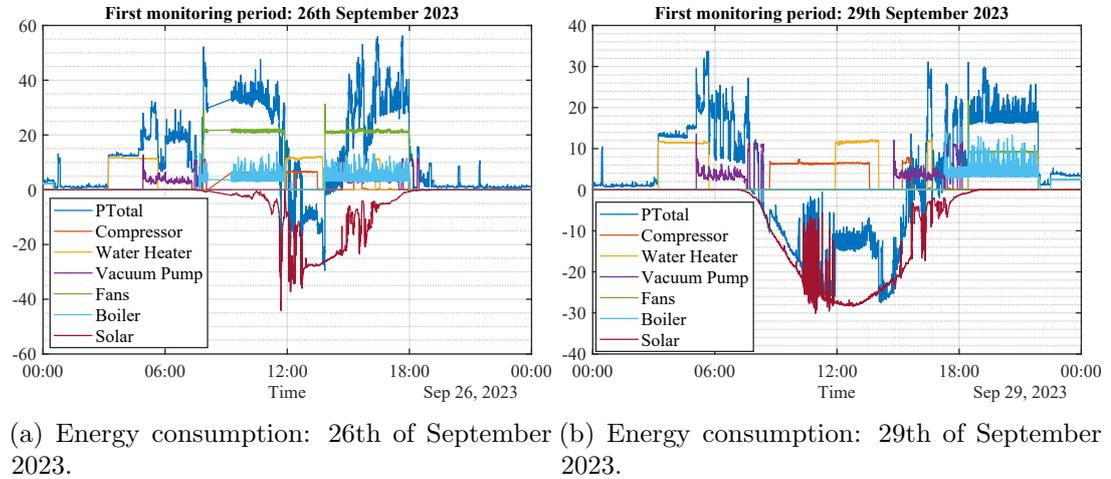
Figure 6.5: Farm I: energy consumption during the development period.

The inefficient energy management in the farm is evident in Figure 6.5, where flexible loads are not actively controlled in order to maximise the self-consumption and/or exploit the dual tariff system. For example, it can be seen from Figure 6.5a, that during the 26th of September 2023, the biomass facility was used in the early morning and late afternoon, i.e., outside the high solar production period, resulting in significant imports from the grid, while at the same time energy which could have been used within the farm is exported to the grid. A similar practice is observed in Figure 6.5b, where during the 29th of September 2023, again the biomass facility was operated in the evening hours after the solar production, and therefore a significant amount of energy was imported from the grid, while excessive energy was exported during the solar hours.

6.2.2   Energy intensive activities disaggregation

Sub-metering equipment, which can be used to monitor individual loads, is rarely installed in farms due to installation costs, the intrusiveness of the installation procedure, and the complexity of installing meters in farming equipment. On the other hand, utility smart meters, collecting power readings in 30-minute intervals (see Section 2.3), are commonly available in agricultural settings. Therefore, the applicability and scalability

of the proposed methodology through NILM was explored, given the envisaged absence of sub-metering devices.

6.2.2.1 DL-based NILM models

The energy-intensive activities of the farm that were identified — and correspond to over 70% of the total energy consumption of the farm — were disaggregated using two state-of-the-art NILM algorithms [23]. The first disaggregation algorithm used is the seq2subseq model [64] that has been demonstrated to have good performance in transferability and generalisability tests (see Subsections 4.3.1.4 & 4.3.1.5), and three-phase load disaggregation in residential settings (see Subsections 3.3.1 & 5.3.2). The second disaggregation algorithm used is the WaveNet model [155], which has been applied for the disaggregation of dairy farming activities [129] and has also been shown to perform well in residential settings (see Subsection 5.3.2). The two algorithms were selected to be benchmarked against each other due to their proven high disaggregation accuracy, their ability to transfer knowledge across different datasets, and due to their different approaches in load disaggregation. More specifically, for the seq2subseq method, given a time window of length ω, the algorithm targets a time-window with a length of $\omega/2$ (see Subsection 3.2.3.1) and for WaveNet, given the same time-window, the algorithm targets the middle point of the time window. It is expected that as seq2subseq targets the half of the time window, it will converge faster — i.e., reduced computational costs — but with reduced accuracy, whereas the WaveNet seq2point model that targets only one point, although converging slower — i.e., higher computational costs — will achieve higher accuracy. As the two approaches are based on supervised learning, FIELD dataset (see Subsection 6.2.1.1), has been used for training purposes for both DNNs (see Subsection 6.3.1.1). Both aggregated and submetered readings were downsampled to the new frequencies with training sets generated from the open dataset and testing sets generated from the monitored sites.

6.2.2.2 DNNs' adaptation and post-processing

Although transfer learning has been highlighted [41] as a viable approach for at-scale load disaggregation, in [129] it was concluded that due to largely non-standardised and distinct nominal power levels across dairy equipment, transfer learning was performing poorly at the disaggregation problem. This was also evident in Chapter 4 and more specifically in transferability test II., where although through transfer learning for EV loads with different characteristics were accurately classified, discrepancies across power levels and duration of events of the training and testing sets resulted in under and/or over estimation of the actual energy consumption when using regression approaches. In order to mitigate this issue and enable transfer learning across farms with similar dairy equipment but with varying load signatures, the following adaptations are proposed.

- The loss function for both DNNs is changed from the commonly MAE (Equation 2.12), used in generic regression problems, to the NILM-specific metric MR (Equation 2.15). As MR measures the overlap between true and estimated energy values (see Subsection 6.3.1.2) performance of the network is expected to increase.

- As it was already stated, transfer learning across similar equipment but with different nominal power levels shows good performance on the classification problem, i.e., identifying the activation period, but rather poorly in the regression problem as the energy consumption is either under- and/or over-estimated. To mitigate this effect the nominal power level ($\tilde{P}_{nom}^i$) of individual loads (i) collected through the qualitative survey and the nominal power level (P_{nom}^i) of the equipment (i) in the training set were used to estimate the ratio:

$$A = \frac{\tilde{P}_{nom}}{P_{nom}} \tag{6.1}$$

  The output of the regressor ($P_{out}^i(t)$) for each load i was then multiplied by the factor A. The resulting timeseries were then adjusted so that $\tilde{P}_{out}^i(t) = min\{A \times P_{out}^i(t), P_{agg}(t)\}$, where $P_{agg}(t)$ is the aggregate power at time t.

The submetered data streams collected across the three farms were used as a bench-

mark for the proposed methodology. Further to that, the disaggregation performance when using transfer learning was benchmarked against training and testing on the same site.

6.2.2.3 Three-phase vs. single-phase NILM

By exploiting the three-phase sub-metering data available (as discussed in Chapter 3), and the symmetry of the loads, for each of the two adopted models, two disaggregation approaches were followed:

I. the energy intensive activities are disaggregated through the aggregated three-phase signal; and,

II. the disaggregation is performed on each phase separately.

Although, the second approach requires the disaggregation of three data streams instead of one, it is expected that the disaggregation accuracy will be higher (see Section 3.4) with only a slight increase in the computational time[2]. In the later approach, i.e., using each phase separately, a post-processing step was followed where non-simultaneous activations of three-phase loads were discarded as FPs (considering that all disaggregated loads are symmetrical).

6.2.3 Production estimation

As described in Subsection 6.2.1, solar generation was monitored in the testing site and used as a benchmark for the production estimation. In order to create a day-ahead load scheduler and optimiser, the next day's granular solar generation was estimated on 30-minute intervals based on the availability of granular 30-minute weather data. A similar approach as described in Subsection 5.2.3 to estimate the solar generation was followed. Forecasted weather data obtained through the Met Office API [156] were used for solar energy generation modelling. The following parameters were used:

- temperature;

---

[2]Please note that the same model is used in all three phases so no additional training time is required in the second approach but rather only testing on three time-series.

- relative humidity; and,

- the direct and diffuse solar irradiance.

The PV estimation was based on the analysis in Subsection 5.2.3.

6.2.4 Load scheduler

The load scheduler is based on a day-ahead optimiser of the time of use of energy-intensive day-to-day equipment in a dairy farm that is considered flexible based on the qualitative data. The optimiser takes into account the quantitative targets and constraints, i.e., the available renewable energy production, the flexibility analysis of the disaggregated activities, the electricity costs and the estimated carbon footprint for the day ahead, and outputs the optimal time to start a particular activity. Note that the intensity of activities (e.g., the number of cows milked) is not optimised, and was considered fixed, which is a practical constraint from the qualitative study.

6.2.4.1 Carbon footprint forecasting via Prophet model

To predict the emitted carbon footprint of the dairy activities and explore the potential of reducing the carbon intensity through load shifting, the granular (30-min) regional carbon footprint as collected through the ESO [96] was used. Data were collected for a period from June 2018 (introduction of the carbon intensity API) until the end of the monitoring period of the study (September 2024). The regional carbon footprint for the South of Scotland region is presented in Figure 6.6. As it can be observed, the carbon footprint varies throughout the day and, more importantly, during the off-peak period when consumers are incentivised to move their flexible loads. As the carbon footprint is not available for the day-ahead, a short term 24-hour forecasting was performed. The Prophet model [157] was used due to its simplicity, its wide applicability, and its ability to decompose the signal into trend, seasonality, and holidays components. It was adapted as follows. The linear trend with changepoints was used for the trend factor with automatic changepoint selections to reflect the changes that occur due to different units contributing to the generation mixture of the grid. The seasonality component was tuned to model the daily, weekly and monthly patterns that affect electricity
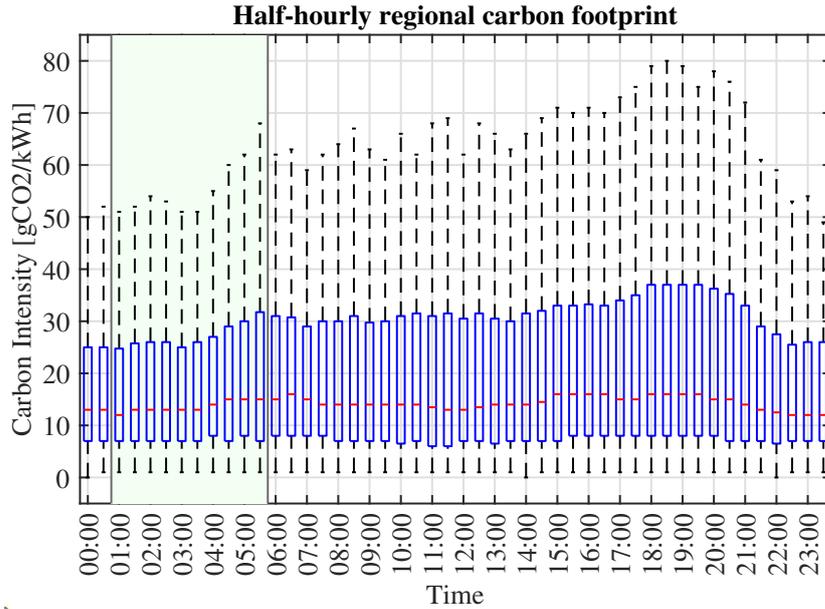
Figure 6.6: South Scotland regional granular carbon footprint over the monitored period. The off-peak tariff period is shaded green.Day tariff (06:00 – 01:00) of $\sim$ 28p/kWh and night tariff (01:00 – 06:00) of $\sim$ 17p/kWh.

consumption. This was done by selecting the additive seasonality component. Lastly, bank holidays observed in the UK were also used as input to "holidays" component, as during these periods abnormal energy consumption and thus carbon footprint occur. More specifically, all occurrences of each individual holiday, both in the past (since 1st of June 2018) and in the future (up to and including September 2024) were inputted as a single data frame to the model. Lower and upper windows (i.e., the period before and after a holiday, such as Christmas Eve) of each individual holiday were included as parameters to the model in order to better model the effect of the days before and after a bank holiday.

6.2.4.2   Multi-objective optimisation

The load scheduler works through multi-objective optimisation, where : (i) the total utility cost for the farm and (ii) the total carbon footprint for the farm are jointly minimised. The function presented in this subsection can be used interchangeably for both the utility cost and the carbon footprint if the latter is considered as cost.

The total cost (in terms of utility or carbon footprint) for the farm is given by:

$$C^i = \sum_{t=t_0}^{t_n} \left(E_{imp}(t) \times C_{imp}^i(t) \right) - \left| E_{exp}(t) \times C_{exp}^i(t) \right|, \tag{6.2}$$

where i is either the utility cost or carbon intensity of the grid, t_0 is the start time, t_n is the end time, $C_{imp}^i(t)$ is the import cost (utility or carbon) at time t, $C_{exp}^i(t)$ is the export cost (utility or carbon) at time t. Utility import and export cost is the import and export tariff at time t. Carbon import cost is the regional granular 30-min generation carbon footprint of each kWh consumed, and the carbon export cost is the difference between the generation carbon footprint at time t and the carbon footprint of each kWh produced from the local energy sources. In the case of PV panels, it is assumed that the carbon footprint is zero. $E_{imp}(t)$ and $E_{exp}(t)$ are the energy import and export at time t, given, respectively, by:

$$E_{imp}(t) = max\{0, E_c(t) - E_p(t)\} \tag{6.3}$$

$$E_{exp}(t) = min\{0, E_c(t) - E_p(t)\}. \tag{6.4}$$

In the above equation, the energy production $E_p(t)$ at t, is calculated by:

$$E_p(t) = \sum_{i=1}^{m} E_i^p(t) = \sum_{i=1}^{m} \int_{t-\Delta t}^{t} P_i^p(\tau)d\tau \tag{6.5}$$

where $E_i^p(t)$ is the energy production of the i-th RES out of the m RES, and $P_i^p(t)$ is the power sample of the i-th RES at t obtained as described in Subsection 6.2.3. The energy consumption $E_c(t)$ at time t is calculated as:

$$E_c(t) = B_L(t) + \sum_{i=1}^{n} E_i^c(t), \tag{6.6}$$

where $B_L(t)$ is the base load, i.e., steady state loads that are running throughout the day on a constant power level and non-flexible loads that cannot be shifted throughout

the day, $E_i^c(t)$ is the energy consumption of the i-th load out of n loads and $P_i^c(t)$ is the power sample of the i-th load at t. Considering that a load can either have a variable power level profile or a constant power level profile, the total energy required by appliance i is given by:

$$E_i^c = \sum_{t=t_i^{start}}^{t_i^{start}+t_i^{duration}} \int_{t-\Delta t}^{t} P_i^c(\tau)d\tau, \qquad (6.7)$$

where Δt is the sampling interval, t_i^{start} is the starting time of the activity and $t_i^{duration}$ is the total duration of the event estimated in the following way. The starting time was estimated based on the required time for a process to be completed — as obtained through the qualitative survey — and the duration of the activity. For processes that do not involve heating and/or cooling $t_i^{duration}$ is the time required to complete a specific operation cycle of equipment as obtained through the semi-structured interviews and validated during the development period from the quantitative sub-metering data (see Subsection 2.1.1). For heating and/or cooling processes, the duration of the activation of the heating/cooling elements is estimated through thermodynamic analysis. Given the energy required to increase a substance by 1 degree as: $Q = m \times c \times \Delta T$ $[J]$, where m is the water mass in kg, $c = 4184$ $[^{J \times Kg}/_K]$ is the specific heat capacity of the water and $\Delta T = T_{stop} - T_{start}$ $[K]$ is the temperature difference between the start and stop states of the water heating/cooling process, and assuming a constant heating/cooling element with a power level of P $[W]$, the total duration of the heating and/or cooling process can be calculated as:

$$t_i^{duration} = \frac{Q}{P_E} = \frac{m \times \frac{c}{3600} \times \Delta T \times L}{P \times COP}, \qquad (6.8)$$

where COP is the coefficient of performance as obtained through the qualitative survey and the manufacturers' specifications; with $COP = 1$ for standard resistive heating elements or greater than 1 for heat pumps/compressors.

