Technical and socio-economic evaluation of energy disaggregation using Non-Intrusive Load Monitoring (NILM) for the characterization of residential appliances' profiles: the Netherlands case study.

Federica Blasioli

Thesis to obtain the Master of Science Degree in

Energy Engineering and Management

Supervisors: Prof. Duarte de Mesquita e Sousa

Ir. Arne Kaas

Examination Committee

Chairperson: Prof. Susana Isabel Carvalho Relvas

Supervisor: Prof. Duarte de Mesquita e Sousa

Member of the Committee: Dr. Henrique Ramalho Monteiro Latourrette Pombeiro

January 2021

I declare that this document is an original work of my own authorship and that it fulfils all requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

ABSTRACT

In a context where 97% of the European buildings are energy inefficient, numerous countries are investing in smart meters rollouts, and a severe intensification of the electricity demand from the residential sector is expected, energy disaggregation could serve as a valuable resource for leveraging advanced meter infrastructures and creating opportunities for innovative services.

In this thesis, the Non-Intrusive Load Monitoring (NILM) technology has been evaluated for the residential use-case. The Netherlands was set as reference scenario for the two energy disaggregation models aimed at the recognition of household's appliances fingerprints. An event-based approach was designed using a cost-effective system that allowed the identification of the fridge signature and its duty cycle, bringing the potential to alert consumers about misfunctioning and inefficient devices. Still, compared to the event-less approach using supervised learning algorithms, the proposed model showed less accurate results and appliances classification. Currently, the concerns for data security and the existing technological limits are preventing the NILM technology to be deployed with a high range of accuracy on a larger scale.

Finally, despite a promising market potential, NILM technology was evaluated as still too expensive for direct investments of residential consumers. With utilities being the ones deploying NILM solutions in the households, a broader implementation could be possible instead. Additionally, a detailed appliance-specific breakdown together with a customized feedback system could help energy consumers to save up to 5% of their energy consumption and increase their awareness on energy conservation.

Keywords: Energy Disaggregation; Non-Intrusive Load Monitoring; Appliances fingerprint recognition; NILM economic impact; Energy feedbacks; Residential energy savings.

RESUMO

Considerando o cenário em que 97% dos edifícios europeus são ineficientes em energia, vários países têm investindo na instalação de medidores inteligentes. Sendo esperado um aumento significativo do consumo de energia elétrica pelo setor residencial, a desagregação dos consumos de energia pode servir para alavancar as infraestruturas com medidores inteligentes e a criação de serviços inovadores.

Nesta dissertação, a tecnologia de Monitorização de Carga Não Intrusiva (NILM) foi avaliada para o caso de uso residencial. A Holanda foi selecionada como cenário de referência para os dois modelos de desagregação de energia destinados ao reconhecimento de padrões de eletrodomésticos. Uma abordagem baseada em eventos foi projetada usando um sistema de baixo custo que permitiu a identificação padronizada do frigorífico e do seu ciclo de funcionamento, incluindo a funcionalidade para alertar os consumidores sobre dispositivos com mau funcionamento e ineficientes. Em comparação com a abordagem sem eventos usando algoritmos preditivos, o modelo proposto apresentou resultados menos precisos. Atualmente, as preocupações com a segurança dos dados e algumas limitações tecnológicos estão a impedir que a tecnologia NILM possa ser mais precisa.

A penetração no mercado da tecnologia NILM, apesar de promissora, é dispendiosa como investimento direto de consumidores residenciais. Com políticas públicas para implementação residencial de soluções NILM, uma implementação em maior escala poderia ser possível. Além disso, uma descrição apropriada do aparelho, conjuntamente com um modelo de retorno personalizado, pode ajudar os consumidores a economizar até 5% de seu consumo de energia e aumentar a sua consciencialização sobre poupança de energia.

Palavras-chave: Desagregação energética; Monitorização de carga não intrusiva (NILM); aparelhos de reconhecimento de impressão digital; Impacto económico da NILM; retorno energético; poupança de energia.

Content

Table of Figures

List of Tables

Acronyms

AAL	Ambient Assisted Living
ALM	Appliance Load Monitoring
ANN	Artificial Neural Networks
B2B2C	Business-to-Business-to-Consumer
B2C	Business-to-Consumer
CVD	Continuously Variable Devices
DR	Demand Response
DRED	Dutch Residential Energy Dataset
DSM	Demand Side Management
DSO	Distribution System Operators
EU	European Union
EV	Electric Vehicle
FHMM	Factorial Hidden Markov Models
FN	False Negative
FP	False Positive
FSM	Finite State Machine
GT	Ground Truth
HAN	Home Area Network
HDF5	Hierarchical Data Format
HEMS	Home Energy Management Systems
HMM	Hidden Markov Models
HVAC	Heating, Ventilation and Air-Conditioning
Hz	Hertz
I	Current
IHD	In-Home-Display
ILM	Intrusive Load Monitoring
IoT	Internet of Things
k-NN	k-nearest neighbour
kWh	Kilowatt hour
NAM	Norm Activation Model
NILM	Non-Intrusive Load Monitoring
NILMTK	Non-Intrusive Load Monitoring Toolkit
NILMTK –DF	NILM Toolkit Data Format
NZEB	Net Zero Energy Buildings
P	Real power
PF	Power Factor
PGMs	Probability graphical models
PV	Photo Voltaic
R&D	Research and Development

REDD	Reference Energy Disaggregation Dataset
RQ	Research Question
S	Apparent power
SaaS	Software as a Service
SVM	Support Vector Machines
TMC	Transtheoretical Model of Change
TP	True Positive
TPB	Theory of Panned Behaviour
TSO	Transmission System Operators
UK	United Kingdom
US	United States
V	Voltage
VBN	Value – Beliefs – Norm Theory of Environmentalism
WAN	Wide Area Network

1. Introduction

1.1 Context

In an era characterized by the rising concern about anthropogenic global warming, clear and ambitious commitments are urgently needed to ensure reduced energy consumption and a faster transition towards clean energy technologies. The building sector is globally responsible for more than half of the total electricity consumption and therefore buildings are at the centre of the debate when talking about decarbonization [1].

In the EU, buildings are the largest energy consumers accounting for 40% of the total energy consumption [2]. Recent studies indicate that 97% of the European building stock is currently inefficient while at the same time 75-90% of those buildings will remain in use until 2050 [3]. Although it will become mandatory from 2021 for all new buildings in the EU to be nearly zero-energy buildings (NZEB), it is imperative to develop cost-effective renovation strategies to assure high performance and efficiency of existing buildings [4].

One successful example to boost the rate of building renovations is represented by the project *Energiesprong* (literally "Energy Jump") which originated in the Netherlands in 2009 as a government-funded innovation program to renovate social rental houses and make them *"zero-on-the-meter"*. This definition implies that the renovated houses, typically refurbished with prefabricated facades, insulated rooftops, solar panels, and smart heating, have zero net energy consumption throughout the year. The required retrofits are partially financed by the future savings on energy so that tenants, instead of paying their energy bills, pay an energy service plan fee of a similar amount to the housing associations. This fee is then used to complete the renovations [5].

Due to the collaboration of all the stakeholders in the sector including municipalities, housing corporations, building companies and suppliers, guaranteed by the non-profit organization *Stroomversnelling*, in 2019 the number of zero-on-the-meter houses passed 10.000 units in the Netherlands. Currently, the Dutch model of the *Energiesprong* program is being replicated with great success in many other countries such as England, France, Germany, Italy, Luxemburg, and several US states such as New York, Ontario and British Colombia [6].

Another important initiative being strongly promoted by the EU is the rollout of smart meters which allows an improved and automated monitoring of the electricity consumption in every household, aiming at providing more detailed information to both consumers and utilities and thus, promote a reduction of energy consumption. The rollout of this technology is expected to reach almost 77% of the European consumers by the end of 2024 while several countries are already targeting a penetration of 95% (Italy, Sweden, Finland and the Netherlands) [7][8].

Nevertheless, recent investigations found out that utilities are failing to leverage the advantages of these smart metering infrastructures despite the potentiality of the technology and the high investments made which in Europe accounted for over €21 billion only up to 2020 [8][9]. The lack of achieved energy savings is in part caused by regulatory, administrative and technological challenges. However, utilities are missing the opportunity to pair smart meters with other technologies that would provide the customers with more engaging tools to motivate them to adopt a pro environmental energy behaviour. Among the possible use-cases of the data obtained from the smart meters, feedbacks and energy disaggregation were revealed to have high potential to encourage energy consumers to exploit many of the benefits of this advanced technology and thus, save energy [9].

In this context, aggravated by an expected severe intensification of the electricity demand of the residential sector, especially for cooling, appliances and plug loads such as electric vehicles, the implementation of an algorithm that can disaggregate the specific electric loads in a household, being able to provide real-time insights to the tenants and detailed tailored information about the specific appliance's usage, could contribute not only to a better energy management and smart meter leveraging but also to foster the consumers' engagement in the cause.

Inspired by numerous progress in the field of artificial intelligence, machine learning, networks and data communication, research efforts in this area have lately been focusing on the concept of Appliance Load Monitoring (ALM). This method consists of performing a breakdown of the total energy consumption data, measured at house level by the smart meter, to foster the development of energy conservation strategies such as demand response, to forecast energy demand or to provide a better understanding of the usage patterns of each individual appliance and consecutively of the consumers' behaviours [10].

Every appliance is characterized by a signature, or *fingerprint*, representing its specific consumption profile that can be detected from the overall electricity consumption of a house in a process called *disaggregation*. When the data for disaggregation are provided by adding just one sensor measurement to the smart meter, the approach is called *Non-Intrusive Load Monitoring* (NILM).

Despite the concerns for privacy and the current technological limits that still prevent this technology to be deployed with high range of accuracy, if broadly implemented, NILM could serve as a valuable resource for understanding the variability of energy consumption and creating an opportunity for innovative services to increase awareness on the issue of excess energy consumption, finally both reducing the large impact of energy utilization in buildings and expanding grid flexibility capacity.