6.2.4.3   Optimisation through exhaustive search

In summary, under farming equipment power constraints, and based on the variable — peak/off-peak — import energy tariff, the proposed optimisation aims to minimise the joint utility and carbon cost (see Equation 6.2). The mixed-integer nonlinear multi-objective minimisation problem can be written as:

$$min_{t_{start}^j}\{C^{utility}(t_{start}), C^{carbon}(t_{start})\} \tag{6.9}$$

where $C^{utility}(t_{start})$ and $C^{carbon}(t_{stop})$ are given by Equation 6.2, over all possible combination of start times, such that:

$$t_j^{lower,start_{req}} \leq t_j^{start} \leq t_j^{upper,start_{req}}, \quad \forall j \in \{1,...,k\} \tag{6.10}$$

where $t_j^{lower,start_{req}}$ and $t_j^{upper,start_{req}}$ are, respectively, the earliest and latest possible start of the j-th out of k flexible activities in order to be completed within the required time window, as obtained through the qualitative constraints.

The mixed-integer non-linear optimisation problem was solved using an exhaustive search method to avoid getting suboptimal results due to entrapment into local optima. The selection of an exhaustive search was selected over other heuristic and meta-heuristic approaches — that cannot guarantee a global optimum solution — in order to isolate and assess the effect of the disaggregation accuracy on the load scheduler. The number of possible combinations of the optimisation approach capped by Equation 6.11 as:

$$C_n^k \leq n^k, \tag{6.11}$$

where n is the total number of possible time slots for the start of each activity, which is capped by the granularity of the PV production estimation, here 30-min, i.e., 48 time slots for an activity that can happen any time during the day, and the total number of flexible loads (k), in the case of Farm I and III three (water heaters, fans, and boilers) whereas for Farm II two (chiller and water heaters). The number of time slots n is further dynamically constrained on a day-to-day basis by Equation 6.10 that dictates

the earliest and latest possible start time (e.g., loads that are related to activities that happen evenly twice a day, such as milking, can be shifted by up to 12 hours).

6.3   Results

In this section, we first present the results of energy disaggregation comparing the two different NILM approaches as discussed in Subsection 6.2.2, and then present our results obtained via the proposed load scheduling optimisation, described in Subsection 6.2.4.

6.3.1   Load disaggregation of energy intensive equipment

6.3.1.1   Experimental setup

The parameters used for the seq2subseq model were selected as in Subsection 3.2.3. The parameters selected for the WaveNet model were the same as the default model parameters in [155]. More specifically, the input and output sequences were 127, 32 skip channels were used, and the batch size was equal to 128. The training was performed with 300,000 iterations using a decaying learning rate. Both DNNs were adapted as discussed in Subsection 6.2.2.2. The default loss function of MAE was changed to the NILM-specific MR (see Equation 2.15 and the discussion afterwards) for both WaveNet and seq2subseq networks. Further to that, a post-processing step to mitigate over- and/or underestimation of energy usage was performed.

A mixture of farms available in the FIELD dataset (see Subsection 6.2.1.1) was used for the disaggregation of the energy intensive activities by label matching the submetered equipment based on qualitative findings from the co-creation approach and quantitative analysis of the load signal, as summarised in Table 6.3. A subset of the farms and the monitored equipment were selected based on matching activities observed in the monitored farm/testing site.

To evaluate the proposed post-processing for NILM when training on data from farms with different equipment load profile magnitudes and durations, results when the models are trained on the sub-metered data from the three test farms are also included. The period between the 15th of September 2023 and the 9th of October 2023, was

Table 6.3: FIELD data comparison with test farms. ID refers to the farm ID in the FIELD dataset.

Farm	Yearly imports	Yearly exports	Testing equipment	Training equipment: FIELD ID
I	40 MWh	15 MWh	Compressors	Compressor: 18
			Water heaters & Boilers	Heavy duty cleaner: 3 Water treatment: 21
			Vacuum pump	Vacuum pumps: 18, 23
			Calf house ventilation	Ventilation: 5 Fan cowshed: 4 Fan 1 − 4: 11 Fan barn: 23 Ventilation 1 − 2: 27, 28
II	82.5 MWh	11.5 MWh	Chiller	Compressor: 18
			Water heaters	Heavy duty cleaner: 3 Water treatment: 21
			Vacuum pumps	Vacuum pump: 18, 23
III	150 MWh	3.5 MWh	Compressors	Compressor: 18
			Water heaters	Heavy duty cleaner: 3 Water treatment: 21
			Vacuum pumps	Vacuum pump: 18, 23
			Calf house ventilation	Ventilation: 5 Fan cowshed: 4 Fan 1 − 4: 11 Fan barn: 23 Ventilation 1 − 2: 27, 28
			Water pumps	Water pump: 21

used for training, whereas the rest of the dataset was used for testing and validation purposes.

For each of the two approaches, disaggregation was performed both on the per-phase (ϕ) signal and on the aggregate signal (agg). Data of different granularities have been used to test the effect of the disaggregation accuracy on the load optimiser. Specifically, six different levels of granularities were used for the disaggregation of the main consumers of the farm: low-frequency stream with samples every 1 second, 10 seconds and 1 minute data and very low-frequency stream with samples every 5, 10, and 30 minutes. The data streams in all frequency levels contained both the aggregated and the per-phase power signal. As the collected data streams from the testing site and the data streams available from the public dataset were sampled in 1-second intervals, both aggregated readings and submetered readings were down-sampled to the selected granularity.

6.3.1.2   Performance measures

Accuracy of the disaggregation was measured using the NILM-specific metric MR (Equation 2.15) as discussed in Subsection 2.4.1.

The speed performance of the two DNNs has also been explored by comparing them when trained and tested on the same machine under the same load. The machine was equipped with an Intel I9-10900X processor, 32GB RAM and an NVIDIA GeForce RTX 3080 12GB graphics card.

6.3.1.3   Disaggregation Results

The results obtained through the disaggregation of the energy intensive equipment (as described in Subsection 6.2.1) of the farms are presented in Figures 6.7, 6.8, and 6.9 for Farms I, II, and III, respectively. Disaggregation across the three testing sites follows a similar pattern (demonstrating good transferability performance across all three settings). It is worth noting that accuracy is higher in Farm I (smaller farm), which can be attributed to the lower signal noise from other energy activities in the farm. On the other hand, Farm II demonstrates the lowest disaggregation accuracy, but with all methods converging to relatively close values. The worse performance at Farm II was the result of the seasonal operation of the farm, which resulted in multiple FPs of the milking equipment (chillers, water heaters, and vacuum pumps) during the shutdown period that were a result of other aperiodic activity in the farm. Lastly, the large-scale Farm III demonstrates similar performance patterns with Farm I regarding the disaggregation accuracy, though being lower by an average of 5%.

As expected, it can be observed that the Wavenet network outperformed the seq2subseq model in most scenarios for all the farming equipment. Furthermore, disaggregation on a per-phase basis improved the disaggregation accuracy due to the reduced FPs that occurred, based on the proposed mitigation strategy as discussed in Subsection 6.2.2.

The effect of the data granularity is evident across all the dairy equipment with boilers (see Figure 6.7e) and fans (see Figure 6.7d & 6.9d) being the most susceptible

(a) Compressor

(b) Water heaters

(c) Vacuum pumps
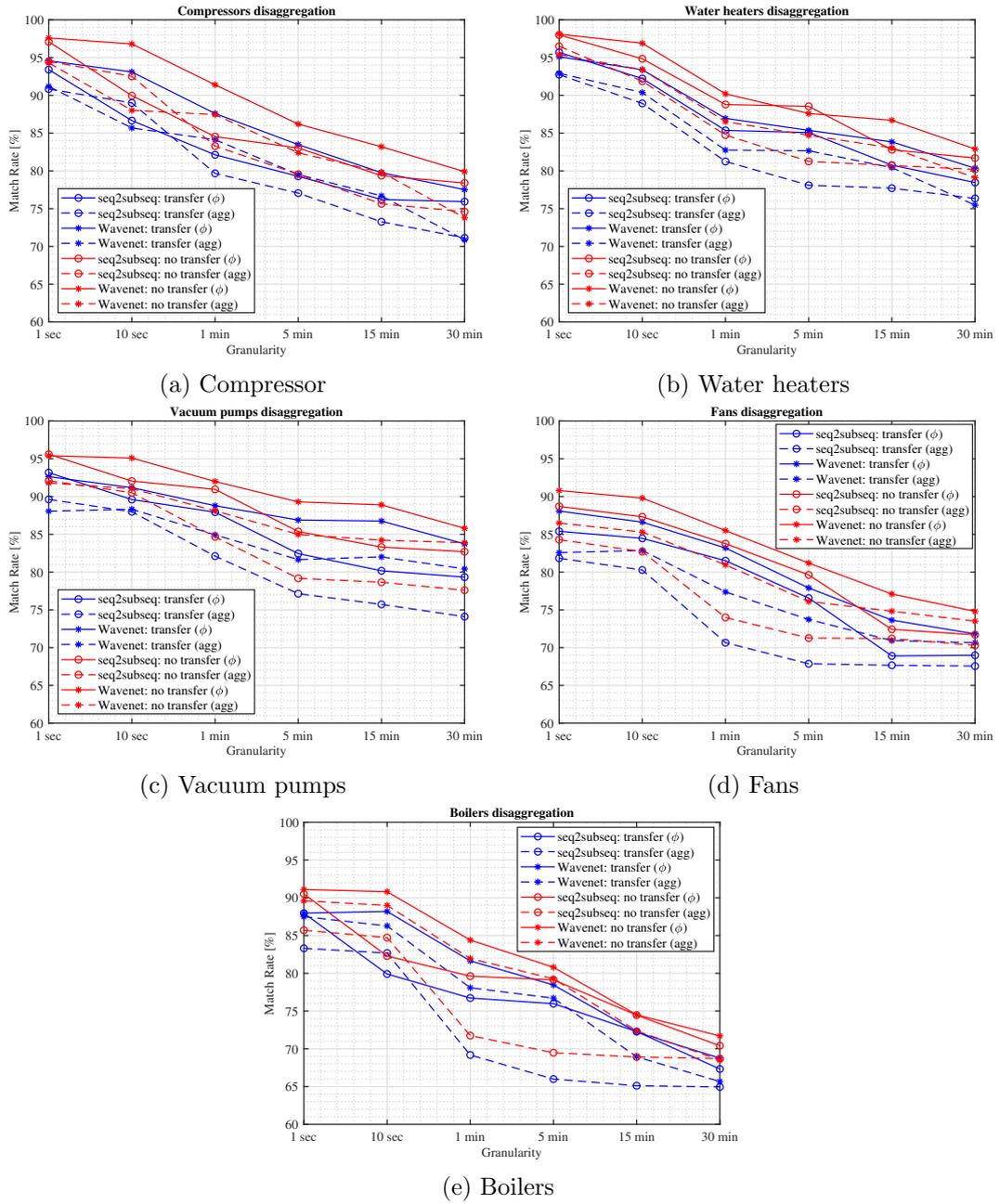
(d) Fans

(e) Boilers

Figure 6.7: Farm I: load disaggregation of dairy equipment.

to reduced disaggregation accuracy. This can be attributed to the more sparse and aperiodic activation of the boilers and fans (running on demand and under different profiles).

Lastly, in both disaggregation approaches, i.e., transfer and non-transfer learning,

(a) Chiller

(b) Water heaters

(c) Vacuum pumps
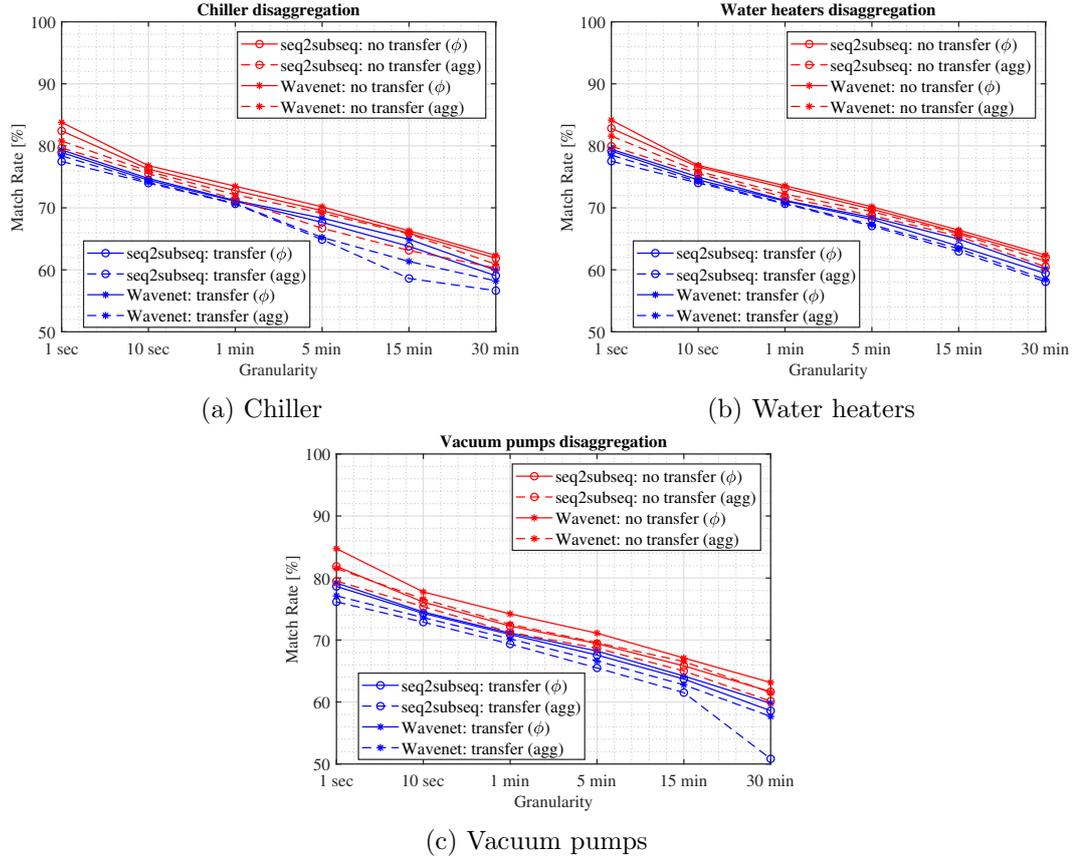
Figure 6.8: Farm II: load disaggregation of dairy equipment.

the disaggregation accuracy deviates from $\sim 2\%$ to $\sim 5\%$, and therefore NILM, where only the aggregate is monitored, is a viable, practical, cheaper method that preserves privacy.