However, despite the many initiatives being proposed, those alone are not enough to accelerate the shift towards a more sustainable, reliable and efficient energy system. In fact, according to a study conducted within the *Energiesprong* project about the renovated houses that were not reaching the status of zero-on-the-meter, the incorrect resident's behaviours were representing as much as 81% of the reasons [11]. It becomes of fundamental importance then, to engage the energy consumers by increasing their awareness on the consequences of their actions and enhance intrinsic motivational strategies to let them naturally transit to a sustainable behavioural change [12].

In conclusion, this thesis hypothesises that in this framework where smart meter rollouts and energy efficiency programs are largely incentivised, energy disaggregation could have the potential of fostering residential energy savings by increasing both the sense of control and awareness of energy consumers on the issue, and by leveraging the smart metering infrastructures.

1.2 Objectives and Research Approach

This thesis aims to provide a comprehensive understanding of the NILM technology from a technical, economic and social point of view. An analysis was performed to disaggregate the total electricity consumption data of a household located in the Netherlands, obtained from smart meters with low-frequency sampling resolution (0,1 Hz), to determine the individual load profiles of common household appliances, identifying their characteristic signatures and power consumption. For this purpose, the fridge is used as the case-study appliance.

Moreover, this technology was analysed from an economic and social prospective in order to evaluate the potential impacts that energy disaggregation could have on the consumers. After defining the customer segmentation and the possible business models, the market potential was studied for three main scenarios: United States, Europe and the Netherlands. A further social analysis integrated some theories of behavioural change to evaluate which are the best feedback mechanisms that could foster customer engagement and consequent energy savings.

The following research questions are defined:

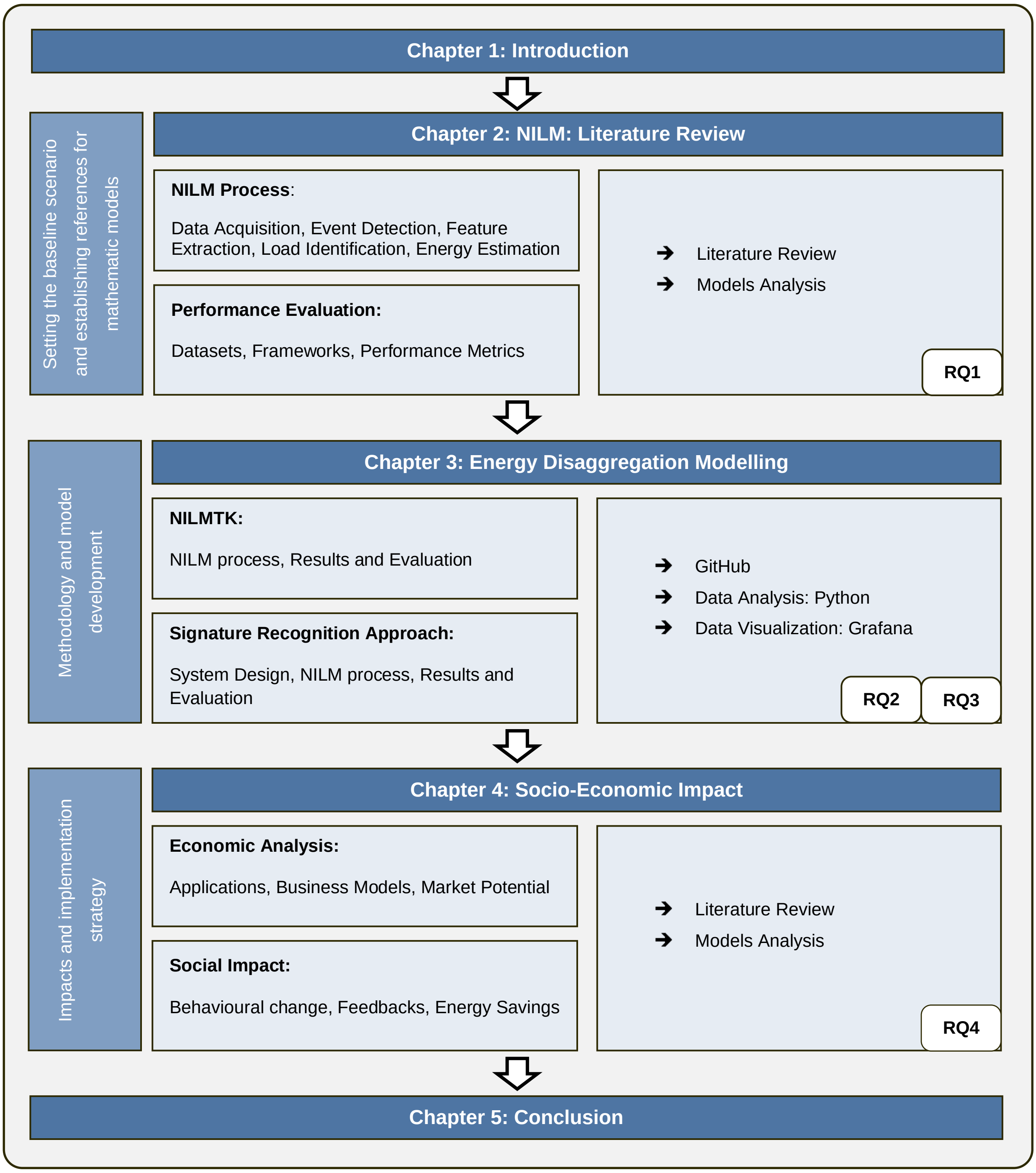
RQ1	*What is the Non-Intrusive Load Monitoring (NILM) technology and how is the NILM process developed?*
RQ2	*Which are the most common algorithms that allow to extract the devices' profiles from the overall electric load profile of a household and how effective are they?*
RQ3	*Is it possible to develop an algorithm, without using a machine learning approach, that can recognize, with suitable accuracy and without the necessity of a training dataset, some of the basic appliances in a household?*
RQ4	*Which are the benefits of energy disaggregation to the consumers? What is the potential socio-economic impact that this technology could bring?*

To this end, the Energiesprong project in the Netherlands is set as a reference scenario to better evaluate the impact that the identification of single appliances consumption has on the energy consumption of a household and how this could lead to major savings for the consumers. The necessary data to perform the study are provided by Watch-E, a Dutch company that monitors and manages the energy performance of zero-on-the-meter houses while providing the residents with relevant insights on their consumption.

1.3 Thesis Outline

Table 1.1 summarizes the main structure of which the thesis is composed, addressing each research question to a specific chapter. After a brief introduction, a review of the literature in chapter 2 includes a detailed description of the modules of which the NILM process is composed and the tools to perform the models' evaluation. Once the baseline scenario is defined, the energy disaggregation models are presented in chapter 3. The technical analysis is performed with two different methods: a toolkit (NILMTK) publicly available on GitHub and developing a signature recognition algorithm. The results and their evaluation are presented for each methodology followed by the challenges and limitations encountered in the process. Afterwards, in chapter 4 the economic and social analysis is performed defining the impacts and the possible implementation strategies for a broader employment of NILM. Finally, the conclusion summarizes the main outcomes of the thesis.

Table 1.1 Thesis Structure



2. Non-Intrusive Load Monitoring (NILM): Literature Review

The recent increase in smart meter installations in households and the numerous initiatives to promote energy efficiency have led to a large interest in load monitoring techniques to provide better services and obtain useful information about appliance usage and user consumption behaviour. The term load monitoring refers to the process of identifying and acquiring load measurements from a power system which can then be exploited to recognize the status, the behaviour and the consumption of every single appliance [13].

The most traditional approach is called *Intrusive Load Monitoring* (ILM) and it relies on using low-end metering devices connected to one or few appliances or loads under interest. The ILM has multiple points of measurements inside the house, thus the term intrusive, and it is able to recognize each appliances signature accurately. Consequently, it provides detailed information, especially for low-power and standby devices which are difficult to recognize using NILM. On the downside, the multiple sensors configurations, the complex installation and the high costs associated with ILM make it unattractive for a widespread implementation. However, ILM still remains useful as a way of evaluating NILM environments and for creating datasets [14].

Hence, despite the more complicated disaggregation process that it implies, NILM remains the most common approach for load monitoring due to its non-intrusive characteristic of one sensor metering that has to be installed in the house without extra equipment. Recent years have shown a growing research effort related to NILM, among which there are a large number of publications in scientific papers and companies investing in this technology and offering NILM related services [15].

With the high adoption rate of electric vehicles (EVs) and photovoltaic (PV) systems, the increasing numbers of prosumers need granular and instant information about their energy consumption and generation to optimize their own energy utilization [16]. Moreover, the great potential of NILM goes far beyond the individual energy savings and the already considerable reduction of the carbon footprint associated with energy consumption. According to the first international NILM workshop in 2016, it will indeed enable the construction of innovative smart-grid services which for example could allow grid operators to improve predictions on energy demand, implement targeted demand-side response, enable a higher accuracy on energy system models, or detect anomalies in the electric loads [17].

Another very interesting focus for NILM's application is in the field of Home Energy Management System (HEMS) and Ambient Assisted Living (AAL) where the possibility of identifying the exact appliances which are in use could foster the research on personalized care, especially for elderly people or patients with health problems [18].

Despite the many possible applications of this technology further mentioned in chapter 4.2.1, this thesis will be mainly dedicated to the residential sector and the understanding of appliances consumption patterns. Also, part of the analysis will include an assessment of the numerous challenges and limitations to be solved for a large-scale, reliable and cost-effective NILM solutions implementation.

2.1 General Framework of NILM

The concept of NILM was first introduced in 1985 by George W. Hart from the Massachusetts Institute of Technology. He defined the aggregated load of a household for a total number of n active appliances at a time *t* as follows in *Equation* (1):

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

Where $P(t)$ is the aggregate power load at the entry point of the meter, $P_i(t)$ is the power consumption of the active individual appliances contributing to the total load and $e(t)$ is a noise or error term [19]. The basic assumption behind this first equation is that every change in the total electrical load of a household is the consequence of an electric device changing its state, such as turning on or off the television.

The following *Figure* 2.1 shows how the method works. The graph shows a representation of the *event-based* approach, associated with the first epoch of NILM, in which the disaggregation was pursued by detecting and labelling every appliance transition in the aggregated signal by using previously trained supervised or semi-supervised learning algorithms. Nowadays, the most commonly explored approach is the *event-less* approach in which the disaggregation is performed through sophisticated machine-learning algorithms which try to match the knowledge of the operation of an individual device or the combination of few of them, without the necessity of having prior labelling and event detection.