6.3.1.4 Complexity Analysis

The training and testing time of the two DNNs was compared, and the results under all tested scenarios are presented in Figure 6.10. ΔSpeed was estimated by comparing the sum of the training and testing time for the two different DNNs as described in Subsection 6.2.2.1 using the same resources (see Subsection 6.3.1.2). It can be estimated as:

$$\Delta\text{Speed}^{i,j,k} = \frac{T_{s2s}^{i,j,k} - T_{s2p}^{i,j,k}}{T_{s2p}^{i,j,k}} \tag{6.12}$$

(a) Compressor

(b) Water heaters

(c) Vacuum pumps

(d) Fans

(e) Boilers
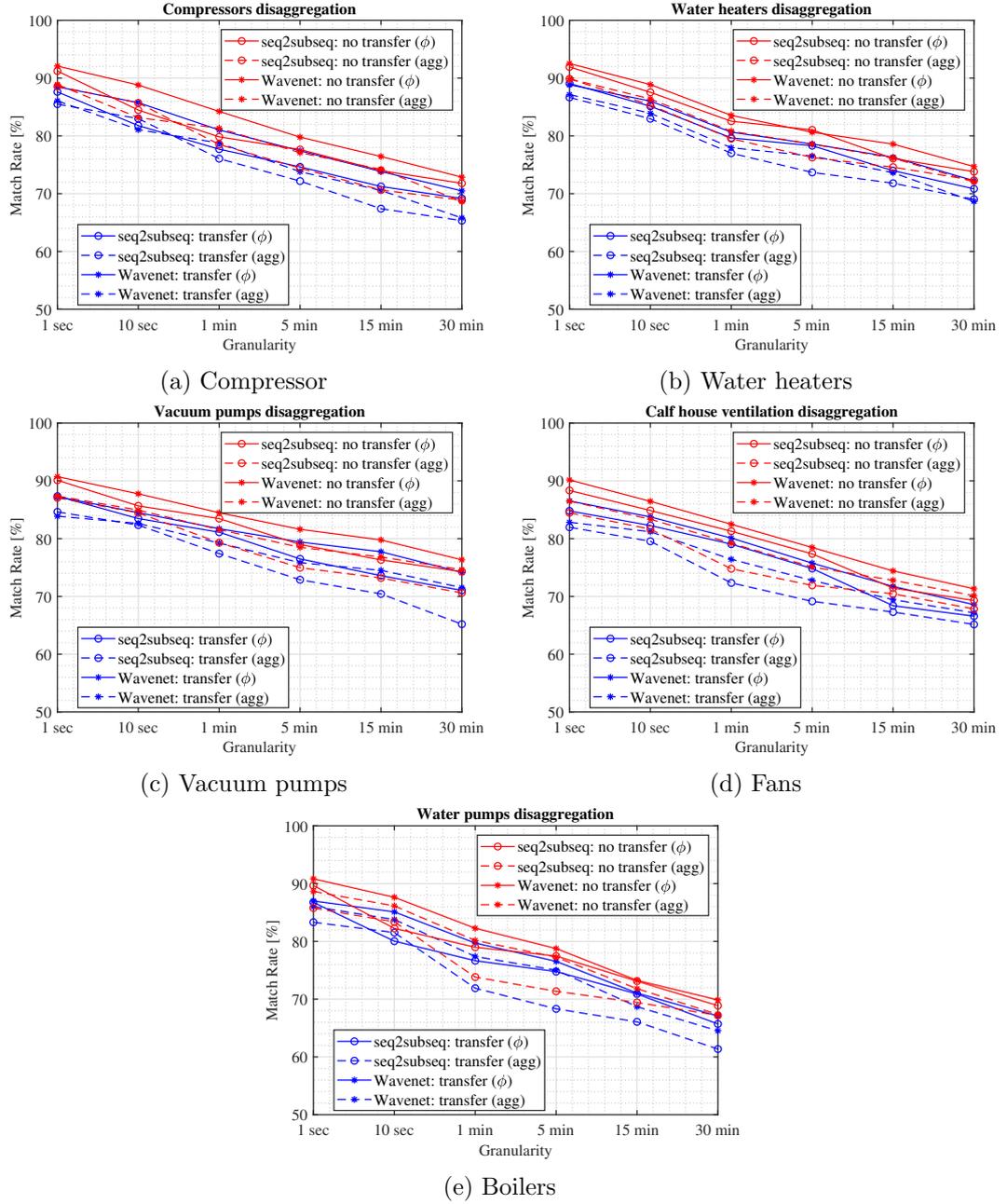
Figure 6.9: Farm III: load disaggregation of dairy equipment.

where $T_{s2s}^{i,j,k}$ and $T_{s2p}^{i,j,k}$ are the average sum of the training and testing time for seq2subseq and seq2point (WaveNet) networks, respectively. i is the group of equipment (i.e., compressors, water heaters, vacuum pumps, fans, boilers), j is either aggregate (agg) or per phase signal (ϕ), and k is the granularity that ranges from 1 second to 30 minutes.
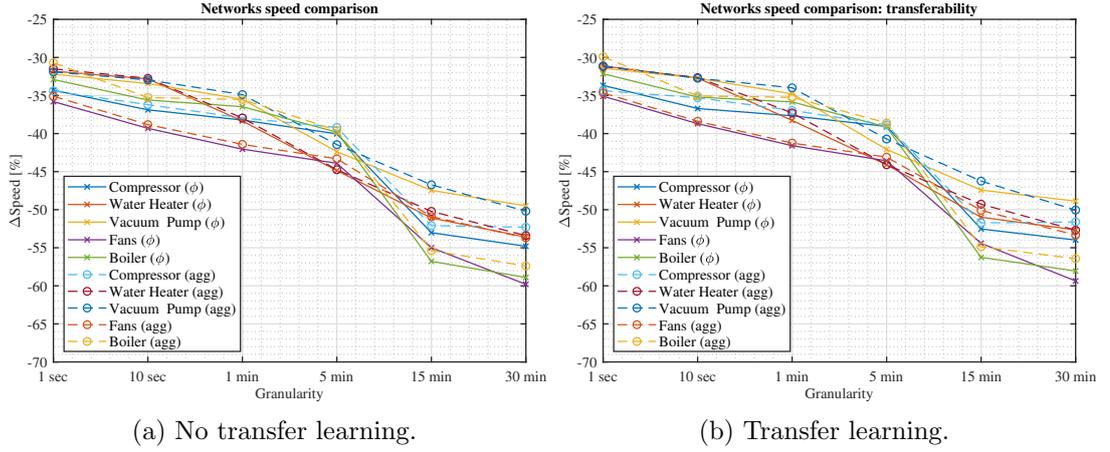
(a) No transfer learning.   (b) Transfer learning.

Figure 6.10: Network average speed comparison across the two DNNs under different loads and different granularities. Please note that the speeds are the averages across the three testing farms. Compressor loads are grouped together with the chiller load of Farm II due to their load similarity and training time.

Though for low frequency (1-sec) data, the seq2subseq outperforms WaveNet in terms of speed by approx. $30\% - 35\%$, this is substantially increased for lower frequency data, with seq2subseq model being as much as 60% faster under a 30-min data scenario. Across the different appliances, there is a slight variation in terms of speed that can be attributed to the number of activations and the time that the network takes to converge. In summary, the WaveNet network shows higher MR than the seq2subseq network, but at the cost of considerably higher computational time. Hence, by combining Figures 6.7, 6.8, 6.9, and 6.10 the trade-off between disaggregation accuracy, the DNN selection (complexity), and data granularity can be derived.

6.3.2   Load Scheduling

The optimisation results obtained following the methodology presented in Subsection 6.2.4 based on the exhaustive search multi-objective optimiser that prioritises the minimisation of the total cost (see Equation 6.2), across the three farms over the 12 months testing period are presented in Figure 6.11. Various disaggregation scenarios are considered, with the day-ahead renewable and the day-ahead carbon footprint estimation (see Subsections 6.2.3 & 6.2.4.1).

As expected, the optimiser performance (the highest energy cost reduction) is the

(a) Farm I

(b) Farm II

(c) Farm III
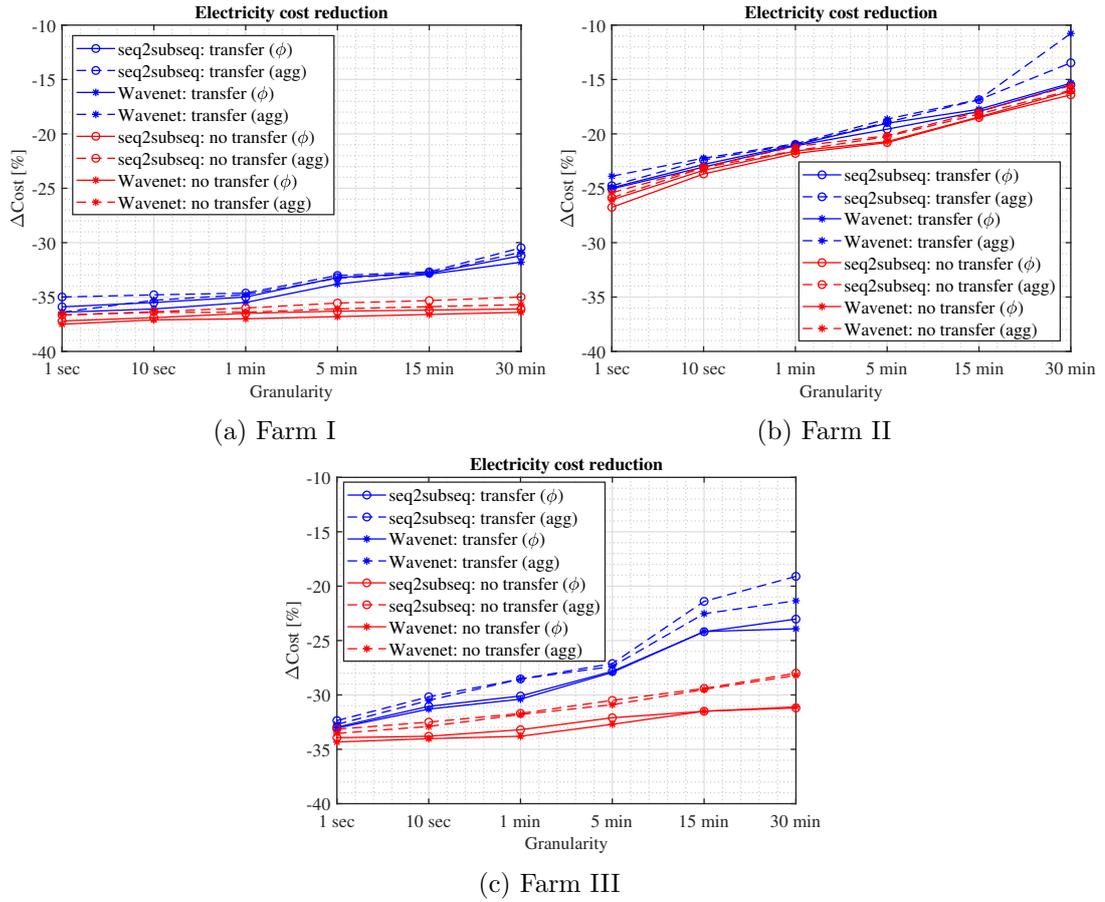
Figure 6.11: Year over year electricity cost reduction compared to the pre-study consumption shown in Table 6.3 across the three farms.

best when using a non-transfer learning scenario and with higher granularity data. It is worth noting that the calculated maximum possible energy bill reduction for Farm I was 41% (by self-consuming all produced energy and also moving the remaining loads from the peak to the off-peak period), which is slightly higher than the proposed system's cost reduction of up to 37.5% through load disaggregation when training and testing on the same site (i.e., a rate of $\sim 91.5\%$) on a low-frequency 1-sec scenario (baseline scenario). The divergence between the theoretically maximum possible savings and the savings with the proposed optimiser can be attributed to the following:

I. not all activities can be shifted to either nighttime or a flexible activity may last longer than the off-peak period;

II. the day ahead solar production estimation as described in Subsection 6.2.3 is not 100% accurate; and,

III. the renewable estimation is based on 30-min forecasted weather data so the optimiser and the start points of each activity is limited by the solar estimation granularity.

Furthermore, the optimiser is less efficient with lower frequency data and in the transfer learning scenario; e.g., a bill reduction in the range of 30.5% – 32% for 30-minute data (i.e., 81.3% – 85.3% compared to the baseline scenario). The divergence between the transfer and no-transfer scenarios is expected due to the non-equally accurate disaggregation of individual loads ($\sim$ 75% compared to $\sim$ 95%), which affectsthe accuracy of the estimation of the expected actual energy consumption, and hence the duration and the starting point of the activity. Energy self-consumption through load shifting contributed approx. 42% to cost reduction, whereas shifting during nighttime contributed to the remaining 58% of cost reduction.

The carbon footprint optimisation was more robust to data granularity across all three farms, with the reduction across all simulated scenarios for Farm I varying only slightly in the range of 28.4% – 30.1% for low and high granular scenarios, with self-consumption contributing approx. 23% and shifting activities out of peak carbon times the remaining 67%. The effects of the different disaggregation approaches on carbon footprint reduction under different granularities are summarised in Figure 6.12.

As can be observed from Figures 6.11 & 6.12 in Farm II, the optimiser performance was significantly reduced compared to Farm I. This can be attributed to the following:

I. there are fewer flexible loads present in the farm;

II. the farm is shutting down during autumn and winter with limited activity taking place; therefore, the energy produced during that period cannot be self-consumed; and,

III. the disaggregation performance (see Figure 6.8 & 6.12b) is poorer.

Farm II also achieved lower carbon savings in the range of 12.9% – 14.8%. On the other hand, Farm III, although significantly larger with more complex energy profile that

(a) Farm I

(b) Farm II

(c) Farm III
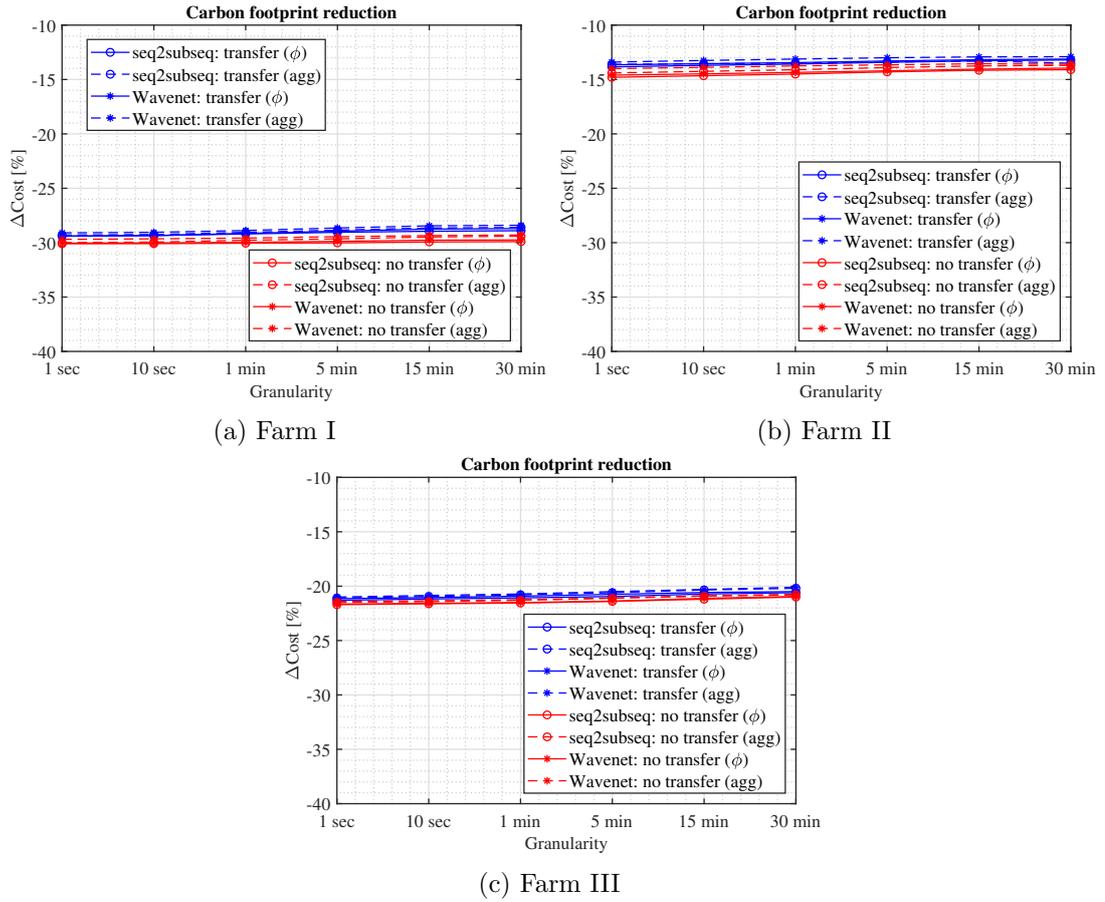
Figure 6.12: Year-over-year carbon footprint reduction compared to the pre-study carbon footprint across the three farms.

affected disaggregation performance (see Figure 6.9 & 6.12c), achieved savings in line with Farm I, with the lower cost savings driven mainly from the reduced renewable energy availability surplus. The savings in carbon ranged from 20.1% to 21.7%.