The NILM process is inclusive of all the signal processing and machine-learning algorithms used to estimate the individual appliance consumption from the total demand of the selected location [20].

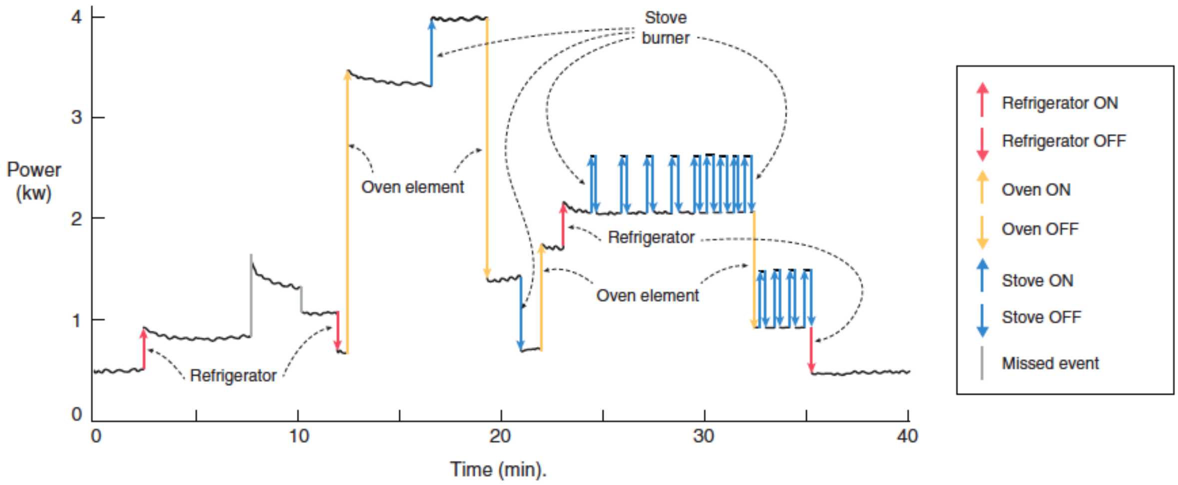


Figure 2.1 Example of event-based energy disaggregation [20]

2.1.1 Categories of Appliances

Due to different internal structures, working patterns and working environments, each device has its own characteristic signature which exhibits a unique load feature when the appliance is in operation. This allows every specific load signature to be discernible from others. Since these signatures, besides the temporal behaviour, also depend on the operational states of the devices, it is important to distinguish between four different types of appliances. The typical power curves for the first three of those appliances types are illustrated in *Figure* 2.2 [21][22].

- *Type I*: ON/OFF appliances. These are the easiest appliances to recognize and are characterized by only two operational states: on or off. Examples include table lamps, toaster, electric kettle, etc.
- *Type II*: Finite state machines (FSM). These are multi-state appliances with a limited number of operating states and repeating switching patterns. Examples include washing machines, stove burners, clothes dryers, etc.
- *Type III*: Continuously variable devices (CVD). These are the most complicated appliances to recognize because their power draw is not fixed and thus, they keep changing states during their operation. Examples include power drills, dimmer lights, etc.
- *Type IV*: Permanent consumer devices. These are appliances that remain regularly active, consuming a constant power over time periods. Examples include hard-wired smoke detectors, telephone sets, cable TV receivers, etc.

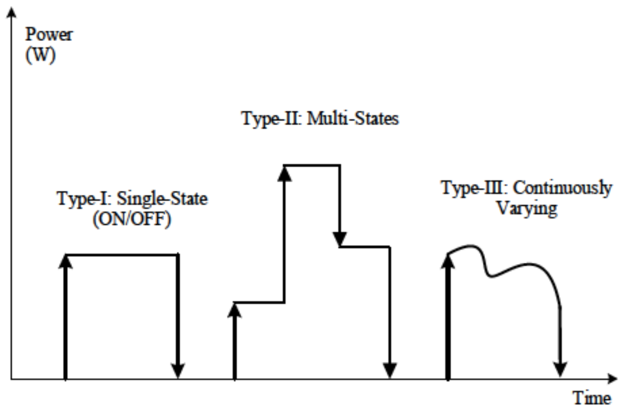


Figure 2.2 Power curves of three types of appliances' loads [10]

Classifying appliances according to the categories listed above is used to improve and facilitate the identification of the load signatures. However, many devices have a multi-working model based on the way consumers use or customize them and thus, cannot be clearly assigned to a single category. A new classification approach based on the main basic units and the working style of residential appliances has been proposed in the literature [23]. Taking this into consideration, the classification presented in this paragraph may lead to larger errors that should be accounted in the NILM process.

2.2 NILM Process

In almost four decades since NILM was firstly introduced, a large amount of literature has been produced over the topic, including academic reviews and many scientific papers proposing different approaches to obtain disaggregation through the NILM methodology. Due to the complexity of the technique, the numerous data requirements and the recent introduction of machine learning and artificial intelligence algorithms, there is no single standardized procedure to perform NILM.

Figure 2.3 shows a flowchart that aims at simplifying the overall NILM process, constituted of five main modules. It is important to specify that the second and third modules, respectively event detection and feature extraction, are only representative of the event-based approaches as it will be clarified in chapter 2.2.2. The following subchapters will provide further information on each of the five modules.

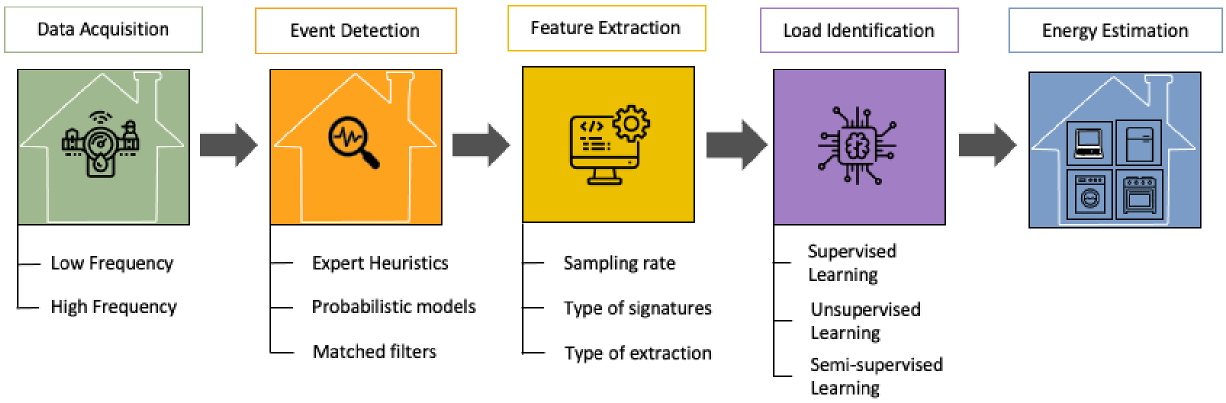


Figure 2.3 Description of NILM process

2.2.1 Data Acquisition

The first module of the disaggregation process consists of collecting and pre-processing the aggregated load measurements from the smart meter connected in the household. This initial stage is of fundamental importance for the rest of the analysis because it will influence the kind of NILM algorithms that can be used and consequently their performance. Two criteria must be taken into account in this step: the sampling rate of the collected data, and the kind of hardware architecture that is implemented [18].

Usually, the power meters measure the alternating current (AC) of the electrical system and consequently the two basic values of voltage (V) and current (I). From these signals, other parameters and their variation over time can be determined such as real power (P), apparent power (S), the power factor (PF) and additional characteristics like the I-V trajectory [13]. This usually happens in the pre-processing phase when the gathered raw data is resampled and adjusted before being collected and stored. An important emphasis though needs to be given to the sampling rate of this collected data which is responsible for the type of information that is extracted from the electrical signals. The sampling rate can be further classified as follows:

- Low Frequency Data: lower than 1 Hz. The most common kind of smart meters that are nowadays commercially available are typically able to collect frequency sampling of minutes or hours. Being a low-cost metering solution, the collected data only presents a limited functionality for further processing.

- High Frequency Data: higher than 1 Hz (kHz to MHz). This kind of frequency sampling has a very high resolution able to capture current and voltage waveforms (at around 10 to 100 MHz) and higher order harmonics of the electric signals. The collection of more fine-grained data for better identification of load features also involves the increase of the computational complexity of the hardware used which often needs to be custom-built. For this reason, the deployment of high frequency smart meters implies high installation and maintenance costs.

Data acquisition presents many challenges. The noise produced by variable speed (colour noise) and permanent (white noise) devices reduce the accuracy of raw data, network malfunctions lead to missing readings, and data compression in the power meters is responsible for the loss of raw data [21].

The auxiliary hardware implemented to collect specific data is responsible for the communication between local devices and the remote computing centres. It should present a minimum of computing capacity (especially for the high frequency sampling) and since it has to be installed inside the household, it should be designed as portable, plug-in and easy to connect. For example, many solutions already deployed in the market are based on current clamps and a set of utilities to attach to the smart meter already installed in the house [18].

The acquisition of data is a very sensitive part of the NILM methodology, being also time-consuming and associated with high costs. However, several databases are already publicly available in the literature and allow to test the different algorithms in diverse environments. Those will be introduced in chapter 2.3.1.

2.2.2 Event Detection

An event is defined as the variation in the signature curve of a device, provoked by a change in the state of the device itself. These events are detected in the second module of the NILM process, where the total power consumption is analysed to identify any load switching operation and to filter out any useful information from the whole collected data, in order to not affect the computational efficiency of the technology [21].

However, before proceeding with the explanation of event detection, it is important to further clarify the difference between event-based and event-less approaches for energy disaggregation. While the event-based approach, as previously described (*Figure* 2.1), requires the detection and the labelling of every appliance transition from the aggregated load, the event-less approach, does not rely on event detection and tries to match each sample of the total consumption with the consumption of one or different combinations of specific devices [15]. The latter approach applies statistical and probabilistic machine-learning methods without the necessity of any labelled information. Thus, the whole process is more straightforward, but at the same time more computationally intensive especially when compared to the event-based approach which necessitates previously trained supervised or semi-supervised learning algorithms [20]. On the other hand, the event-based approach requires less computational power because the feature extraction is only performed when an event is identified and not on the whole dataset, which also leads to a higher possibility of error propagation. Overall, each module of the analysed NILM process is needed for the event-based approaches while for the event-less approaches two modules are not included in the process: the event detection and the feature extraction.