More importantly, the performance of the optimiser in reducing the carbon footprint of the farm is affected to a lesser extent when the NILM-enabled solution with 30-min data is used instead of submetering for Farm I. Though the optimiser's performance in the lower frequency scenarios is lower, the magnitude of error due to disaggregation was reduced while being propagated through the optimiser. More specifically, the optimiser using 1-sec data and dedicated submetering, achieved on Farm I on average 37.5% bill reduction, whereas when using 30-min data with no submetering under a transfer scenario, the achieved savings were on average 32%, i.e., a ratio of 85.3% in utility cost

savings when comparing the two data granularities. By comparing the performance of the cost optimiser (85.3%) to the disaggregation accuracy — which for 30-min data was on average $\sim 75\%$ — it can be deduced that the optimiser is not linearly affected by the disaggregation accuracy. Further to that, the performance of the optimiser in reducing the carbon footprint of the farm is affected to a lesser extent, as when comparing the use of 30-min data to submetering for Farm I, the performance was 94.3%. Therefore, it can be concluded that even with suboptimal disaggregation performance, the optimiser is able to achieve similar utility costs and carbon savings when compared with a highly granular submetered scenario.

(a) Farm I

(b) Farm II

(c) Farm III

Figure 6.13: Representative ante and post load profiles across the three farms.

A sample of representative ante and post load profiles of the three farms included in the testing set is included in Figure 6.13. In all three farms, load scheduling increases self-consumption and reduces imports during peak hours by moving flexible loads to non-peak hours. Comparing the three farms, Farm I benefit the most from the load

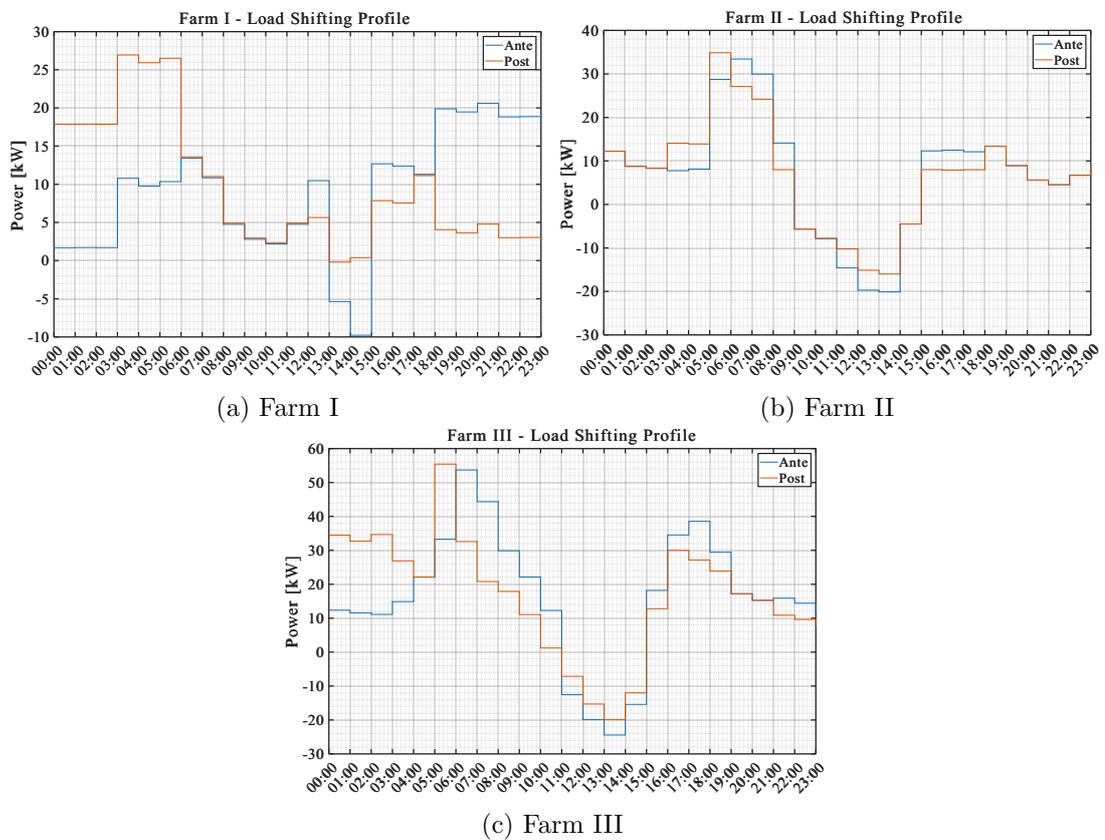scheduler due to the presence of multiple flexible loads. Farms II & III, although both operating a rotary parlour, do not benefit the same from load scheduling. This can be partly attributed to the different equipment available on the two farms and to the different farming practices. As can be observed in Farm II, there is a prominent morning milking round, whereas the second afternoon milking round is less intense compared to Farm III.

In summary, under all scenarios, the disaggregation on a per-phase basis instead of the aggregate of the three phases achieved improved cost savings in the range of 0.1%–2% and improved carbon savings in the range of 0.1% – 0.8%. It is worth noting that due to the lower variability of the carbon footprint (see Figure 6.6) compared to the electricity prices, the cost reduction is proportionally higher compared to the carbon reduction across all three farms, but the carbon reduction is more immune to different granularities and disaggregation models.

6.4   Discussion & conclusions

Figure 6.14 summarises the results of the load optimiser for the per-phase disaggregation scenarios of Farm I. The following conclusions can be made:

- Through the proposed system, a capital intensive approach that involves the installation of specialised per-phase submetering equipment with sampling rates of 1 second that requires no disaggregation, can achieve utility bill savings of $\sim 37\%$, i.e., achieving 92% savings compared to the theoretical maximum of $\sim 41\%$ while at the same time achieving carbon footprint reduction from electricity usage of up to 30.1%.

- Approaches that involve the installation of specialised per phase metering equipment that only collect 1-second readings of the aggregated consumption (and therefore reducing the intrusiveness and the capital intensity of the approach) and through transfer learning achieved almost identical performance with utility bill savings of $\sim 36\%$, i.e., achieving 87% savings compared to the theoretical maximum and carbon savings of 29.9%.
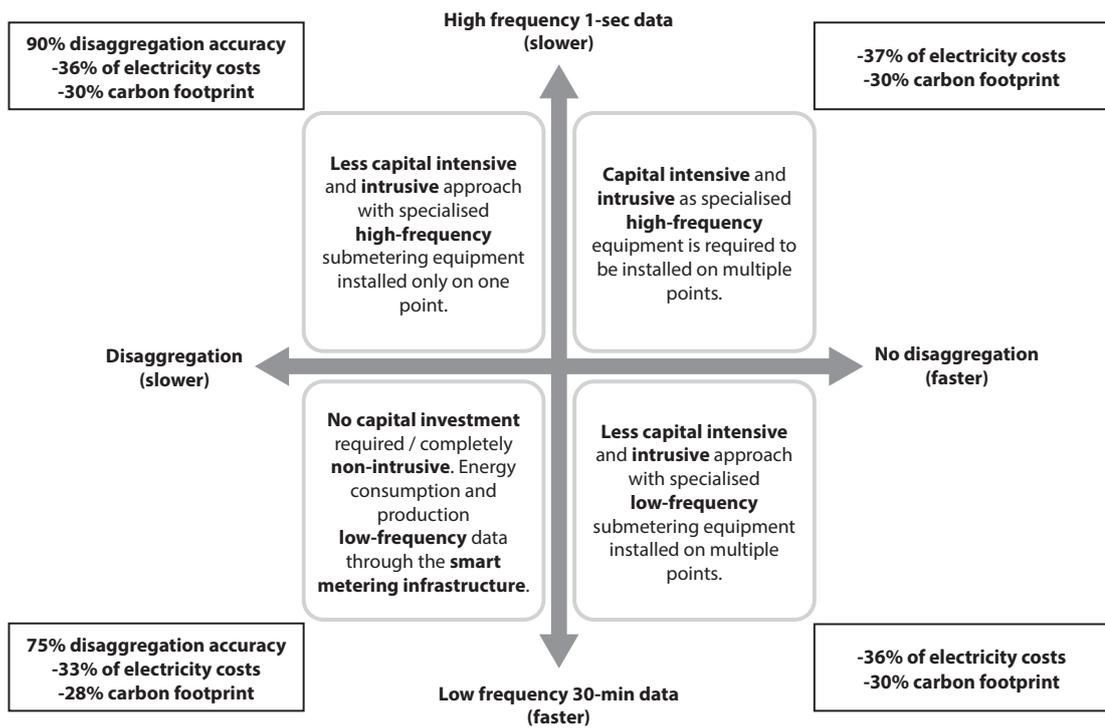
Figure 6.14: Trade-off between data granularity, load sub-metering and cost.

- Similar savings were achieved when specialised per-phase submetering equipment with sampling rates of 30 minutes was installed.

- Lastly, approaches that do not require capital investment, i.e., through the usage of the available 30-min aggregated smart metering infrastructure and disaggregation through transfer learning, can achieve utility bill savings of $\sim 33\%$, i.e., achieving 80% savings compared to the theoretical maximum and reduction in carbon footprint of 28.4%.

A NILM-enabled load scheduling and energy management methodology for improved handling of available renewable energy in dairy farms was presented that takes into account the constraints of causality and time of use of farming activities. The novelty lies in the holistic integration of the various sub-systems, including highly granular solar production and grid-mix based carbon footprint forecasting, which was co-created with stakeholders to define context and therefore resulted in a comprehensive study keeping all stakeholders in the loop. Furthermore, novelty lies in the proposed NILM approach

that can adapt to the heterogeneity of farm equipment load signatures and labelling, which is currently lacking in the NILM literature. The proposed methodology was benchmarked against the ideal capital-intensive approach, where specialised, intrusive monitoring equipment is installed in farms, demonstrating that utility is preserved during load scheduling optimisation via the NILM-enabled approach at up to 30 min smart aggregate granularity.

Although in this chapter, the performance of a NILM-enabled co-created load-scheduling system using transfer learning, very-low frequency utility smart metering infrastructure, and granular regional forecasted carbon footprint data was proposed, and tested across three farms, individual system components were not optimised in terms of computational performance. Further research should be performed towards optimising and adapting each individual subsystem to different agricultural environments, including different dairy settings such as confined dairy houses, as well as in settings with diverse renewables penetration, including wind, anaerobic digestion, and hydro. In principle, different renewables can be integrated in a dairy setting; the diverse generation profiles of different RES, as well as the erratic nature of others that are directly related to physical phenomena, should further be explored. Exhaustive search was introduced in the framework in order to assess and quantify the effect of NILM on load scheduling for the agricultural sector, isolated from the performance of other components. Although this approach enables the assessment of NILM, it may not be practical in more complex systems with more loads or states, as the number of possible combinations increases exponentially. Therefore, heuristic and metaheuristic approaches such as particle swarm optimisation should be explored as a more viable and scalable solution. Lastly, different milking technologies that exhibit different power profiles should be explored, including voluntary milking, where the milking activity is largely a stochastic process as cows decide the time and the duration of the extraction cycle.

Chapter 7

Conclusions

7.1 Summary

In conclusion, this thesis demonstrated the potential of energy management and evaluation through NILM-enabled co-created solutions for the residential and agricultural sectors that enhance energy efficiency, responsible consumption, and sustainable production.

More specifically, in Chapter 3 that explored R.Q.1, the benefits of per-phase load disaggregation in three-phase installations for improved appliance-level disaggregation through reduction of noise from interfering loads across various granularity levels were demonstrated. The improvement of NILM-based data mining methods by exploiting three-phase data was supported through the disaggregation of the ECO dataset [4] using a seq2subseq regression-based DL model for various common residential loads including WMs, DWs, FRZs, FRDs, TDs, MWs and CMs. The gain in disaggregation performance when using the appliance phase compared to the aggregate signal was highlighted, especially for lower sampling rates, with commonly mixed appliances such as TDs and WMs benefiting the most from the per-phase disaggregation. The loss in disaggregation performance when using 30-sec and 1-min data compared to 1-sec data was explored with refrigerating appliances demonstrating the highest level of robustness to lower sampling rates (maximum accuracy drop of 10.40% when using 1 min data and aggregate phase) – whereas appliances such as WMs and TDs were extremeley sensitive

172

to granularities with losses up to 58.70% for 1 min aggregate data. The feasibility of leveraging transfer learning in combination with three-phase information to enable labelling of unlabelled datasets was also demonstrated, with a previously unlabelled household with an EV, labelled with an estimated labelling accuracy of over 90%.

This was followed by Chapter 4, which explored R.Q.2, where the seq2subseq network was adapted for EV load disaggregation, with the performance of different generalisability and cross-domain transferability scenarios quantified. The open-access PECAN Dataport dataset [52] was used, with 10 households from the Austin and New York areas used for the adaptation of the seq2subseq model for EV load disaggregation with different evaluation metrics benchmarked against each other. Under all scenarios, it was observed that commonly used metrics such as MAE and SAE misrepresent the actual regression performance compared to NILM-specific metrics such as *Acc* and MR. A MR of up to 81% was achieved when training and testing on the same household using 1 min data. The usage of very-low frequency (15-min) data greatly affected the disaggregation performance, with regression performance reduced from 4.6% to 44.5% depending on the household. Generalisability results demonstrated that for lower frequency (15 min) data, performance was significantly increased due to the availability of additional training data from multiple households with similar charging levels. On the other hand, transferability tests demonstrated that when training and testing on households with different charging levels, the regression-based approach was under- and/or over-estimating the energy consumption of the EV charger. The results of the low-frequency EV disaggregation were used in conjunction with qualitative data to better inform the usage component of LCA models of EVs based on actual usage patterns and answer R.Q.3. The T&E LCA model [74] was adapted to better reflect end-users' practices and the actual carbon footprint of the grid. It was observed that based on the end-users' practices and the available charging equipment, the total resulting carbon footprint of the vehicle can vary from -12.9% up to 3.8%.

In Chapter 5, a comprehensive approach to assess net-positive houses by integrating both quantiative and qualitative data has been presented in order to answer R.Q.4, answering the "what", "why", and "how" of energy prosumption in residential settings.

Challenges in net-zero housing design were identified, emphasising the need for mixed-method data-centric evaluation approaches that align construction assumptions with end-user practices and dynamic energy pricing to achieve net-positive performance. A holistic evaluation framework, where a mixed-methods approach that combines data-driven analysis with qualitative insights to provide a better understanding of urban energy dynamics was introduced. Different perspectives on energy usage patterns and practices have been uncovered through a combination of granular energy consumption smart metering data and qualitative data from in-person interviews and surveys. The proposed framework can be adapted to different urban contexts, thus informing net-zero and net-positive buildings development. Findings support that an integrated approach, where insights from the end-user are incorporated into quantitative studies, can better inform effective energy policies and urban planning strategies.

Lastly, in Chapter 6, a NILM-enabled load scheduling framework for the energy-intensive dairy sector has been introduced, answering R.Q.5, through a co-created approach with agri-sector stakeholders. The approach was co-created with stakeholders and integrated high-resolution solar and carbon forecasting methods. Transfer learning of NILM models has been demonstrated to be an accurate, scalable, and cost-effective approach compared to the capital-intensive, and intrusive scenario of submetering installation. The framework considered renewable energy availability, time-of-use tariffs, and operational constraints that impose specific time-of-use of energy-intensive equipment. Minimum-intrusiveness approaches that rely on disaggregation of 1-sec data achieved similar savings compared to highly intrusive scenarios (savings of $\sim 36\%$ compared to $\sim 37\%$). The use of only 30-min smart metering data delivered comparable savings, demonstrating that a no-hardware installation approach that relies on NILM can achieve savings of up to $\sim 33\%$, i.e., a $\sim 10\%$ compared to the highly granular (1-sec) scenario.

In summary, the research presented in this thesis has demonstrated the value of downstream applications of NILM-enabled systems across both residential and industrial sectors, highlighting their capacity to improve energy efficiency, responsible consumption, and support sustainable prosumption practices. By bridging advanced data-driven techniques with user-centric insights and constraints, scalable and minimally

intrusive approaches have been demonstrated as an enabler of load scheduling, both for residential and industrial applications, without the requirement of additional hardware. The contributions presented across diverse use cases—from per-phase disaggregation in households and electric vehicle load analysis to mixed-method evaluations of net-positive homes and load scheduling in the energy-intensive dairy sector.