Therefore, the event detection module (only valid for the event-based approach) can use three different approaches [15][18][24] :

- Expert heuristics: focus on the pre-definition of a specific threshold range of the power variation for each specific appliance (usually the most consuming ones), scanning the data and looking for changes that are above these thresholds. This process requires the initialization of certain variables.

- Probabilistic models: calculate the probability of an event occurring at each sample of power signals using training processes and statistical models on specific parameters. The results are later used to decide on the event occurrence.

- Matched filters: used only for high frequency sampling data, this approach extracts the signals waveforms through filtering techniques and without any previous training. It tries to correlate them with known signals in the aggregate consumption.

Finally, the events detected, if correctly representative of the appliance state transitions (for example on/off state), are crucial to determine the most typical features of a certain device. Those features are useful for a later identification of the device itself.

2.2.3 Feature Extraction

Once the event points have been detected, different parameters can be extracted from the vicinity of the power event (features) creating a power signature that will be used in the following module for the load identification of each device. It is important to mention that since this module depends on the event detection, it is only applicable for event-based approaches. According to the literature, there are several ways for classifying the different kinds of features extracted. Hereunder, three categories are presented by combining different methods from the most recent published papers [21][25].

2.2.3.1 Sampling rate

In this method, the sampling rates refer to the ones of the data output after the measurements of the device, not to be confused with the sampling rates classified in the data acquisition module relative to the device input (current and voltage). These rates vary from MHz to hours and are organized first into two main classes: macroscopic or low-frequency and microscopic or high-frequency. These two classes can further be divided into six additional classes grouped as follow [18]:

- *Very slow* (slower than 1 minute) and *slow* (between 1 minute and 1 second): obtained from the time series of power variables such as voltage, current, apparent, active or reactive power, etc., and employed either directly, for statistical moments or for time-domain transformation.

- *Medium* (faster than 1 Hz) and *high* (up to 2 kHz): can detect more characteristics of the transient state, extracting features such as harmonics or frequency bands from the frequency information.

- *Very high* (frequency between 2 and 40 kHz) and *extremely high* (faster than 40 kHz): allow for extremely detailed information about the device’s waveforms.

2.2.3.2 Type of signatures

This method depends on the operating states of the appliances and is the most common type of classification. It can be divided into three categories [10][13][21].

- Steady State Features: when the device is in steady state operation. One way to extract features from type I devices (on/off) is by analysing the variations of real power and reactive power. Other features in this category are obtained from the time-frequency analysis of V-I waveforms, from the V-I trajectory or from steady-state voltage noise but with the necessity of higher sampling rates.

- Transient State Features: when the device is in a transient state. Those are more efficient to deal with overlapping problems compared to the steady-state features although to be extracted these require a high sampling rate. The analysis of transient power allows the recognition of type I, type II and type III devices, even the ones with similar power characteristics. Some other features of multi-state appliances (type II) can be extracted from the start-up current knowing its shape, size and duration or from the transient-state voltage noise but with the need of large computing capacity.

- Non–traditional Features: derive from new external characteristics. Some examples are the time of the day, on and off duration distribution, the frequency of the device usage, the correlation between the usage of more than one appliance, occupancy, outdoor temperature, and even the user's behaviours.

2.2.3.3 Type of extraction

The last method, proposed in [15], classifies all the features in two main categories according to the origins of the extraction:

- Engineered features: extracted from the data. These are the ones extracted from the electrical power information and the knowledge of the appliance characteristics, requiring that the events have been previously detected. In this classification, all the steady-state features obtained from the variation of power characteristics and the V-I trajectory analysis, and some other features obtained from the frequency domain content are included.

- Data driven features: learned from the data. These are the ones extracted from the measurements surrounding the event and are previously detected by using unsupervised feature learning applied directly to the data. This classification includes for example transient-state features obtained from the current waveforms.

2.2.4 Load Identification

Once the features have been extracted from the total aggregated load, it is necessary to determine and classify which appliances are operating at a specific time and also in which state they are operating. This becomes even more difficult when considering that in a typical household, many devices are active contemporarily with some of them having a similar signature. In some cases, several operating appliances of the same kind will be used in the same house such as two televisions or two fridges, which complicate the load identification even more [25]. Thus, the module of load identification is one of the most critical in the whole NILM process and for this reason a multitude of different approaches have been proposed. Currently, the main disaggregation algorithms in use are usually classified into supervised or unsupervised learning according to whether the data were trained or not before the classification. Furthermore, since training the data is a process that requires the availability of many pre-labelled data, and consequently, an increase in time, cost and effort, novel researches focus on semi-supervised approaches that only require partial training of the data [13].

These methods for load identification are briefly presented and further classified. Yet, the aim of this section is not to provide an exhaustive analysis of the mathematics behind those algorithms but rather to offer a general overview of the different kinds of approaches currently in use.

2.2.4.1 Supervised Learning

Most of the research work for NILM is focused on supervised learning. This necessitates labelled data for training classification algorithms, that at the same time constitutes itself as one of the major obstacles due its scarce and costly availability. Another important distinction to be made is that most of the supervised algorithms rely on the detection of events to perform the classification, being then typical for event-based approaches.

Supervised learning algorithms can be further divided into [10][13][21]:

- Optimization: the extracted features are compared to the load features stored in a database in order to find the most accurate match of appliances to have caused the output, minimizing the errors. One of the main disadvantages, besides the large computational capacity that this approach requires, is that the recognition becomes more complicated when there are loads in the aggregate power that are unknown to the database. Some techniques in this category are integer programming and genetic algorithms.

- Pattern recognition: the extracted features are compared to the load features stored in a database in order to recognize the event associated with the operation of an appliance. The advantage, in this case, lies in the introduction of state transition and temporal information while the main drawbacks are the lack of data to test and the overlapping signatures for appliances with low-power consumption. This method, which is indeed the most commonly employed for NILM, uses machine learning techniques with offline training from a database that contains multiple specific features information for future predictions. Some examples are Hidden Markov Models (HMM), Artificial Neural Networks (ANN), Naïve Bayes classifier, Support Vector Machines (SVM), k-nearest neighbour (k-NN).

2.2.4.2 Unsupervised Learning

As anticipated in chapter 2.2.1, the scarcity of trained data availability, the time-consuming labelling of data and the high cost associated with this process, pushed research to move towards unsupervised learning which instead only requires unlabelled data for training classification algorithms. Unsupervised techniques only require a short training phase, low set-up costs and no previous information that would allow a reduced intrusiveness for the data access and easy deployment. Since these methods do not use previously trained data and preserve the representation of the whole system state instead of dealing with individual events, they are typical of event-less approaches. However, this characteristic leads to the disadvantage of being very computationally expensive to implement because every part of the sample has to be constantly taken into account. Moreover, the techniques using unsupervised machine learning algorithms are still delivering a low accuracy performance.

Unsupervised algorithms can be classified as follows [10][13][15][18]:

- Motif mining: this approach works by searching for power change events from the aggregate load signature, identifying recurring episodes that have to fulfil certain conditions to be matched with a specific device. Those episodes are composed of a sequence of power events that can be attributed to the on/off operational states of the device. This method is still not well developed for continuously variable devices and for appliances of similar signatures.

- Blind source separation: this technique implies that an individual signal with very little information regarding its source is separated from the main signal known to be composed of a set of other signals. When applied to NILM, the real and reactive power changes have been clustered each representing an appliance state transition. Then, a matching pursuit algorithm tries to minimize the distance between the unknown event to be recognized and the possible clusters. Multi-state appliances represent a large limitation due to the high number of different clusters that those would involve.

- Probability graphical models (PGMs): those are probabilistic approaches that only require the power consumption and ultimately some non-power features related to the time of the device's operation. The basic behind these techniques is that the total power consumption is characterized at a specific instant *t* by the consumption of several appliances in a specific state of operation. Then, the solution should find the best possible sequence of states that describes the observed aggregated power. These models have been developed using many different variations of the factorial Hidden Markov Models (FHMM), which is constituted by a series of several independent HMMs modelled by non-powered features. While some of the models have developed high average precision, the algorithms are not able to recognize small loads in the presence of larger ones and furthermore their accuracy still drastically decreases with the increasing number of target appliances.

2.2.4.3 Semi-supervised Learning

The semi-supervised learning are less explored methods whose aim is to reduce the disadvantages of both supervised and unsupervised techniques, addressing the problem of the scarcity of data. It requires a compromise between both labelled and unlabelled data for training classification algorithms, reducing the efforts of the training phase. This can happen in two ways. Either some additional information is used to assist an unsupervised algorithm (semi-supervised clustering), or some unlabelled information is adjusted to fit into a supervised algorithm (semi-supervised classification) [26].

2.2.5 Energy Estimation

Energy estimation is the final phase of the NILM process. It includes all that detailed information regarding the disaggregated appliances that are provided to the consumers. First of all, the total electricity consumption per device and its relative cost is recognized and calculated. Furthermore, based on the results, some personalized advice or forecasts on the electricity consumption will be offered to the customers with the possibility of involving also the energy companies which are interested in the energy demand. More of these feedbacks will be discussed and analysed in chapter 4.3.2.

2.3 Performance Evaluation of NILM

The performance assessment of the numerous and variegated solutions proposed to obtain disaggregation represents one of the major challenges for researchers. The introduction of a common formal method to evaluate the results could enable the generalization of research findings and lead to a larger adoption of this technology. Only recently, many authors started focusing on standardizing existing datasets and creating a consensus over the evaluation metrics to be adopted to report the results of the different disaggregation algorithms [20]. In this chapter, publicly available energy datasets, frameworks and performance metrics will be introduced.