7.2   Limitations & future work

While advancements in load disaggregation, energy management, and sustainability across residential, agricultural, and smart housing contexts through co-creation has been demonstrated, there are still multiple unexplored sectors (especially in commercial and industrial settings) where NILM-enabled approaches could improve energy management, efficiency, and carbon savings at scale. Lack of domain knowledge was demonstrated to be a systematic blocker of optimising and tailoring NILM algorithms to handle different settings, mainly due to the variable and stochastic load profiles that equipment of different sectors have.

The reliance on qualitative data, the collection of which is a lengthy procedure, is a key limitation of the research presented in this thesis. Further to that, inaccuracies may be introduced due to potential errors in self-reported practices. Privacy concerns or unintentional omissions in surveys and interviews could lead to incomplete or imprecise data, impacting the accuracy of load disaggregation and, consequently, the overall methodology. While triangulation and cross-validation of qualitative and quantitative data helped mitigate these issues, scalability remains a challenge, as the methodology depends on qualitative data and the availability of specific datasets, such as smart meter readings, installed RES capacity, and time-of-use tariff information. Additionally, variations in data collection methods and processing, including different NILM approaches, solar models, or cost estimation techniques, may influence findings, highlighting the need for further comparative analyses to refine and validate the methodology across the different sectors.

Future efforts should aim to enhance the scalability and transferability of the proposed approaches by further evaluating the generalisability of algorithms across

different geographic regions, infrastructures, and housing or farm types, especially those with unique and non-standardised energy consumption patterns. Further to that, the computational efficiency of load scheduling optimisation approaches should be explored, through the introduction of heuristic and meta-heuristic especially for highly complex environments. The feasibility of using increasingly lower granularity data (15-minute intervals or higher) for disaggregation and energy management, particularly in privacy-conscious applications, should be explored. This could involve developing robust methods to compensate for the loss of data granularity while maintaining the utility and trust of the data mining methods. Future steps should investigate how NILM-enabled systems can dynamically adapt to real-time changes in renewable energy production and carbon footprint data by integrating predictive models for renewable energy generation and load demand to improve real-time load scheduling while taking into account end-user routines and day-to-day practices. Mixed-method approaches that combine quantitative data with qualitative insights can inform policies and regulations to ensure the design and operation of net-zero and net-positive developments achieve their intended outcomes. By addressing these areas, future research can refine and expand the impact of NILM-enabled energy management solutions, driving further progress toward energy efficiency, sustainability, and the broader adoption of smart/low-carbon energy technologies.

Appendix A

NorPEN: A Norwegian Positive Energy Neighbourhood dataset of electrical measurements and interviews on energy practices

A.1   Introduction

This appendix supplements Subsection 5.2.1, where the curation and release of the first energy dataset of a PEN, including consumption and production energy data, granular weather data, transcribed semi-structured interviews with the householders and time-of-use questionnaires, was presented. Through this dataset, the effects of hourly variable energy tariffs in end-users' consumption practices can be explored, with the dataset already being used to explore the effects of smart energy technologies and how these

affect household practices and demand shifting [93] as well as on a mixed-methods data-driven approach for energy-centric evaluation of net-positive households to answer the "what", "why" and "how" of energy prosumption in net-positive dwellings [92] (see Chapter 5). Validation and quality of the dataset, records, storage, insights and source code pertaining to NorPEN dataset are included in this appendix.

A.2   Validation & quality

In this dataset, energy data from 6 different households with a total of 34,852,816 samples collected from a total expected energy-related samples of 35,697,648, which corresponds to a missing data rate of 2.36%. A summary of the actual versus the expected timestamps during which data were collected is presented in Table A.1. Each timestamp corresponds to 10 readings for the metering data and 4 readings for the billing data. As expected, the hourly billing vector has statistically fewer missing values, partly due to the fact that billing data are collected every hour (compared to 10 seconds for the metering data) and partly due to the prioritisation of the collection of billing data to preserve utility and accurate billing. In addition to data availability,

Table A.1: Energy data availability as the number of available (actual) and expected samples. ©2024 IEEE

	Metering data		Billing data	
ID	**Actual/Expected**	**Ratio**	**Actual/Expected**	**Ratio**
1	505,708/509,400	99.28%	1,415/1,416	99.93%
3	505,073/509,400	99.15%	1,415/1,416	99.93%
5	486,485/509,400	95.50%	1,416/1,416	100%
6	505,835/509,400	99.30%	1,415/1,416	99.93%
7	490,564/509,400	96.30%	1,414/1,416	99.86%
9	499,672/509,400	98.09%	1,415/1,416	99.93%
Total	3,481,320/3,565,800	97.63%	9,904/9,912	99.92%

the quality of the collected data was explored by estimating the length of missing data gaps. Table A.2 includes the percentage of the data that which the maximum gap

interval does not exceed a period spanning 10 seconds to 6 hours. The majority of the gaps within the dataset are in the range of 10 seconds to 1 minute, with very few gaps having a duration greater than 1 minute per household. There has been no gap interval of more than 6 hours for any of the 6 households. Small gaps can be filled through an interpolation method, whereas longer gaps can be filled from average historical consumption data. All the collected and generated disaggregated streams were manually inspected (visual inspection) by an energy expert to assess the validity of the data. No erroneous spikes were identified in the collected energy readings. The expected accuracy of the disaggregated data is expected to be similar to the demonstrated accuracy of the used disaggregation algorithms, as already demonstrated in the literature [6, 64, 67, 155]. As electricity tariff data were obtained through the energy provider, the tariff data do not suffer from missing values, and all 1,415 hourly pricing readings are available. PV data were cross-validated through the prediction of worldwide energy resource (POWER) [158] portal with an average deviation of the solar irradiance data of less than 1%. The technical validation of qualitative data, such as interviews, is not as objective

Table A.2: Quality of energy data: length data gaps as % of samples with data gap lengths less than the given period. ©2024 IEEE

House ID	10-sec	1-min	30-min	1-hour	6-hour
1	99.18%	99.93%	99.93%	99.93%	100%
3	99.00%	99.88%	99.88%	99.88%	99.88%
5	95.32%	99.74%	99.84%	99.84%	99.84%
6	99.10%	99.83%	99.84%	99.84%	99.84%
7	96.13%	99.79%	99.84%	99.84%	99.84%
9	97.97%	99.90%	99.91%	99.91%	99.91%

as the quantitative data. The results that may be obtained from this data rely on the content analysis techniques that will be deployed, as well as the theoretical framework chosen by researchers. Nonetheless, since the interview guideline was created based on a theoretical framework of social practices, relevant connections between materials, skills, and meanings (i.e., the elements of practices according to [159]) that can be obtained from the interviews can be highlighted. The data reveal several meanings ascribed to

Appendix A. NorPEN dataset

RES, EVs, and SHTs, as well as different ways of engagement with such technologies for energy management. Competencies and skills to handle such technologies can also be found throughout the interviews. A few themes that can be potentially explored in the data are summarised below:

- **Materials:** PVs, ground source heat pumps, SHTs, EVs, smart apps in general.
- **Skills/Competencies:** Basic tech skills are needed to run the smart home. Some households enjoy acquiring tech skills through interaction with technology, while others prefer/need digital or in-person technical assistance. As a community, the relationship with neighbours in the process of acquiring knowledge on energy technologies was also uncovered. The smart system's complexity and load of information can exclude certain households, such as the elderly and others who do not have time to learn how to handle such devices and mobile apps. The systems can be complex even for households with previous knowledge of energy and IT.
- **Meanings/engagement:** Affordability, energy efficiency, and convenience are some of the meanings that may be found in the interviews as drivers of households' engagement with their smart homes, EVs, and neighbourhood. Exclusionary design, unmet expectations, technical issues, time demanded to set up/learn how to set up features and automation, as well as gender issues in handling smart technologies, can be found as some of the reasons for households' disengagement with energy demand.

A.3   Records & storage

Adhering to the FAIR principles [160], the recommended file formats by the UK data service for data sharing, reuse, and preservation [161], and the practices in NILM literature [162], the data are made available in the form of CSV and TXT files. There are 4 CSV files for each household, one containing the household total energy consumed data, one containing the disaggregated activities, one containing the solar production data and one containing the utility billing vector data. Figure A.1 represents the structure of the dataset. All timestamps are in Coordinated Universal Time (UTC) [YYYY-MM-DD

Appendix A. NorPEN dataset
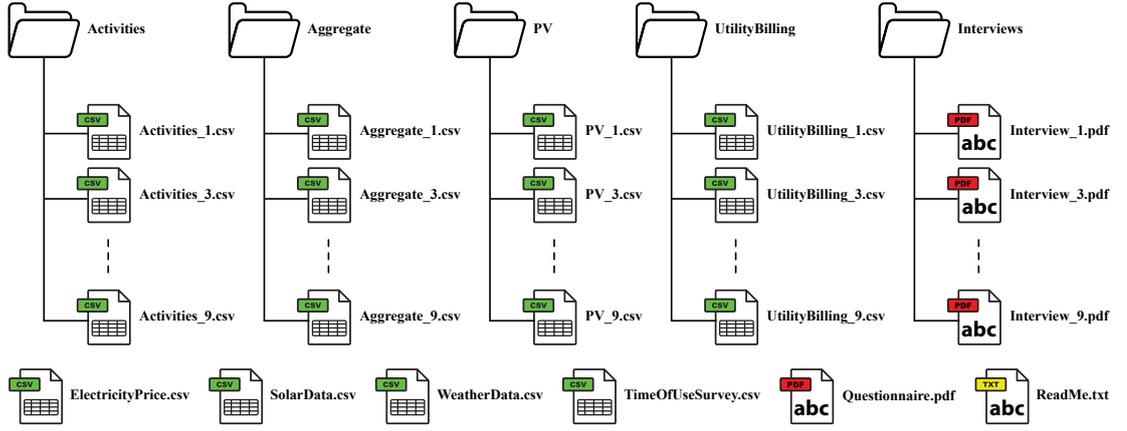
Figure A.1: Dataset structure. ©2024 IEEE

HH:mm:ss] format. UTC format was selected as it is the primary global standard to regulate time. The CSV files ("Aggregate_#.csv") containing the aggregate data are described in Table A.3. The CSV files ("Activities_#.csv") containing the disaggregated

Table A.3: Description of aggregate ("Aggregate_#.csv").

Feature	Type	Description
Timestamp	DateTime	The timestamp in UTC
ActivePowerPositive	Float	Positive aggregated active power [W]
ActivePowerNegative	Float	Negative aggregated active power [W]
ReactivePowerPositive	Float	Positive aggregated reactive power [VAR]
ReactivePowerNegative	Float	Negative aggregated reactive power [VAR]
PhaseOneCurrent	Float	Phase 1 current [A]
PhaseTwoCurrent	Float	Phase 2 current [A]
PhaseThreeCurrent	Float	Phase 3 current [A]
PhaseOneVoltage	Float	Phase 1 voltage [V]
PhaseTwoVoltage	Float	Phase 2 voltage [V]
PhaseThreeVoltage	Float	Phase 3 voltage [V]

activities data are described in Table A.4 The CSV files ("PV_#.csv") containing the solar production data is described in Table A.5. The CSV files ("UtilityBilling_#.csv") containing the utility billing vector energy data are described in Table A.6.

Further to these, a single CSV file containing the hourly billing vector, a single CSV

181

Table A.4: Description of activities ("Activities_#.csv").

Feature	Type	Description
Timestamp	DateTime	The timestamp in UTC
Heating	Float	Estimated heating power [W]
Cooking	Float	Estimated cooking power [W]
LaundryDishwashing	Float	Estimated laundry / dishwashing power [W]
EV	Float	Estimated EV charging power [W]

Table A.5: Description of solar data ("PV_#.csv").

Feature	Type	Description
Timestamp	DateTime	The timestamp in UTC
PV	Float	Estimated PV power production [W]
EV	Float	Estimated EV charging power [W]

Table A.6: Description of utility billing ("UtilityBilling_#.csv").

Feature	Type	Description
Timestamp	DateTime	The timestamp in UTC
CumulativeActiveImportEnergy	Float	Aggregated active energy imported from the grid from the date of the installation of the meter [Wh]
CumulativeActiveExportEnergy	Float	Aggregated active energy exported to the grid from the date of the installation of the meter [Wh]
CumulativeReactiveImportEnergy	Float	Aggregated reactive energy imported from the grid from the date of the installation of the meter [VARh]
CumulativeReactiveExportEnergy	Float	Aggregated reactive energy exported to the grid from the date of the installation of the meter [VARh]

file that includes the weather variables and a single CSV file containing the solar data
are provided.

The CSV file ("ElectricityPrice.csv") containing the hourly billing vector is described
in Table A.7.

Table A.7: Description of electricity price data ("ElectricityPrice.csv").

Feature	Type	Description
Timestamp	DateTime	The timestamp in UTC
Price	Float	Hourly market price [NOK] excluding VAT

The CSV file ("WeatherData.csv") containing the weather data is described in
Table A.8.

Table A.8: Description of weather data ("WeatherData.csv").

Feature	Type	Description
Timestamp	DateTime	The end timestamp in UTC
Temperature	Float	Air temperature in Celsius [°C]
RelativeHumidity	Float	Atmospheric relative humidity [%]
Dewpoint	Float	Dewpoint obtained from Eq. 5.5 [°C]
SurfacePressure	Float	Surface pressure [hPa]
Precipitation	Float	Rain precipitation [mm/h]
WindSpeed	Float	Wind speed at 10 meters from the ground [m/s]
WindDirection	Float	Wind direction at 10 meters from the ground [°]

The CSV file ("SolarData.csv") containing the solar data is described in Table A.9.

Further to the quantitative data, the qualitative data are organised in interviews
and questionnaire data as follows: (i) the semi-structured interviews are organised in
a single folder containing the 4 interviews in PDF format; and (ii) the questionnaire
template and the replies to the time of use survey are made available in PDF and CSV
file accordingly. Lastly, there is a single TXT read-me file that summarises the content
of the dataset.

Table A.9: Description of solar data ("SolarData.csv").

Feature	Type	Description
Timestamp	DateTime	The end timestamp in UTC
GHI	Float	Global solar irradiance, i.e., the total irradiance on a horizontal surface at ground level [W/m²]
BHI	Float	Direct solar irradiance, i.e., the beam irradiance on a horizontal surface at ground level [W/m²]
DHI	Float	Diffuse solar irradiance, i.e., the diffuse irradiance on a horizontal surface at ground level [W/m²]
BNI	Float	Direct solar irradiance on a mobile plane at normal incidence that follows the sun [W/m²]
Zenith	Float	Solar zenith angle [°]
Albedo	Float	Reflective coefficient on ground [%]
CloudCoverage	Float	Cloud coverage [%]
CloudType	Float	Cloud type, -1 = no value, 0 = no clouds, 5 = low-level cloud, 6 = medium-level cloud, 7 = high-level cloud, and 8 = thin cloud

A.4   Insights & notes

The dataset is made available in CSV format, which can be easily accessed by the majority of the scientific computing packages, including MATLAB, SPSS, R and Python.

A.5   Source code & scripts

The code was developed using MATLAB and Python 3.8 and deployed on a Windows machine. Code from public repositories that have been used in this dataset can be accessed at: `https://github.com/DLZRMR/seq2subseq` [64], `https://github.com/jiejiang-jojo/fast-seq2point` [155] and `https://github.com/renewables-ninja/gsee` [130].

Appendix B

FIELD: A comprehensive FarmIng Electrical LoaD measurements dataset from 30 three-phase dairy farms in Germany

B.1   Introduction

This appendix supplements Subsection 6.2.1.1, where the curation and release of the comprehensive electrical loads measurement FIELD dataset for a diverse range of typical energy-intensive activities including detailed labelling and load characteristics information that improves the understanding of the diverse dairy farming activities, is introduced and presented. Validation and quality of the dataset, records, storage, insights and source code pertaining tothe FIELD dataset are included in this appendix.

B.2 Validation & quality

A summary of the aggregate and submetered data availability across the 30 farms (see Subsection 6.2.1.1) comparing the actual and the expected number of samples is presented in Table B.1. Please note that at each time point, three active power readings, one from each of the three phases, were collected (both for aggregated and submetered samples).

Table B.2 highlights large gaps due to metering infrastructure disconnection or due to meter malfunctioning. The majority of the metering disconnections lasted from a few hours to a couple of days. Farm 28 contains a notable gap that spans over 4 months — which also affects the data availability (see Table B.1). Although only 20.68% of the timestamps were collected on this farm, the farm was not excluded from the dataset as it contained equipment that was not monitored in other farms. The identified gaps in Table B.2 were left unfilled with NaN values included in the dataset. These large gaps could be filled through interpolation of average historical consumption data, but this step should be decided based on the use of the specific dataset. All the collected data streams were manually inspected (visual inspection) by an energy expert to assess the validity of the data. Negative values that were produced due to rounding errors, due to the post-processing of the metered data with absolute values ranging from E-14 to E-15 Watts, were set to zero. No erroneous spikes were identified in the collected energy readings. Spikes observed during the starting of equipment (especially inductive motors, such as compressors and pumps) that are the result of the inrush currents were not removed and are included in the dataset. Farms with only submetered loads and no aggregate (i.e., farms with IDs 8, 13, 14, and 19) were manually inspected and compared with similar loads from the dataset. As farm 13 and 14 contain unique submetered equipment that is not available on other farms, these could not be cross-validated through other farms in the dataset. The industrial scale WM load signature of farm 14 was compared with similar residential WMs from publicly available datasets [46]. Although the industrial WM present in Farm 14 had a much higher power level, with the water heating element distributed across the three phases, the load pattern matched

Appendix B. FIELD dataset

Table B.1: Data availability.

	Aggregate samples			Submetering samples		
ID	**Expected**	**Available**	**[%]**	**Expected**	**Available**	**[%]**
1	92,173,260	92,062,074	99.88	368,693,040	368,248,296	99.88
2	38,142,882	35,641,851	93.44	152,571,528	142,567,404	93.44
3	103,161,603	102,856,554	99.70	309,484,809	308,569,662	99.70
4	102,025,005	101,562,177	99.55	408,100,020	406,248,708	99.55
5	103,161,603	91,050,837	88.26	412,646,412	364,203,348	88.26
6	102,496,083	102,496,083	100.00	102,496,083	102,496,083	100.00
7	102,503,064	102,503,064	100.00	102,503,064	102,503,064	100.00
8	-	-	-	102,503,010	102,503,010	100.00
9	52,680,501	50,179,728	95.25	210,722,004	200,718,912	95.25
10	93,436,665	93,436,665	100.00	280,309,995	280,309,995	100.00
11	47,402,796	47,402,796	100.00	189,611,184	189,611,184	100.00
12	49,285,647	49,205,520	99.84	147,856,941	147,616,560	99.84
13	-	-	-	92,977,263	92,977,263	100.00
14	-	-	-	92,986,602	92,986,602	100.00
16	92,984,934	92,984,934	100.00	185,969,868	185,969,868	100.00
17	92,169,429	92,169,429	100.00	184,338,858	184,338,858	100.00
18	92,171,640	92,171,640	100.00	368,686,560	368,686,560	100.00
19	-	-	-	92,171,409	92,171,409	100.00
20	83,607,951	81,692,028	97.71	83,607,951	81,692,028	97.71
21	75,410,346	55,095,387	73.06	301,641,384	220,381,548	73.06
22	52,679,829	50,179,062	95.25	52,679,829	50,179,062	95.25
23	83,603,913	81,667,818	97.68	334,415,652	326,671,272	97.68
24	92,984,421	92,984,421	100.00	371,937,684	371,937,684	100.00
25	65,539,086	65,539,086	100.00	65,539,086	65,539,086	100.00
26	98,616,387	98,616,387	100.00	98,616,387	98,616,387	100.00
27	100,682,016	100,682,016	100.00	302,046,048	302,046,048	100.00
28	48,080,736	9,942,039	20.68	96,161,472	19,884,078	20.68
29	98,619,216	98,619,216	100.00	197,238,432	197,238,432	100.00
30	98,616,084	78,697,698	79.80	98,616,084	78,697,698	79.80
31	98,616,006	98,616,006	100.00	98,616,006	98,616,006	100.00
Total	**2,160,851,103**	**2,058,054,516**	**95.24**	**5,905,744,665**	**5,644,226,115**	**95.57**

the residential ones. Lastly, the sum of the submetered loads was compared with the aggregate load on a per-phase basis to ensure that the total submetered power is less than or equal to the aggregate power at each sampling point.

Table B.2: Data gaps.

ID	Start	Stop	Duration
1	21/01/2021 04:00:30	21/01/2021 13:02:50	9h 2m 20s
2	13/06/2020 16:32:25	23/06/2020 08:05:55	9d 15h 33m 30s
3	17/02/2020 07:37:51	18/02/2020 11:52:33	1d 4h 14m 42s
4	03/09/2020 14:55:36	05/09/2020 09:44:04	1d 18h 48m 28s
5	16/12/2020 16:10:04	01/02/2021 08:34:39	1mo 15d 16h 24m 35s
9	13/06/2020 16:32:24	13/06/2020 08:05:34	9d 15h 33m 10s
12	03/07/2020 20:01:06	04/07/2020 03:26:14	7h 25m 8s
20	13/06/2020 16:58:55	17/06/2020 05:11:47	3d 12h 12m 52s
	31/12/2020 10:50:40	04/01/2021 08:01:47	3d 21h 11m 7s
21	13/06/2020 16:58:54	17/06/2020 05:11:48	3d 12h 12m 54s
	09/07/2020 12:54:50	18/09/2020 12:31:35	2mo 8d 23h 36m 45s
	31/12/2020 10:50:41	04/01/2021 08:01:52	3d 21h 11m 11s
22	13/06/2020 16:32:24	23/06/2020 08:05:32	9d 15h 33m 8s
23	13/06/2020 16:58:55	17/06/2020 05:11:53	3d 12h 12m 58s
	31/12/2020 10:50:41	04/01/2021 08:01:54	3d 21h 11m 13s
27	20/07/2020 19:29:43	17/09/2020 12:02:35	1mo 27d 32m 52s
28	20/02/2020 15:39:38	22/06/2020 18:54:12	4mo 2d 3h 14m 34s
	02/07/2020 09:13:54	03/07/2020 14:17:21	1d 5h 3m 27s
	20/07/2020 19:29:44	12/08/2020 14:33:19	22d 19h 3m 35s
30	29/07/2020 10:37:04	31/07/2020 07:44:22	1d 21h 7m 18s
	06/08/2020 11:52:26	20/10/2020 11:02:48	2mo 13d 23h 10m 22s

B.3   Records & storage

In line with the FAIR principles [160], the UK Data Service's recommended file formats for data sharing, reuse, and preservation [161], and established practices in NILM literature [162], the data are provided as comma-separated values (CSV) and text (TXT) files. For each farm where aggregated data are available (see Table 5.1), there is a

Appendix B. FIELD dataset
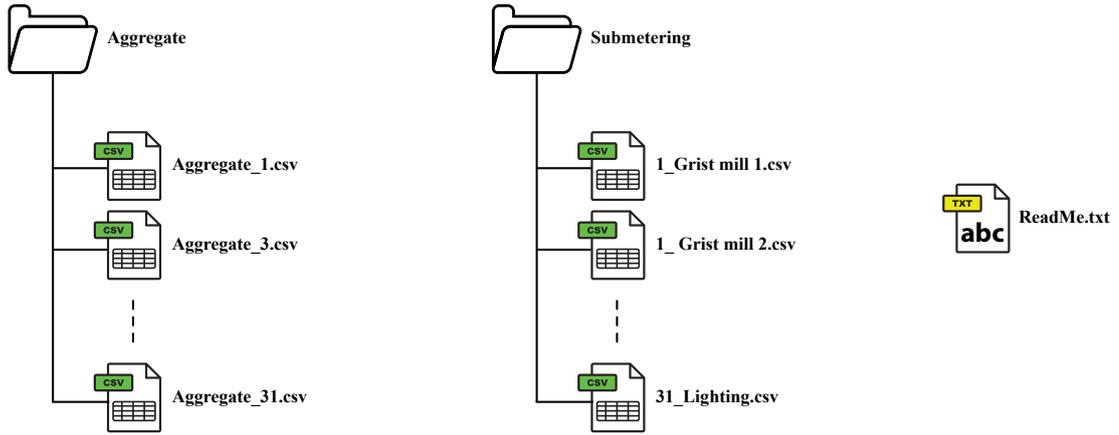
Figure B.1: Dataset structure.

single CSV file associated with the farm that contains the aggregated data readings. For each farm for which submetered data are available (see Table 5.1), there is a CSV file associated per submetered dairy equipment that contains the submetered readings. Figure B.1 represents the structure of the dataset. The columns of the CSV files ("Aggregate_*id*.csv") containing the aggregate data of each farm "*id*" are described in Table B.3.

Table B.3: Description of aggregate (`Aggregate_#.csv`)

Feature	Type	Description
Time	DateTime	The timestamp of the collected data in UTC [YYYY-MM-DD HH:mm:ss]
phase1	Float	The aggregated active power level on phase 1 in Watts [W]
phase2	Float	The aggregated active power level on phase 2 in Watts [W]
phase3	Float	The aggregated active power level on phase 3 in Watts [W]

The columns of the CSV files ("*id*_equipment.csv") containing the submetered equipment data of each farm "*id*" are described in Table B.4.

A single ReadMe TXT file is included to provide additional information about the structure of the dataset. The format of the ReadMe file is the following:

- Dataset description
- Licensing

189

Appendix B. FIELD dataset

Table B.4: Description of submetering ("#_equipment.csv")

Feature	Type	Description
Time	DateTime	The timestamp of the collected data in UTC [YYYY-MM-DD HH:mm:ss]
phase1	Float	The submetered active power level on phase 1 in Watts [W]
phase2	Float	The submetered active power level on phase 2 in Watts [W]
phase3	Float	The submetered active power level on phase 3 in Watts [W]

- Naming conventions
- File formats
- Farming equipment per site

The post-processed dataset is provided through the University of Strathclyde's PURE data repository and can be accessed at: `https://doi.org/10.15129/1211ae7c-9b70-4a39-b3ce-318d81583749`.

B.4 Insights & notes

The dataset is made available in CSV format, which can be easily accessed by the majority of the scientific computing packages, including MATLAB, SPSS, R and Python. The included README file explains the structure of each individual CSV file, including its contents and any known issues.

B.5 Source code & notes

The code to prepare, identify gaps, visualise and store the data was developed using MATLAB R2024a and deployed on a Windows machine. The algorithm used to post-process the data is presented in Section 6.2.1.1.

Bibliography

[1] S. Makonin and F. Popowich, "Nonintrusive load monitoring (NILM) performance evaluation," *Energy Efficiency*, vol. 8, no. 4, pp. 809–814, 2015.

[2] P. A. Schirmer and I. Mporas, "Non-Intrusive Load Monitoring: A Review," *IEEE Transactions on Smart Grid*, vol. 14, no. 1, pp. 769–784, 2023.

[3] G.-F. Angelis, C. Timplalexis, S. Krinidis, D. Ioannidis, and D. Tzovaras, "NILM applications: Literature review of learning approaches, recent developments and challenges," *Energy and Buildings*, vol. 261, p. 111951, 2022.

[4] C. Beckel, W. Kleiminger, R. Cicchetti, T. Staake, and S. Santini, "The ECO Data Set and the Performance of Non-Intrusive Load Monitoring Algorithms," in *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, ser. BuildSys '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 80–89.

[5] Department of Energy & Climate Change. (2014) Smart Metering Equipment Technical Specifications: Second Version. [Online]. Available: https://www.gov.uk/government/consultations/smart-metering-equipment-technical-specifications-second-version

[6] A. Vavouris, L. Stankovic, V. Stankovic, and J. Shi, "Benefits of Three-Phase Metering for Load Disaggregation," in *Proceedings of the 9th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*, ser. BuildSys '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 393–397.

Bibliography

[7] J. Brozovsky, A. Gustavsen, and N. Gaitani, "Zero emission neighbourhoods and positive energy districts–A state-of-the-art review," *Sustainable Cities and Society*, vol. 72, p. 103013, 2021.

[8] T. Harputlugil and P. de Wilde, "The interaction between humans and buildings for energy efficiency: A critical review," *Energy Research & Social Science*, vol. 71, p. 101828, 2021.

[9] P. Shine, J. Upton, P. Sefeedpari, and M. D. Murphy, "Energy Consumption on Dairy Farms: A Review of Monitoring, Prediction Modelling, and Analyses," *Energies*, vol. 13, no. 5, p. 1288, 2020. [Online]. Available: https://www.mdpi.com/1996-1073/13/5/1288

[10] M. Kaselimi, E. Protopapadakis, A. Voulodimos, N. Doulamis, and A. Doulamis, "Towards trustworthy energy disaggregation: A review of challenges, methods, and perspectives for non-intrusive load monitoring," *Sensors*, vol. 22, no. 15, p. 5872, 2022.

[11] Tamara, L. Sobot, V. Stankovic, and Stankovic, "Electricity consumption measurements from three dairy farms in Germany," 2022. [Online]. Available: https://dx.doi.org/10.15129/cc3d1dcd-0ffb-4c38-aab7-b47712938d13

[12] Hugo Pro, "AI to help cut energy usage," 2024. [Online]. Available: https://blog.hugoenergypro.co.uk/ai-to-help-cut-energy-usage/

[13] The Association For Renewable Energy and Clean Technology. (2018) The feasibility, costs and benefits of three phase power supplies in new homes. [Online]. Available: https://www.r-e-a.net/new-homes-shouldnt-be-held-back-by-pre-wwii-electrical-standards/

[14] J. F. Martins, A. G. Pronto, V. Delgado-Gomes, and M. Sanduleac, "Smart meters and advanced metering infrastructure," in *Pathways to a smarter power system*. Elsevier, 2019, pp. 89–114.

Bibliography

[15] B. K. Sovacool, A. Hook, S. Sareen, and F. W. Geels, "Global sustainability, innovation and governance dynamics of national smart electricity meter transitions," *Global Environmental Change*, vol. 68, p. 102272, 2021.

[16] B. K. Sovacool, P. Kivimaa, S. Hielscher, and K. Jenkins, "Vulnerability and resistance in the United Kingdom's smart meter transition," *Energy Policy*, vol. 109, pp. 767–781, 2017.

[17] T. Knayer and N. Kryvinska, "An analysis of smart meter technologies for efficient energy management in households and organizations," *Energy Reports*, vol. 8, pp. 4022–4040, 2022.

[18] Å. L. Sørensen, K. B. Lindberg, I. Sartori, and I. Andresen, "Analysis of residential EV energy flexibility potential based on real-world charging reports and smart meter data," *Energy and Buildings*, vol. 241, p. 110923, 2021.

[19] A. Weigert, K. Hopf, S. A. Günther, and T. Staake, "Heat pump inspections result in large energy savings when a pre-selection of households is performed: A promising use case of smart meter data," *Energy Policy*, vol. 169, p. 113156, 2022.

[20] Comité Européen de Normalisation Électrotechnique. (2022) Smart Grids and Meters. [Online]. Available: https://www.cencenelec.eu/areas-of-work/cen-cenelec-topics/smart-grids-and-meters/smart-meters/

[21] European Commission and Directorate-General for Energy, C. Alaton, and F. Tounquet, *Benchmarking smart metering deployment in the EU-28 – Final report*. Publications Office, 2020.

[22] G. W. Hart, "Nonintrusive appliance load monitoring," *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1870–1891, 1992.