2.3.1 Public datasets

A household energy dataset for disaggregation is composed of electrical energy measurements of a realistic scenario in which everyday life is not distorted. The kind of measurements that can be collected are either the aggregated load of the house (from the mains), or the single loads of each device obtained directly at the plug-level or at the circuit-level. In the latter case, the house's circuits are often connected to multiple appliances at a time which means that these kinds of data do not always ensure the availability of all the individual devices' consumption. Only the datasets that contain both aggregated and individual load consumption are suitable to be used as training and testing data. The datasets with just individual appliance consumption can only be used for training the data since assuming the sum of the different loads to be the aggregate power could lead to inaccuracies and unrealistic precision in the disaggregation which will overestimate the quality of the results [15].

Moreover, there are many other factors that influence the selection of a dataset according to the kind of algorithm that is intended to be used. Data can be collected at low or high frequency and the measurements can include different types of features such as real power, reactive power, V-I trajectory or temperature. Additionally, the location is also an important element to consider when choosing a dataset, not only due to the different weather conditions, consumption patterns and behaviours, but also for the kind of appliances used, and in some cases even for the discrepancy in the frequency of the national electric grid itself. Other important factors that influence the final disaggregation results are the duration of the measurements and the number of analysed households. For example, longer periods of training data availability can benefit the identification of periodic consumption behaviours.

Furthermore, there are many complications that arise when creating an accurate dataset. The process of capturing and compiling a large amount of data in a long range of time (months or also years) is time-consuming, the hardware is expensive because many sensors must be installed at the appliance-level (especially for high frequency data) and finally privacy concerns can limit its implementation [21]. For this reason, to date there are only 27 publicly available datasets that are suitable to perform NILM, 18 of which contain both aggregated and individual load consumption.

The first to be introduced was the Reference Energy Disaggregation Dataset (REDD) in 2011, which is still the most widely adopted dataset. The data was collected at both low frequency (3 seconds) and high frequency (15 kHz) in the state of Massachusetts USA, for six different households. Each house has its whole-home aggregate and circuit consumption measured for a duration that varies from days to several months [27].

The Dutch Residential Energy Dataset (DRED) has been released in 2015 and is of particular interest for this thesis because it contains six months data for one household located in the Netherlands. The peculiarity of this low-frequency dataset (one-second or one-minute sampling rate) lays in the extra information provided such as occupancy and ambient parameters of the households (temperature, appliance-location mapping, etc.) [28].

Released in 2020, the latest IDEAL Household Energy Dataset contains data from 255 UK houses in the period of two years. Some of the houses also include other information about occupancy, inhabitants behaviours and extra characteristics of the devices [29].

Since the disaggregation results are strictly dependent on the data source and the type of dataset selected for the analysis, it is of fundamental importance to choose an appropriate dataset. For this purpose, many references provide tables with detailed information on the different datasets that aim at facilitating the decision-making process [20][22].

2.3.2 Frameworks

Despite the great efforts to release new datasets, there is a lack of a homogeneous structure and a universal format between them that would allow an easier understanding, replicability and evaluation practice.

One of the earliest important attempts to create such a uniform approach was released in April 2014 and it consists of an open-source toolkit to perform energy disaggregation using existing datasets and algorithms with a unified interface: the NILMTK [30]. Written in Python and freely available on GitHub [31], it proposes a standard data format (NILMTK–DF) based on the Hierarchical Data format (HDF5) modelled over six main datasets, a series of benchmark algorithms and a set of standard metrics functions [32]. Since then, the toolkit has been simplified becoming the main reference library for energy disaggregation and driving the NILM research community to drastic changes and improvements [33].

After that, other related projects were proposed. An example is the NILM Metadata, a hierarchical metadata schema to improve the representation of other typical information for energy disaggregation such as appliances, electricity meters, measurements, buildings and datasets [34].

Another approach is the Location-aware Energy Disaggregation Framework (LocED) which was released together with the DRED dataset in 2015 [35]. The authors objective was to use the information about the occupants' location inside the household in order to reduce the number of potential appliances that are operating in a specific moment. Thus, these appliances are excluded from the NILM algorithm in the period of time in which the person was not present in the room associated with the device's activity, consequently lowering the computational complexity of the algorithm. The occupancy is obtained using indirect and cost-effective approaches relying on Wi-Fi and Bluetooth signals from the inhabitants' smartphones [28].

2.3.3 Performance Metrics

With the implementation of public datasets and frameworks to test their algorithms, researchers need some established accuracy metrics to evaluate their works and compare them with the ones of others. The term disaggregation accuracy has been used over the years to refer to many different observations and thus, it became inconsistent as a proper evaluation comparison.

One of the proposed approaches distinguishes between two possible event detection errors, Type I when a detection occurs but the appliance is not operating, and Type II when the detection is missing but the appliance is operating. The method suggests three accuracy measures: detection accuracy (which consider the ratio of correct detected events over the total detections), disaggregation accuracy (which does not take into account the effect of wrongly detected events) and overall accuracy (which considers both wrong and missed detections) [36].

Other metrics for this event-based evaluation approach derive from the confusion matrix in which the number of intervals when an appliance state is confused or correctly identified is counted (for example calculating rates or percentages from true positives, false positives, true negatives and false negatives). Some of those statistical metrics are precision, recall and F_{score}. The *Precision* metric is the proportion of the true positive events against all the events that were reported to be positive, interpreted as the measure of how many detected events are relevant. The *Recall* metric is the proportion of the true positive events against all the truly positive events, interpreted as the measure of how many relevant events are properly detected [37]. The two metrics are reported in *Equation (2)* and *Equation (3*) respectively, where TP are the true positive events, FP are the false positive and FN are the false negative ones [30]:

$$Precision = \frac{TP}{(TP + FP)} \qquad (2)$$

$$Recall = \frac{TP}{(TP + FN)} \qquad (3)$$

The F_{score} is given by the harmonic mean of precision and recall metrics in *Equation (4*) and it's a number that ranges from 0 (either precision or recall is zero) to 1 which indicates the perfect accuracy of state detection. In this case, both the precision and recall metrics have the same weight on the formula (F_1) [30].

$$F_{score} = \frac{2xPrecision\ x\ Recall}{(Precision + Recall)} \qquad (4)$$

The F_{score} metric is used to indicate how good the algorithm is in classifying the operating states of an appliance, but it does not provide any information about the power consumption of the appliance [37].

In case of event-less approaches where there is no event detection module, some methods are necessary to evaluate the final energy estimation. Those include metrics based on both basics and advanced statistics such as root mean squared error (RMSE), error of the total energy assigned (EA), and standard deviation error (SDE) as well as other metrics specifically proposed for NILM [20].

In particular, the RMSE of the assigned power is the root mean square error between the actual aggregated power of the *n* appliance in each instant of time *t* ($y_t^{(n)}$) and the estimated power demand of that *n* appliance ($\hat{y}_t^{(n)}$), with a number of total samples recorded T. This metric is expressed in *Equation (5*) and it represents how concentrated the estimated values are around the true values:

$$RSME = \sqrt{\frac{1}{T}\sum_t \left(y_t^{(n)} - \hat{y}_t^{(n)}\right)^2} \qquad (5)$$

This metric is able to estimate the error of assigned power of each appliance and it's a number greater than 0 which represents the perfect fit with the data. The RSME though, is not a normalized measure and for this reason, it is hard to compare how the disaggregation of different appliances performed.

The error in total energy assigned (EA) in *Equation (6*) is given by the absolute value of the difference between the actual energy consumed by the *n* appliance over the dataset and the estimated power consumption of that appliance [30]:

$$EA = \left|\sum_t y_t^{(n)} - \sum_t \hat{y}_t^{(n)}\right| \qquad (6)$$

In the literature it is also possible to find very comprehensive tables summarizing the different metrics currently employed to evaluate energy disaggregation performances [20].

In conclusion, if the aim of disaggregation is to correctly identify the operational state of each device, then the metrics used should be the ones evaluating the accuracy of the NILM algorithm such as precision and recall. If the aim is to provide a detailed electricity bill, then the metrics used should be the ones checking the percentage of energy correctly identified.

3. Energy disaggregation Modelling

In this section of the thesis, a technical analysis on NILM is performed using two different energy disaggregation models. The first one includes the execution of the NILMTK for REDD and DRED datasets to understand the performances of two of the most common algorithms already publicly implemented within the toolkit. The second model aims at the development of an event-based algorithm able to identify the signatures of common appliances in a Dutch household, by analysing the differences in power in the total electricity consumption.

3.1 Methodology

The energy disaggregation was gradually performed in five steps:

1. NILMTK: energy disaggregation of REDD dataset.
2. NILMTK: energy disaggregation of DRED dataset.
3. Collection of data from Dutch households.
4. Analysis and visualization of the collected data.
5. Development of a signature recognition algorithm.

Publicly available on GitHub, NILMTK was used as a tool to understand the quality of performance of two common NILM algorithms: Combinatorial Optimization (CO) and Factorial Hidden Markov Models (FHMM). The analysis was first performed using the standard REDD dataset, with which the toolkit was designed, and then executed with the Dutch DRED dataset in order to tailor it to the target area of this thesis.

In a second phase, a system to collect the data from multiple households, already provided with smart meters, was developed and the data was visualized in Grafana and saved in SQL databases. Moreover, an attempt was made to implement the collected data within Watch-E in the NILMTK. This failed due to a mismatch between the new dataset missing both the individual device's measurements and metadata, and the proposed HDF5 format.

Additionally, an algorithm was developed and implemented to perform energy disaggregation on the collected Dutch household's data, without having any data about the loads of individual appliances. The home conceived code is also available on GitHub and it was written using the programming language Python [38].

Due to the differences in the two proposed approaches, the methodologies for each of them are separated and structured according to the five phases of the NILM process presented in the literature review. The description of the two models is thus directly followed by the respective results and their evaluation to avoid confusion and increase the fluidity of the reading.

3.2 NILMTK

As anticipated in chapter 2.3.2, the NILMTK was designed to enable the analysis of different datasets and algorithms and to create a uniform and simple interface to perform those algorithms and implement others. It is an open-source software available online on GitHub with detailed documentation that guides both the users and the developers through each step of the process from the installation of the NILM environment to the disaggregation algorithms and the evaluation metrics.

In this chapter, the NILMTK model of energy disaggregation with publicly available datasets is proposed. Being the built-in algorithms implemented for event-less approaches, the method only includes the phases data acquisition and load identification. Afterwards, the results of the analysis for the two datasets are presented with a comparison of their performances.

3.2.1 Data Acquisition

The first phase of the NILM process is data acquisition, or more specifically for this case, the selection of the dataset, its conversion into an appropriate file format and the understanding of the collected data and metadata.

Originally modelled over the REDD dataset, the data format (NILMTK–DF) proposed by the authors currently includes importers for twelve different datasets. The remaining available datasets were excluded due to the lack of either aggregate data, individual devices consumption or for time constraints. Once imported, the dataset, to be described in all its aspects, is converted into HDF5 binary file that is used to store both the data itself (usually CSV format) and the metadata (YAML format). In addition, this type of data format allows the nomenclature of each dataset to be standardized to match a specific vocabulary for disaggregation labels, from the name of the appliances to the one of other descriptive parameters [30].

In this regard, the NILM Metadata schema (introduced in chapter 2.3.2), enables a better understanding of the information contained in the dataset and it helps to store the different measured parameters in a more organized way. It classifies the whole metadata in the dataset into two parts each of which contains different classes. The central metadata includes common information about countries, categories for each appliance type, appliance's properties and general knowledge about their variables (power, duration, correlation with other appliances). The second part of the metadata, contains the number of buildings, a list of all the appliances for each house with the respective specifications and number of components, the electricity meters (whole-house meter or individual meter) and a map describing which appliance is connected to which specific meter [34][39].

Once the data has been converted and loaded into NILMTK, the following step of exploring the data is of fundamental importance to understand the characteristics of the dataset and the kind of data that was measured. For this reason, the toolkit provides some diagnostic and statistics functions to observe for example the rates of dropout samples, gaps in measurements or proportions of the energy sub-metered and fraction of energy consumption per appliance [30].

The first analysed dataset was REDD which contains data for six different homes in the USA in 2011. Despite the fact that the dataset only provides several weeks of power data at the frequency of 1 Hz, it includes both active power at the meter-device level and apparent power for the whole-house measurements. The disaggregation was only performed on one house (building 1) which is composed of two site meters describing the total electricity consumption per circuit, and eighteen monitored devices. Some of those appliances recur more than once, such as the sockets and the lights which are measured in different circuits of the house. In *APPENDIX I* a graph representing several typical days of electricity consumption for the REDD dataset is reported.

The second dataset, the DRED, is constituted of five months data of a single household in the Netherlands in 2015. The sampling frequency of the data is 1 Hz, the apparent power is measured both at the appliance level and at the site meter, and there is no measurement of active power. Despite the longer range of measurements, the dropout rate of the measured values is very low, and the structure of the data is simpler than that of the REDD dataset. It is composed of 12 monitored devices and one site meter for the overall consumption. Also, it is important to mention that the file converted in HDF5 format does not include extra occupancy and ambient information that were collected for the LocED framework to allow a better comparison between the two selected datasets [28].

In *Figure* 3.1 the electricity consumption for three typical days in July 2015 from the DRED dataset is shown. In blue, the site meter measures the consumption of the whole house which is discernibly composed by the sum of the other devices. Besides others, it can be worth notice the fridge, visualized by the yellow line, which is constantly on. Some other devices such as the cooker, the microwave and the electric oven instead, turn on occasionally during the days.

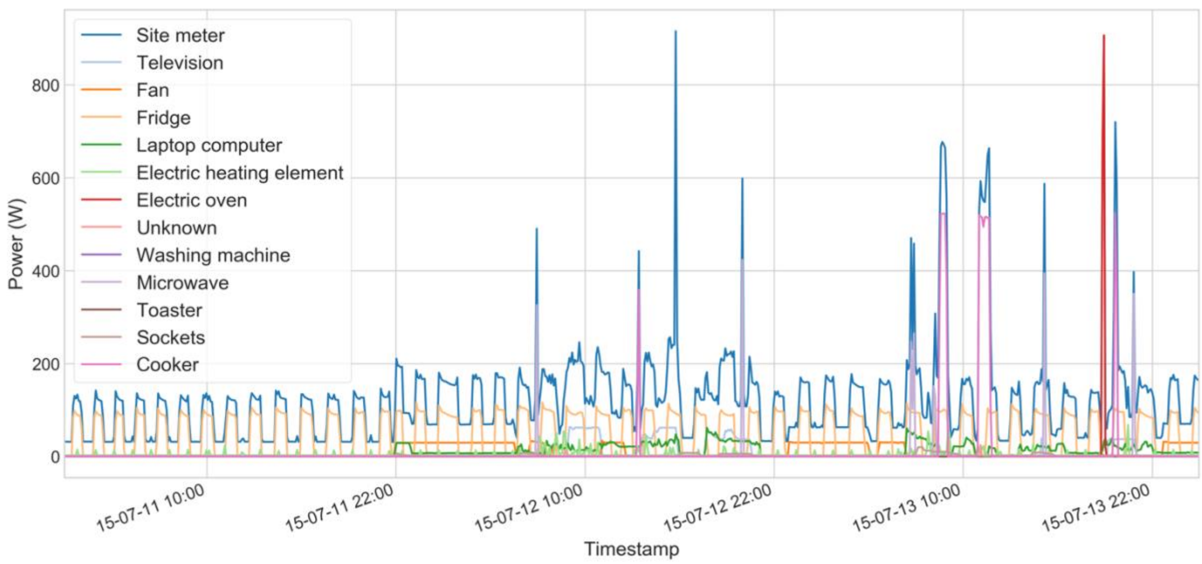


Figure 3.1 DRED whole-house (site meter) and individual appliances' consumption for three days during July 2015

Finally, before performing the disaggregation, the acquired data must be pre-processed. This phase consists of moderating some issues that could emerge from the diversity of the datasets. Some examples are down sampling the variation in the frequency into a specified one, filtering missing data, voltage normalization due to the collection in different countries and the reduction to the "top-k" appliances. This method of only selecting a k-number of most consuming appliances in the dataset, usually five, is a good method to simplify the disaggregation model while at the same time obtaining just the most valuable devices features, considering the remaining appliances as noise [30]. From performing this analysis to both the datasets, it is immediately possible to identify that the five most energy consuming appliances in a typical Dutch house are completely different from the ones of an American household. In fact, while in the Netherlands the fridge appears to consume more than 40% of the total electricity consumption followed by the electric heating and other cooking appliances such as cooker, microwave and electric oven, in the US the fridge constitutes only 19% of the total consumption less than the consumption from lights, sockets, and the washer dryer and the dishwasher together.

3.2.2 Load Identification

The NILMTK has implemented four benchmark disaggregation algorithms: Combinatorial Optimization (CO), Factorial Hidden Markov Model (FHMM), an implemented version of the first George Hart's disaggregation algorithm from 1985, and Maximum Likelihood Estimation (MLE). All these belong to the supervised learning category in which training with previously labelled data is necessary to perform disaggregation. The main reason why the authors decided to include these common and well-known algorithms and not new state-of-the-art disaggregation methods is to provide an established approach to enable a uniform comparison for novel algorithms' results [30].

At the same time the NILMTK-contrib, a repository for disaggregation algorithms currently in development, gives the possibility of accessing and contributing to newly proposed models while NILMTK-eval is another framework used to evaluate the performance of those algorithms in different scenarios. Two others third-party NILM algorithms that work with the toolkit are also proposed in GitHub [31].

In this study, the two firstly introduced algorithms on NILMTK were selected: CO and FHMM. The labelled data is represented by the individual consumption of each device *n* at a specific time *t* during the whole dataset duration. The percentage of data that is used as training data was varied to perform a sensitivity analysis on the different results and algorithms. A brief explanation of the logic behind the two chosen algorithms is provided in the following subchapters.

3.2.2.1 Combinatorial Optimization (CO)

This optimization algorithm tries to estimate the state of each appliance, minimizing the difference between the household total power demand measured at a specific time and the summatory of all the appliances' estimated power demand at the same time. The logic of CO is represented in the following *Equation (7)* [30]:

$$\hat{x}_t^{(n)} = argmin_{\hat{x}_t^{(n)}} \left| \bar{y}_t - \sum_{n=1}^{N} \hat{y}_t^{(n)} \right| \qquad (7)$$

Where $\hat{x}_t^{(n)}$ is the estimated state of the *n* appliance at time *t*, $\bar{y}_t$ is the aggregated actual power demand of the household at time *t*, and $\hat{y}_t^{(n)}$ is the estimated power demand of the *n* appliance at time *t.*

The CO finds the best combination of these appliances' states selecting the ones closer to the aggregate signal. Before that though, as shown by the *Equation (7*), every time slice and consequently every state is assumed to be an independent cluster, increasing the possibility of error propagation.

As anticipated in chapter 2.2.4.1, the complexity of the optimization algorithms increases exponentially the more devices are included in the computation. Thus, it is more efficient to identify a smaller number of appliances to perform the disaggregation.

3.2.2.2 Factorial Hidden Markov Model (FHMM)

The Hidden Markov Model (HMM) and its variations are the most commonly used methods for NILM algorithms. The approach is based on the Markov Model according to which a number of finite states of a process can be described by a final state, or output. This means that the output values are strictly dependent on the states of the process at the same instant of time. In *Figure* 3.2*a,* the state of the process at a specific time is called s_t and it is unknown, thus hidden state. The observation O_t is the actual output of the system and it is known. If this is applied to energy disaggregation, the observation would be the actual aggregated power measured at time *t* and s_t would be the anonymous state of an appliance. Since the total consumption in every instant is described as the change of states of many different devices in the household and not only one, a combined model of many HMM is necessary to perform NILM. The Factorial Hidden Markov Model (FHMM) (*Figure* 3.2*b*) is composed of k-HMMs, where k is the number of appliances that evolve simultaneously and independently over time. Each output corresponds to a different combination of states of every single appliance, calculated by the Viterbi algorithm to find the most likely sequences of states [40][41].