[23] P. Huber, A. Calatroni, A. Rumsch, and A. Paice, "Review on Deep Neural Networks Applied to Low-Frequency NILM," *Energies*, vol. 14, no. 9, 2021.

[24] L. Pereira and N. Nunes, "Performance evaluation in non-intrusive load monitoring: Datasets, metrics, and tools—A review," *Wiley Interdisciplinary Reviews: data mining and knowledge discovery*, vol. 8, no. 6, p. e1265, 2018.

[25] H. Rafiq, P. Manandhar, E. Rodriguez-Ubinas, O. A. Qureshi, and T. Palpanas, "A review of current methods and challenges of advanced deep learning-based non-intrusive load monitoring (NILM) in residential context," *Energy and Buildings*, p. 113890, 2024.

[26] G. Tanoni, E. Principi, and S. Squartini, "Non-Intrusive Load Monitoring in industrial settings: A systematic review," *Renewable and Sustainable Energy Reviews*, vol. 202, p. 114703, 2024.

[27] A. Yaniv and Y. Beck, "Advances in non-intrusive load monitoring for the industrial domain: Challenges, insights, and path forward," *Renewable and Sustainable Energy Reviews*, vol. 210, p. 115136, 2025.

[28] B. Zhao, K. He, L. Stankovic, and V. Stankovic, "Improving event-based non-intrusive load monitoring using graph signal processing," *IEEE Access*, vol. 6, pp. 53 944–53 959, 2018.

[29] K. He, L. Stankovic, J. Liao, and V. Stankovic, "Non-intrusive load disaggregation using graph signal processing," *IEEE Transactions on Smart Grid*, vol. 9, no. 3, pp. 1739–1747, 2016.

[30] S. Wang, L. Du, J. Ye, and D. Zhao, "A deep generative model for non-intrusive identification of EV charging profiles," *IEEE Transactions on Smart Grid*, vol. 11, no. 6, pp. 4916–4927, 2020.

[31] F. D. Garcia, W. A. Souza, I. S. Diniz, and F. P. Marafão, "NILM-based approach for energy efficiency assessment of household appliances," *Energy Informatics*, vol. 3, pp. 1–21, 2020.

[32] H. Rashid, P. Singh, V. Stankovic, and L. Stankovic, "Can non-intrusive load monitoring be used for identifying an appliance's anomalous behaviour?" *Applied energy*, vol. 238, pp. 796–805, 2019.

[33] H. Rashid, V. Stankovic, L. Stankovic, and P. Singh, "Evaluation of non-intrusive load monitoring algorithms for appliance-level anomaly detection," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, pp. 8325–8329.

[34] A. Ruano, A. Hernandez, J. Ureña, M. Ruano, and J. Garcia, "NILM techniques for intelligent home energy management and ambient assisted living: A review," *Energies*, vol. 12, no. 11, p. 2203, 2019.

[35] B. Zhao, M. Ye, L. Stankovic, and V. Stankovic, "Non-intrusive load disaggregation solutions for very low-rate smart meter data," *Applied energy*, vol. 268, p. 114949, 2020.

[36] K. Basu, V. Debusschere, S. Bacha, U. Maulik, and S. Bondyopadhyay, "Nonintrusive load monitoring: A temporal multilabel classification approach," *IEEE Transactions on industrial informatics*, vol. 11, no. 1, pp. 262–270, 2014.

[37] J. Liao, G. Elafoudi, L. Stankovic, and V. Stankovic, "Non-intrusive appliance load monitoring using low-resolution smart meter data," in *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, 2014, pp. 535–540.

[38] A. U. Rehman, T. T. Lie, B. Vallès, and S. R. Tito, "Low complexity non-intrusive load disaggregation of air conditioning unit and electric vehicle charging," in *2019 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia)*. IEEE, 2019, pp. 2607–2612.

[39] O. Parson, S. Ghosh, M. Weal, and A. Rogers, "Non-intrusive load monitoring using prior models of general appliance types," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 26, no. 1, 2012, pp. 356–362.

[40] S. Makonin, F. Popowich, I. V. Bajić, B. Gill, and L. Bartram, "Exploiting HMM Sparsity to Perform Online Real-Time Nonintrusive Load Monitoring," *IEEE Transactions on Smart Grid*, vol. 7, no. 6, pp. 2575–2585, 2016.

[41] D. Murray, L. Stankovic, V. Stankovic, S. Lulic, and S. Sladojevic, "Transferability of neural network approaches for low-rate energy disaggregation," in *ICASSP 2019-2019 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. IEEE, 2019, pp. 8330–8334.

[42] S. Yi, X. Yin, Y. Diao, B. Wang, and P. Wu, "A New Event-detection Method Based on Composite Windows in NILM for Industrial Settings," in *2019 IEEE Sustainable Power and Energy Conference (iSPEC)*, 2019, pp. 2768–2771.

[43] N. Batra, M. Gulati, A. Singh, and M. B. Srivastava, "It's Different: Insights into Home Energy Consumption in India," in *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, ser. BuildSys'13. New York, NY, USA: Association for Computing Machinery, 2013, pp. 1—-8.

[44] T. Kriechbaumer and H.-A. Jacobsen, "BLOND, a building-level office environment dataset of typical electrical appliances," *Scientific data*, vol. 5, no. 1, pp. 1–14, 2018.

[45] J. Kelly and W. Knottenbelt, "The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes," *Scientific data*, vol. 2, no. 1, pp. 1–14, 2015.

[46] D. Murray and L. Stankovic, "REFIT: Electrical load measurements (cleaned)," 2016. [Online]. Available: https://doi.org/10.15129/31da3ece-f902-4e95-a093-e0a 9536983c4

[47] J. Z. Kolter and M. J. Johnson, "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. Citeseer, 2011, pp. 59–62.

[48] T. Jasiński, "Modelling the disaggregated demand for electricity in residential buildings using artificial neural networks (deep learning approach)," *Energies*, vol. 13, no. 5, p. 1263, 2020.

[49] H. Rafiq, H. Zhang, H. Li, and M. K. Ochani, "Regularized LSTM Based Deep Learning Model: First Step towards Real-Time Non-Intrusive Load Monitoring," in *2018 IEEE International Conference on Smart Energy Grid Engineering (SEGE)*, 2018, pp. 234–239.

[50] H. Rafiq, X. Shi, H. Zhang, H. Li, and M. K. Ochani, "A deep recurrent neural network for non-intrusive load monitoring based on multi-feature input space and post-processing," *Energies*, vol. 13, no. 9, p. 2195, 2020.

[51] Z. Zhang, J. H. Son, Y. Li, M. Trayer, Z. Pi, D. Y. Hwang, and J. K. Moon, "Training-free non-intrusive load monitoring of electric vehicle charging with low sampling rate," in *Iecon 2014-40th annual conference of the ieee industrial electronics society*. IEEE, 2014, pp. 5419–5425.

[52] Pecan Street Inc., "Dataport," 2022. [Online]. Available: https://www.pecanstreet.org/dataport/

[53] A. A. Munshi and Y. A.-R. I. Mohamed, "Unsupervised nonintrusive extraction of electrical vehicle charging load patterns," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 1, pp. 266–279, 2018.

[54] A. F. M. Jaramillo, D. M. Laverty, J. M. del Rincón, J. Hastings, and D. J. Morrow, "Supervised non-intrusive load monitoring algorithm for electric vehicle identification," in *2020 IEEE international instrumentation and measurement Technology conference (I2MTC)*. IEEE, 2020, pp. 1–6.

[55] A. F. Moreno Jaramillo, J. Lopez-Lorente, D. M. Laverty, J. Martinez-del Rincon, D. J. Morrow, and A. M. Foley, "Effective identification of distributed energy resources using smart meter net-demand data," *IET Smart Grid*, vol. 5, no. 2, pp. 120–135, 2022.

[56] D. Murray, L. Stankovic, and V. Stankovic, "An electrical load measurements dataset of United Kingdom households from a two-year longitudinal study," *Scientific Data*, vol. 4, no. 1, p. 160122, 2017.

[57] Y. Amara-Ouali, Y. Goude, P. Massart, J.-M. Poggi, and H. Yan, "A review of electric vehicle load open data and models," *Energies*, vol. 14, no. 8, p. 2233, 2021.

[58] International Energy Agency. (2024) EV Outlook 2024. [Online]. Available: https://www.iea.org/reports/global-ev-outlook-2024

[59] UK Government. (2008) Climate Change Act 2008. [Online]. Available: https://www.legislation.gov.uk/ukpga/2008/27/contents

[60] International Energy Agency. (2021) Heat Pumps Analysis. [Online]. Available: https://www.iea.org/reports/heat-pumps

[61] Electric Nation. (2019) Smart Charged. [Online]. Available: https://www.electric nation.org.uk/wp-content/uploads/2019/07/Smart-Charged-Presentations.pdf

[62] Department for Business, Energy & Industrial Strategy. (2018) Clean Growth Strategy: executive summary. [Online]. Available: https://www.gov.uk/governm ent/publications/clean-growth-strategy/clean-growth-strategy-executive-sum mary

[63] Operations Directorate of Energy Networks Association. (2012) Engineering Recommendation G83 Issue 2. [Online]. Available: https://www.ofgem.gov.uk/sit es/default/files/docs/2012/08/er-g83-2-_v5--the-master-09-07-12-inc-ofgem-c omments---clean-version_0.pdf

[64] Y. Pan, K. Liu, Z. Shen, X. Cai, and Z. Jia, "Sequence-To-Subsequence Learning With Conditional Gan For Power Disaggregation," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 3202–3206.

[65] GitHub-DLZRMR/seq2subseq. (2019) Seq2subseq Method for NILM. [Online]. Available: https://github.com/DLZRMR/seq2subseq

[66] D. R. Chavan and D. S. More, "A Systematic Review on Low-Resolution NILM: Datasets, Algorithms, and Challenges," *Electronic Systems and Intelligent Computing*, pp. 101–120, 2022.

[67] A. Vavouris, B. Garside, L. Stankovic, and V. Stankovic, "Low-Frequency Non-Intrusive Load Monitoring of Electric Vehicles in Houses with Solar Generation: Generalisability and Transferability," *Energies*, vol. 15, no. 6, 2022.

[68] C. Zhang, M. Zhong, Z. Wang, N. Goddard, and C. Sutton, "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.

[69] M. D'Incecco, S. Squartini, and M. Zhong, "Transfer learning for non-intrusive load monitoring," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1419–1429, 2019.

[70] United Nations, "UN Climate Change Conference (COP26) at the SEC—Glasgow," 2021. [Online]. Available: https://ukcop26.org

[71] Department for Transport, "Transport and Environment Statistics: Autumn 2021," 2021. [Online]. Available: https://www.gov.uk/government/statistics/transport-and-environment-statistics-autumn-2021/transport-and-environment-statistics-autumn-2021

[72] Society of Motor Manufactures and Traders, "Electric Vehicle and Alternatively Fuelled Vehicle Registrations," 2025. [Online]. Available: https://www.smmt.co.uk/vehicle-data/evs-and-afvs-registrations/

[73] Department for Transport, "Electric vehicle charging infrastructure statistics," Department for Transport, Tech. Rep., 2025. [Online]. Available: https://www.gov.uk/government/collections/electric-vehicle-charging-infrastructure-statistics

[74] Transport & Environment, "UPDATE - T&E's analysis of electric car lifecycle CO2 emissions," 2022. [Online]. Available: https://www.transportenvironment.org/wp-content/uploads/2022/05/2022_05_TE_LCA_update-1.pdf.

[75] B. Marmiroli, M. Messagie, G. Dotelli, and J. Van Mierlo, "Electricity generation in LCA of electric vehicles: A review," *applied sciences*, vol. 8, no. 8, p. 1384, 2018.

[76] S. Verma, G. Dwivedi, and P. Verma, "Life cycle assessment of electric vehicles in comparison to combustion engine vehicles: A review," *Materials Today: Proceedings*, vol. 49, pp. 217–222, 2022.

[77] T. Alquthami, A. Alsubaie, M. Alkhraijah, K. Alqahtani, S. Alshahrani, and M. Anwar, "Investigating the impact of electric vehicles demand on the distribution network," *Energies*, vol. 15, no. 3, p. 1180, 2022.

[78] R.-C. Leou, C.-L. Su, and C.-N. Lu, "Stochastic analyses of electric vehicle charging impacts on distribution network," *IEEE Transactions on Power Systems*, vol. 29, no. 3, pp. 1055–1063, 2013.

[79] J. Yang, X. Long, X. Pan, F. Wu, X. Zhan, and Y. Lin, "Electric vehicle charging load forecasting model considering road network-power grid information," in *2019 International Conference on Technologies and Policies in Electric Power & Energy*. IEEE, 2019, pp. 1–5.

[80] M. Afzalan and F. Jazizadeh, "A machine learning framework to infer time-of-use of flexible loads: Resident behavior learning for demand response," *IEEE Access*, vol. 8, pp. 111 718–111 730, 2020.

[81] F. Khosrojerdi, S. Taheri, H. Taheri, and E. Pouresmaeil, "Integration of electric vehicles into a smart power grid: A technical review," in *2016 IEEE electrical power and energy conference (EPEC)*. IEEE, 2016, pp. 1–6.

[82] International Organization for Standardization (ISO), "ISO 14040: 2006 Environmental Management–Life Cycle Assessment–Principles and Framework (2)," 2009.

[83] European Commission, "European Platform on LCA," 2023. [Online]. Available: https://eplca.jrc.ec.europa.eu/index.html#menu1

[84] X. Xia and P. Li, "A review of the life cycle assessment of electric vehicles: Considering the influence of batteries," *Science of the Total Environment*, p. 152870, 2022.

[85] H. C. Kim, T. J. Wallington, R. Arsenault, C. Bae, S. Ahn, and J. Lee, "Cradle-to-gate emissions from a commercial electric vehicle Li-ion battery: a comparative analysis," *Environmental science & technology*, vol. 50, no. 14, pp. 7715–7722, 2016.

[86] D. Burchart-Korol, S. Jursova, P. Folega, J. Korol, P. Pustejovska, and A. Blaut, "Environmental life cycle assessment of electric vehicles in Poland and the Czech Republic," *Journal of Cleaner Production*, vol. 202, pp. 476–487, 2018.

[87] P. Girardi, A. Gargiulo, and P. C. Brambilla, "A comparative LCA of an electric vehicle and an internal combustion engine vehicle using the appropriate power mix: the Italian case study," *The International Journal of Life Cycle Assessment*, vol. 20, pp. 1127–1142, 2015.

[88] Y. Bicer and I. Dincer, "Life cycle environmental impact assessments and comparisons of alternative fuels for clean vehicles," *Resources, Conservation and Recycling*, vol. 132, pp. 141–157, 2018.

[89] C. Klemenjak, A. Faustine, S. Makonin, and W. Elmenreich, "On metrics to assess the transferability of machine learning models in non-intrusive load monitoring," *arXiv preprint arXiv:1912.06200*, 2019.

[90] Department for Business, Energy and Industrial Strategy, "Electric Vehicle Smart Chargepoint Survey," Department for Business, Energy and Industrial Strategy, Tech. Rep., 2022. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1129104/electric-vehicle-smart-charging-survey-2022.pdf

[91] Department for Transport, "Electric Vehicle Charging Research," Department for Transport, Tech. Rep., 2022. [Online]. Available: https://assets.publishing.service.

Bibliography

gov.uk/government/uploads/system/uploads/attachment_data/file/1078871/df
t-ev-driver-survey-summary-report.pdf

[92] A. Vavouris, F. Guasselli, L. Stankovic, V. Stankovic, K. Gram-Hanssen, and
S. Didierjean, "A complex mixed-methods data-driven energy-centric evaluation
of net-positive households," *Applied Energy*, vol. 367, p. 123404, 2024.

[93] F. Guasselli, A. Vavouris, L. Stankovic, V. Stankovic, S. Didierjean, and K. Gram-
Hanssen, "Smart energy technologies for the collective: Time-shifting, demand
reduction and household practices in a Positive Energy Neighbourhood in Norway,"
Energy Research & Social Science, vol. 110, p. 103436, 2024.