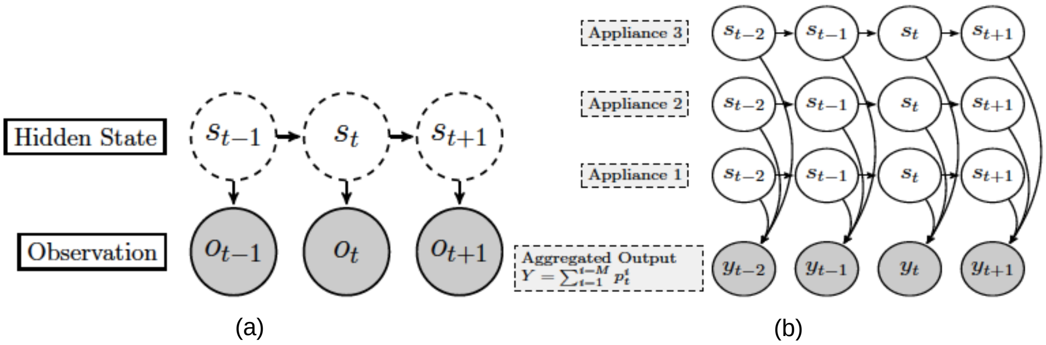


Figure 3.2 The Hidden Markov Model HMM (a) and the combined Factorial model HMM (b) [41]

As the number of introduced parameters increases, FHMM is more complex than CO because of the additional state transition and temporal information that were missing in the previous model. In order to make the algorithm work more efficiently, also in this case, the top-k appliances pre-processing data approach is proposed and recommended. Furthermore, the HMM-based approach performs better with controlled multi-states devices rather than variable devices [30].

3.2.3 Results and Performance Metrics

The final part of the NILM process involves analysing the obtained results, evaluating the performances of the algorithms using different metrics and getting an estimation of the energy consumed by the appliances. Due to the high computational capacity of the selected algorithms, the disaggregation was only performed for the five most consuming appliances of each dataset.

Besides, particular attention was given to the training phase. A significant part of the data needs to be used for training the algorithm while the remaining part is used as a validation of the results. For the REDD dataset, which is composed of only a few weeks of data, the NILMTK proposed to train just 32% of the entire dataset. At first, the same percentage was also applied for training the DRED dataset, however it appeared to perform better when 42% of the data was set as training. Furthermore, this latter option outperformed even higher percentages of trained data respectively 66% and 80%, and for this reason the results shown in this chapter for the DRED dataset are the ones obtained with 42% of data used for training. The REDD results used for the evaluation metrics are the ones obtained for the 32% of training data. The remaining results obtained for the DRED dataset and the graphs showing the results of the REDD dataset at both 42% and 32% of training are reported in *APPENDIX I*.

In *Figure* 3.3 the disaggregation of the fridge, being the most consuming appliance in the DRED dataset is shown for one typical day of September. The upper graph represents the ground truth (GT), which is the actual consumption profile of the fridge. The other two graphs are respectively the results for the CO and FHMM algorithms. As it can be observed, the FHMM performs better than the former one, being able to recognize, despite some noises, all the events. The CO instead, clustering different groups of appliances' states, results in more disturbance on the lower and most frequent powers, not returning a very sharp fridge profile.

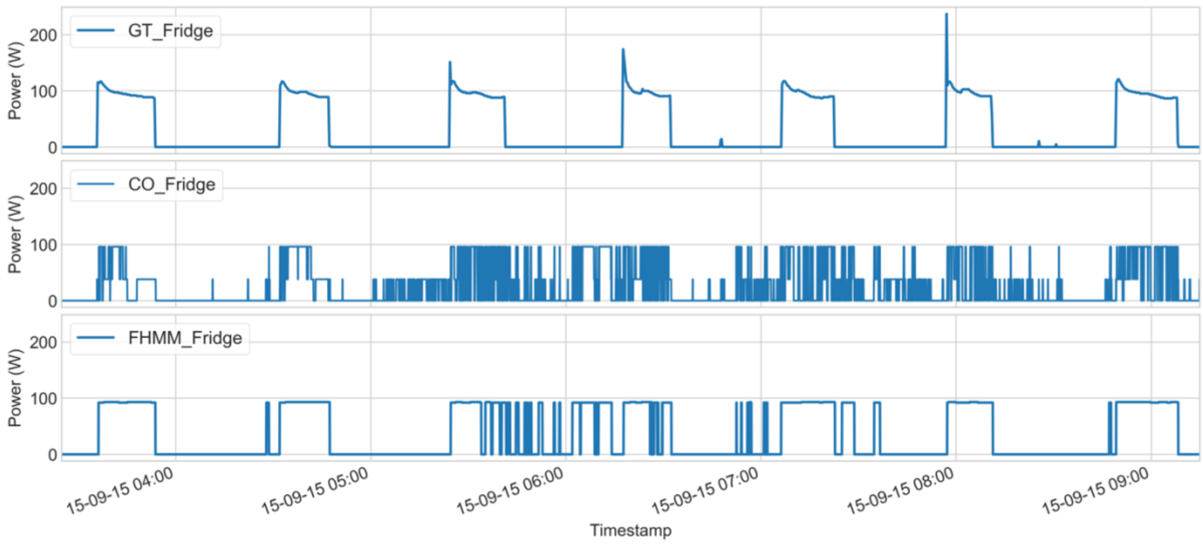


Figure 3.3 DRED. Fridge disaggregation

A similar behaviour is observed with the results of the aggregated five most consuming appliances in *Figure* 3.4. The FHMM in the lower graph has more defined profiles for the less consuming appliances compared to the CO which present a lot of noises around the device's signatures. However, for the cooker at high power, the CO algorithms provide a more accurate prediction.

Figure 3.4 DRED: The five most consuming appliances' disaggregation

3.2.3.1 Performance Metrics

As described in chapter 2.3.3, the evaluation of results is a very challenging phase of NILM because there is no uniform way to measure and compare the accuracy of the disaggregation algorithms. The most common approach, already included in the NILMTK, is to use the root mean square error (RMSE) expressed by *Equation (5)*. This metric represents how convergent the predicted values are to the actual measurements. The lower this error, the more accurate the estimated values will be.

Figure 3.5 shows the RMSEs obtained for the DRED and REDD datasets performing the CO and FHMM disaggregation algorithms for the five most consuming appliances. It is worth noticing that the major devices in the two datasets are not the same, except for the fridge and the microwave.

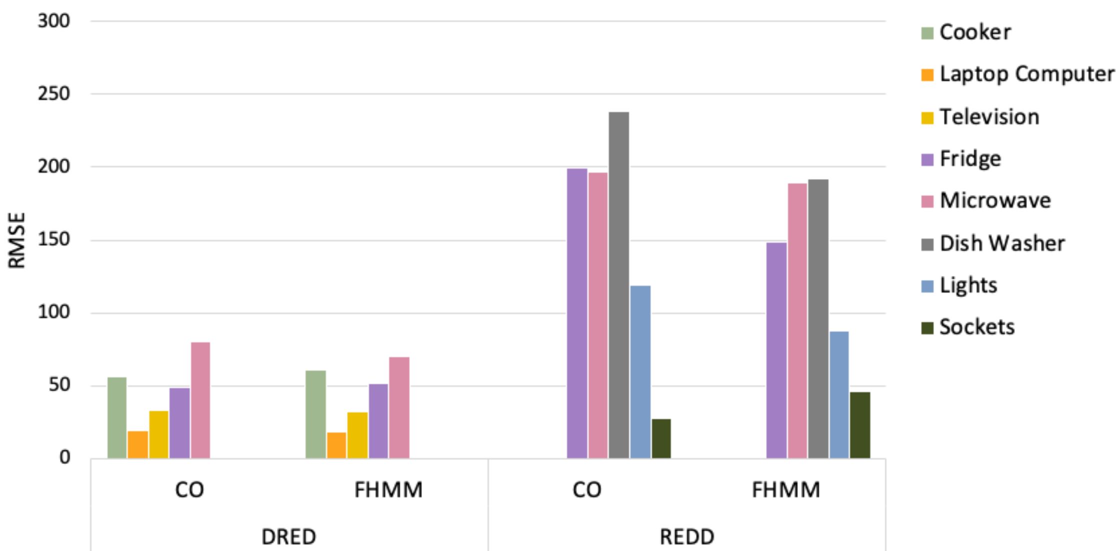


Figure 3.5 RMSE for the energy disaggregation of the top-5 appliances of the DRED and REDD datasets

Overall, the REDD dataset performed worse than the Dutch dataset for both the algorithms, likely caused by the shorter time period of the dataset. Moreover, it is interesting to notice that the most intermittent appliances such as television, computer and sockets received the lowest errors. This is due to their irregular but recurrent usage which allowed the algorithms to averagely detect their mean consumption. Finally, it is important to clarify that the RMSE is not a normalized measurement. Thus, the comparison between the results for different appliances is influenced by this factor.

In *Table* 3.1 two other evaluation metrics calculated for the two datasets are presented. The F_{score} is given by the average of the F_{score} for each of the top-5 appliances and calculated according to *Equation (4).* The fraction of energy assigned correctly is computed as the percentage of the predicted energy that was accurately recognized over the total actual consumption.

Table 3.1 F_{score} and fraction of energy assigned correctly metrics results for DRED and REDD datasets

	DRED		REDD	
	CO	**FHMM**	**CO**	**FHMM**
F_{score}	0,172	0,250	0,275	0,326
Fraction of energy assigned correctly (%)	66,1	70,0	71,5	88,6

The evaluations obtained from these two metrics alter the RMSE results in *Figure* 3.5. The REDD dataset, which received the worst convergence of estimated values around the true values, is the one with the highest percentage of energy assigned correctly and with the highest F_{score}. This means that even though the predictions of the DRED dataset are more concentrated around the true values, less true positive events are actually recognized, as pointed out by the low F_{score}. On the opposite, the REDD dataset correctly recognized a high number of events, including a higher number of true positives (higher F_{score}).

Finally, the FHMM resulted to be the algorithm with the better performances for all the metrics considered in this chapter. The results for CO are likely worsened by the lower power consumption appliances which are most prone to be mismatched during the optimization process. Besides, the FHMM algorithm includes temporal information of the state transitions which are not considered for the CO algorithm.

3.3 Signature Recognition Algorithm

This event-based approach is intended to perform disaggregation with low-frequency data (0,1 Hz) in order to recognize the basic appliances in a household by employing different event detection methods. The signature recognition approach consists of defining the main characteristics for each of those devices' consumption profiles and evaluating their changes over time. For this method, experimental measurements were carried out in several buildings in the Netherlands that provided a dataset with the sampling resolution of 10 seconds data.