[94] A. Vavouris, L. Stankovic, and V. Stankovic, "Smart meter electricity of a
Household in Germany with Electric Vehicle Charging Annotation," 2016. [Online].
Available: https://doi.org/10.15129/c41a6a02-5df5-4ed7-b8e6-6488895d43f7

[95] Department for Transport, "Renewable fuel statistics 2022: Third provisional
report," Department for Transport, Tech. Rep., 2022. [Online]. Available:
https://www.gov.uk/government/statistics/renewable-fuel-statistics-2022-third
-provisional-report/renewable-fuel-statistics-2022-third-provisional-report

[96] National Grid ESO, "Carbon Intensity API," 2023. [Online]. Available:
https://carbonintensity.org.uk

[97] European Network of Transmission System Operators for Electricity,
"ENTSOE Transparency Platform," 2023. [Online]. Available: https:
//transparency.entsoe.eu/dashboard/show

[98] S. Schlömer, T. Bruckner, L. Fulton, E. Hertwich, A. McKinnon, D. Perczyk,
J. Roy, R. Schaeffer, R. Sims, P. Smith *et al.*, "Annex III: Technology-specific
cost and performance parameters," in *Climate change 2014: Mitigation of climate
change: Contribution of working group III to the fifth assessment report of the
Intergovernmental Panel on Climate Change.* Cambridge University Press, 2014,
pp. 1329–1356.

Bibliography

[99] World Economic Forum, "Here's how we can build net-positive homes," 2023. [Online]. Available: https://www.weforum.org/agenda/2023/01/davos23-home-energy-savings-net-positive/

[100] International Energy Agency, "Electricity Market Report 2023," 2023. [Online]. Available: https://www.iea.org/reports/electricity-market-report-2023

[101] Directorate-General for Justice and Consumers, "Consumer Conditions Scoreboard 2023 Edition," European Commission, B-1039 Brussels, Report, 2023. [Online]. Available: https://commission.europa.eu/system/files/2023-10/consumer_conditions_scoreboard_2023_v1.1.pdf

[102] R. Debnath, R. Bardhan, D. U. Shah, K. Mohaddes, M. H. Ramage, R. M. Alvarez, and B. K. Sovacool, "Social media enables people-centric climate action in the hard-to-decarbonise building sector," *Scientific Reports*, vol. 12, no. 1, p. 19017, 2022.

[103] C. Wang, J. Song, D. Shi, J. L. Reyna, H. Horsey, S. Feron, Y. Zhou, Z. Ouyang, Y. Li, and R. B. Jackson, "Impacts of climate change, population growth, and power sector decarbonization on urban building energy use," *Nature Communications*, vol. 14, no. 1, p. 6434, 2023.

[104] Historic England, "Heritage and the environment 2020," 2020. [Online]. Available: https://historicengland.org.uk/content/heritage-counts/pub/2020/heritage-environment-2020/

[105] ISO Central Secretary, "Energy performance of buildings — Methods for expressing energy performance and for energy certification of buildings," International Organization for Standardization, Geneva, CH, Standard ISO 16343:2013, 2013. [Online]. Available: https://www.iso.org/standard/56224.html

[106] K. Gram-Hanssen and S. Georg, "Energy performance gaps: promises, people, practices," *Building Research & Information*, vol. 46, no. 1, pp. 1–9, 2018.

Bibliography

[107] British Broadcasting Corporation (BBC), "UK weather: Ratcliffe-on-Soar power station readied to boost supply," 2023. [Online]. Available: https://www.bbc.co.uk/news/uk-england-nottinghamshire-65879949

[108] British Broadcasting Corporation (BBC), "Emergency coal power plants used for first time as UK sees cold snap," 2023. [Online]. Available: https://www.bbc.co.uk/news/business-64879044

[109] Hannah Fearn, "It may be hot, but most British homes don't need aircon. Switch it off," 2023. [Online]. Available: https://www.theguardian.com/commentisfree/2023/jun/14/british-homes-aircon-coal-fired-power-station-extravagance

[110] U. Ali, M. H. Shamsi, C. Hoare, E. Mangina, and J. O'Donnell, "Review of urban building energy modeling (UBEM) approaches, methods and tools using qualitative and quantitative analysis," *Energy and Buildings*, vol. 246, p. 111073, 2021.

[111] K. Amasyali and N. M. El-Gohary, "A review of data-driven building energy consumption prediction studies," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 1192–1205, 2018.

[112] M. V. Bavaresco, S. D'Oca, E. Ghisi, and R. Lamberts, "Methods used in social sciences that suit energy research: A literature review on qualitative methods to assess the human dimension of energy use in buildings," *Energy and Buildings*, vol. 209, p. 109702, 2020.

[113] B. K. Sovacool, J. Axsen, and S. Sorrell, "Promoting novelty, rigor, and style in energy social science: Towards codes of practice for appropriate methods and research design," *Energy Research & Social Science*, vol. 45, pp. 12–42, 2018, special Issue on the Problems of Methods in Climate and Energy Research.

[114] P. X. Zou, X. Xu, J. Sanjayan, and J. Wang, "A mixed methods design for building occupants' energy behavior research," *Energy and Buildings*, vol. 166, pp. 239–249, 2018.

Bibliography

[115] J. W. Creswell and J. D. Creswell, *Research design: Qualitative, quantitative, and mixed methods approaches.* Sage publications, 2017.

[116] A. P. Neto-Bradley, R. Rangarajan, R. Choudhary, and A. B. Bazaz, "Energy transition pathways amongst low-income urban households: A mixed method clustering approach," *MethodsX*, vol. 8, p. 101491, 2021.

[117] O. Guerra-Santin, N. Romero Herrera, E. Cuerda, and D. Keyson, "Mixed methods approach to determine occupants' behaviour – Analysis of two case studies," *Energy and Buildings*, vol. 130, pp. 546–566, 2016.

[118] L. Stankovic, V. Stankovic, J. Liao, and C. Wilson, "Measuring the energy intensity of domestic activities from smart meter data," *Applied Energy*, vol. 183, pp. 1565–1580, 2016.

[119] K. Gram-Hanssen, "New needs for better understanding of household's energy consumption – behaviour, lifestyle or practices?" *Architectural Engineering and Design Management*, vol. 10, no. 1-2, pp. 91–107, 2014.

[120] L. Canale, B. Peulicke Slott, S. Finsdóttir, L. R. Kildemoes, and R. K. Andersen, "Do in-home displays affect end-user consumptions? A mixed method analysis of electricity, heating and water use in Danish apartments," *Energy and Buildings*, vol. 246, p. 111094, 2021.

[121] S. Athanasoulias, F. Guasselli, N. Doulamis, A. Doulamis, N. Ipiotis, A. Katsari, L. Stankovic, and V. Stankovic, "The Plegma dataset: Domestic appliance-level and aggregate electricity demand with metadata from Greece," *Scientific Data*, vol. 11, no. 1, p. 376, 2024.

[122] A. Vavouris, F. Guasselli, L. Stankovic, V. Stankovic, K. Gram-Hanssen, and S. Didierjean, "Descriptor: A norwegian positive energy neighbourhood dataset of electrical measurements and interviews on energy practices (norpen)," *IEEE Data Descriptions*, 2024.

Bibliography

[123] T. Bjørner, "Why'Qualitative Methods for Consumer Research'?" in *Qualitative Methods for Consumer Research: The value of the qualitative approach in theory and practice.* Hans Reitzels Forlag, 2015, pp. 11–15.

[124] Meteorologisk Institutt, "Norsk Klima Service Senter," 2023. [Online]. Available: https://seklima.met.no

[125] O. L. Brown, "The clausius-clapeyron equation," *Journal of Chemical Education*, vol. 28, no. 8, p. 428, 1951.

[126] Copernicus Atmosphere Monitoring Service (CAMS) Atmosphere Data Store (ADS), "CAMS solar radiation time-series," 2023. [Online]. Available: https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-solar-radiation-t imeseries?tab=overview

[127] W. Schüepp, "Die Bestimmung der Komponenten der atmosphärischen Trübung aus Aktinometermessungen," *Archiv für Meteorologie, Geophysik und Bioklimatologie, Serie B*, vol. 1, pp. 257–346, 1949.

[128] G. Trotta, A. Hansen, L. Aagaard, and K. Gram-Hanssen, *Survey questionnaire on households' use of smart home technology and their time of use of electric appliances*, 1st ed. Institut for Byggeri, By og Miljø (BUILD), Aalborg Universitet, Feb. 2023. [Online]. Available: https://vbn.aau.dk/ws/portalfiles/por tal/554031181/Survey_questionnaire_on_households_use_of_smart_home_technol ogy_and_their_time_of_use_of_electric_appliances.pdf

[129] T. Todic, L. Stankovic, V. Stankovic, and J. Shi, "Quantification of dairy farm energy consumption to support the transition to sustainable farming," in *2022 IEEE International Conference on Smart Computing (SMARTCOMP)*. IEEE, 2022, pp. 368–373.

[130] S. Pfenninger and I. Staffell, "Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data," *Energy*, vol. 114, pp. 1251–1265, 2016.

206

Bibliography

[131] T. Huld, R. Gottschalg, H. G. Beyer, and M. Topič, "Mapping the performance of PV modules, effects of module type and data averaging," *Solar Energy*, vol. 84, no. 2, pp. 324–338, 2010.

[132] Nord Pool SA, "Nord Pool AS," 2023. [Online]. Available: https://www.nordpoolgroup.com

[133] United Nations, Department of Economic and Social Affairs, Population Division, "World Population Prospects 2022: Summary of Results," 2022.

[134] OECD and Food and Agriculture Organization of the United Nations, "OECD-FAO Agricultural Outlook 2023-2032," p. 359, 2023.

[135] International Energy Agency, "Global energy-related CO2 emissions by sector in 2020 and 2050," Online at https://www.iea.org/data-and-statistics/charts/global-energy-related-co2-emissions-by-sector-in-2020-and-2050, 2020.

[136] The Cool Farm, "Cool Farm® Tool," 2024. [Online]. Available: https://coolfarm.org

[137] Department for Environment Food & Rural Affairs, "Agri-climate report 2022," 2022. [Online]. Available: https://www.gov.uk/government/statistics/agri-climate-report-2022/agri-climate-report-2022

[138] Department for Environmnet Food & Rural Affairs, "Farm practices survey February 2023 – greenhouse gas mitigation: Emissions," 2023. [Online]. Available: https://www.gov.uk/government/statistics/farm-practices-survey-february-2023-greenhouse-gas-mitigation/emissions

[139] National Farm Union of Scotland, "Farming Facts, Scottish Farming," 2023. [Online]. Available: https://www.nfus.org.uk/farming-facts.aspx

[140] Agriculture and Horticulture Development Board, "Improving energy efficiency on dairy farms," 2024. [Online]. Available: https://ahdb.org.uk/knowledge-library/improving-energy-efficiency-on-dairy-farms

Bibliography

[141] A. Yadav, A. Sinha, A. Saidi, C. Trinkl, and W. Zörner, "NILM based Energy Disaggregation Algorithm for Dairy Farms," in *Proceedings of the 5th International Workshop on Non-Intrusive Load Monitoring*, ser. NILM'20. New York, NY, USA: Association for Computing Machinery, 2020, p. 16–19.

[142] B. Paris, F. Vandorou, D. Tyris, A. T. Balafoutis, K. Vaiopoulos, G. Kyriakarakos, D. Manolakos, and G. Papadakis, "Energy Use in the EU Livestock Sector: A Review Recommending Energy Efficiency Measures and Renewable Energy Sources Adoption," *Applied Sciences*, vol. 12, no. 4, 2022.

[143] M. I. Malliaroudaki, N. J. Watson, R. Ferrari, L. N. Nchari, and R. L. Gomes, "Energy management for a net zero dairy supply chain under climate change," *Trends in Food Science & Technology*, vol. 126, pp. 153–167, 2022.

[144] A. Moerkerken, S. Duijndam, J. Blasch, P. van Beukering, and A. Smit, "Determinants of energy efficiency in the Dutch dairy sector: dilemmas for sustainability," *Journal of Cleaner Production*, vol. 293, p. 126095, 2021.

[145] E. Martinsson and H. Hansson, "Adjusting eco-efficiency to greenhouse gas emissions targets at farm level – The case of Swedish dairy farms," *Journal of Environmental Management*, vol. 287, p. 112313, 2021.

[146] M. Gargaro, A. Hastings, R. J. Murphy, and Z. M. Harris, "A cradle-to-customer life cycle assessment case study of UK vertical farming," *Journal of Cleaner Production*, vol. 470, p. 143324, 2024.

[147] A. Rotz, R. Stout, A. Leytem, G. Feyereisen, H. Waldrip, G. Thoma, M. Holly, D. Bjorneberg, J. Baker, P. Vadas, and P. Kleinman, "Environmental assessment of United States dairy farms," *Journal of Cleaner Production*, vol. 315, p. 128153, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0959652621023714

[148] K. Buchanan, R. Russo, and B. Anderson, "The question of energy reduction: The problem(s) with feedback," *Energy Policy*, vol. 77, pp. 89–96, 2015.

[149] V. Mitchell, T. Ross, A. May, R. Sims, and C. Parker, "Empirical investigation of the impact of using co-design methods when generating proposals for sustainable travel solutions," *CoDesign*, vol. 12, no. 4, pp. 205–220, 2016.

[150] J. Trischler, S. J. Pervan, S. J. Kelly, and D. R. Scott, "The value of codesign: The effect of customer involvement in service design teams," *Journal of Service Research*, vol. 21, no. 1, pp. 75–100, 2018.

[151] L. J. Bannon and P. Ehn, "Design matters in participatory design," *Routledge international handbook of participatory design*, vol. 711, pp. 37–63, 2012.

[152] C. Eastwood, F. Turner, and A. Romera, "Farmer-centred design: An affordances-based framework for identifying processes that facilitate farmers as co-designers in addressing complex agricultural challenges," *Agricultural Systems*, vol. 195, p. 103314, 2022.

[153] P. Bandeira de Mello Martins, V. Barbosa Nascimento, A. R. de Freitas, P. Bittencourt e Silva, and R. Guimarães Duarte Pinto, "Industrial Machines Dataset for Electrical Load Disaggregation," 2018. [Online]. Available: https://dx.doi.org/10.21227/cg5v-dk02

[154] S. Uski, E. Rinne, and J. Sarsama, "Microgrid as a cost-effective alternative to rural network underground cabling for adequate reliability," *Energies*, vol. 11, no. 8, p. 1978, 2018.

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

[156] Met Office, "Weather DataHub," 2025. [Online]. Available: https://datahub.metoffice.gov.uk

[157] S. J. Taylor and B. Letham, "Forecasting at scale," *The American Statistician*, vol. 72, no. 1, pp. 37–45, 2018.

Bibliography

[158] National Aeronautics and Space Administration, "POWER Data Access Viewer,"
2023. [Online]. Available: https://power.larc.nasa.gov/data-access-viewer/

[159] E. Shove, M. Pantzar, and M. Watson, *The dynamics of social practice: Everyday
life and how it changes.* Sage, 2012.

[160] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton,
A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne *et al.*,
"The FAIR Guiding Principles for scientific data management and stewardship,"
Scientific data, vol. 3, no. 1, pp. 1–9, 2016.

[161] UK Data Service, "Recommended formats: File formats recommended by the UK
Data Service," 2024. [Online]. Available: https://ukdataservice.ac.uk/learning-h
ub/research-data-management/format-your-data/recommended-formats/

[162] C. Klemenjak, A. Reinhardt, L. Pereira, S. Makonin, M. Bergés, and W. Elmenre-
ich, "Electricity consumption data sets: Pitfalls and opportunities," in *Proceedings
of the 6th ACM international conference on systems for energy-efficient buildings,
cities, and transportation*, 2019, pp. 159–162.

Bibliography