The method has been tested on typical Dutch houses with PV production which are not part of the renovated zero-on-the-meter houses of the *Energiesprong* project. These houses were not selected due to their low-frequency data availability of only five minutes resolution, not enough to perform an accurate energy disaggregation. However, the solution is intended to be scalable to this type of households to provide the customers with improved and more tailored services.

In this chapter, the system design, the NILM process, and a specific case study for the fridge detection using the signature recognition approach are presented.

3.3.1 System Design

The hardware necessary for this system is very simple and requires low investments. It consists of a smart meter already installed in the households, a router with wi-fi connection to collect the data and to send them to the server and a Raspberry PI with an SD card.

The Raspberry PI is a small single-board computer that does not include external accessories such as a keyboard, a case or a monitor [42]. Due to its low-cost, portability and functionality, it is a very good solution to easily obtain good resolution data from the smart meter, saving these data in the SD card, and send them to Influx DB [43]. The Influx DB is an open-source time-series database designed to handle a lot of query loads that in this case are represented by the timestamp received every 10 seconds and the loads and demand information necessary to execute the algorithm which has been coded in Python language. Consequently, both live data and disaggregated data are visualized with Grafana dashboards to monitor and analyse the results of the process. The schema proposed in the following *Figure* 3.6 shows the entire procedure of the approach.

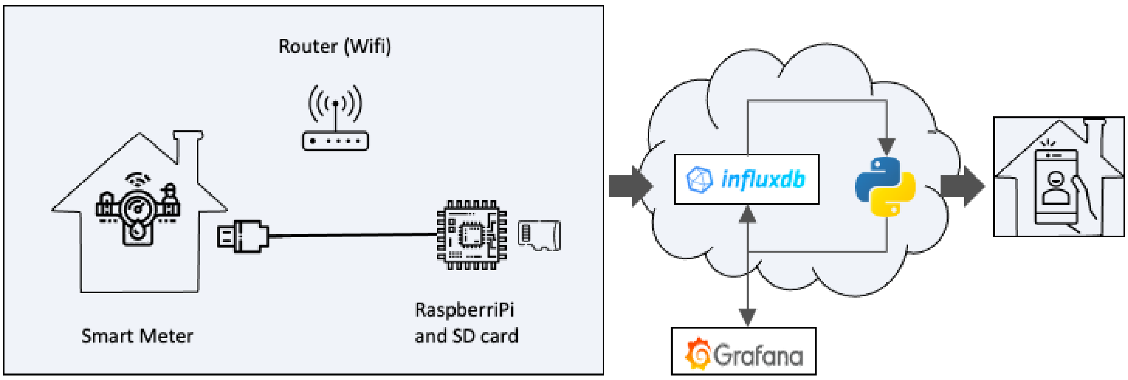


Figure 3.6 Diagram of System Design

3.3.2 Data Acquisition

The data has been collected with 10 seconds resolution of active power (low-frequency sampling rate of 0,1 Hz), from standard detached houses in the Netherlands used for a four people family. The houses have PV production and three phase powers system. The data acquisition is classified as circuit-level data which implies the collection of power measurements from smaller circuits inside the household. This method simplifies the energy disaggregation since fewer appliances are connected to the circuit in contrast to the overall electricity consumption received from the one whole-house circuit [10]. In the specific, one of the houses has been considered as a reference for the development of the algorithm due to the copious availability of information regarding the appliance's specifications and the devices' connection to each circuit-level of the household. However, the measurements for the individual consumption of each device are not available. The data has been collected since April 2020 at a non-constant rate due to an unstable connection. Nevertheless, the signal became stable from August 2020. Also, some previous databases of the same house from 2019 were used as a first experimental data to perform the energy disaggregation algorithm.

As an example, the total electricity consumption of the house during a Saturday of September 2020 is shown in *Figure* 3.7. The negative power production in *phase I* (green line) represents the connection to the solar panels, while *phase II* (yellow line) is coupled to the main appliances such as dishwasher, washing machine, oven, boiler for hot water, etc. *Phase III* (blue line) is connected with the fridge, the clothes dryer (detectable in *Figure* 3.7 between 12.45 pm and 1.30 pm), several plugs and some lights.

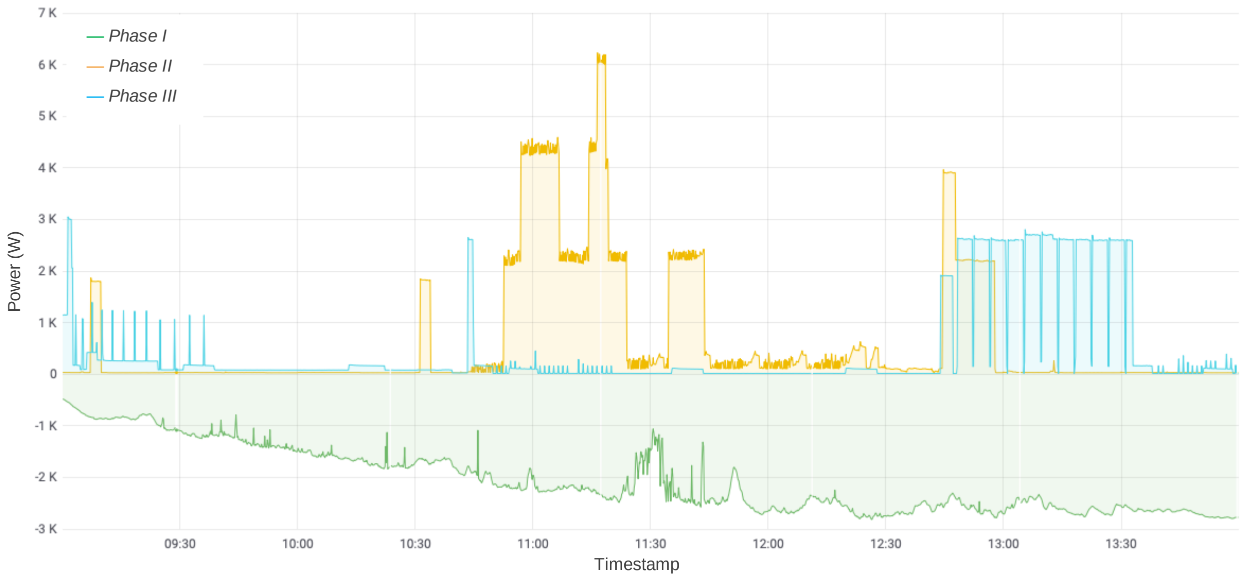


Figure 3.7 Total electricity consumption of a typical house in the Netherlands on a Saturday in September 2020. The consumption is shown according to the three phases of the electrical system of the house (phase I, phase II and phase III). Obtained from Grafana.

3.3.3 Event Detection: Fridge

The case study for the signature recognition approach was selected to be the fridge (with a freezer included). In fact, it belongs both to type IV devices which are always on throughout the day, and to the continuously variable devices (type III). In the latter case, the consumption varies either when the fridge's door is opened by the consumer, letting the cold air exit the device, or according to the number of items contained in the fridge itself. It is important to know that the refrigerator works by using gas. This gas is compressed by a pump, cooled down and transported in narrow pipes through the sides of the fridge to cool its interior. When the gas becomes warm enough, it re-enters the compressor and repeats the cycle to set the internal temperature of the fridge again.

During the night, the energy consumption of the fridge is undisturbed by the consumer's actions, and therefore its signature is more easily detectable. *Figure* 3.8 shows a typical aggregate load during the night where the consumption is only comprised of the baseload and the type IV devices which in this case are the fridge in blue (*phase III)* and the water boiler in orange *(phase II)*. The baseload is represented by other constantly active appliances in both the phases.

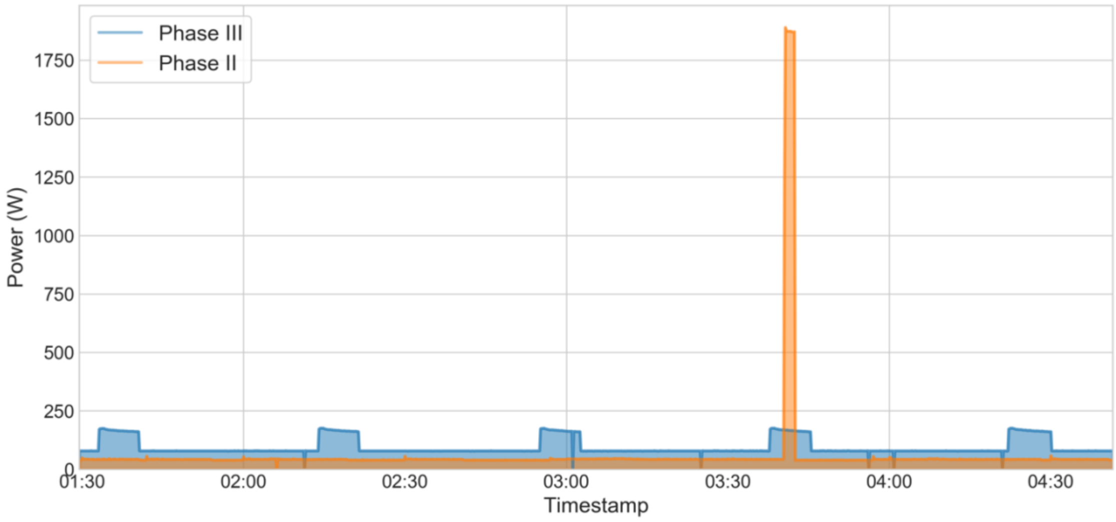


Figure 3.8 Total electricity consumption during the night of the measured house in the Netherlands: Fridge and Boiler loads. The consumption is shown for two of the three phases of the electrical system in the house (phase II and phase III)

3.3.4 Feature Extraction: Fridge

The characteristics of a load signature are here referred as those periodical features which are representative of the operational profile of a device and therefore allow its unambiguous recognition.

For simplification, as it can be observed from *Figure* 3.9, some letters have been attributed to the typical characteristics of the fridge. The symbols which directly refer to the active power from *phase III* are *B*, *L* and *G*. They represent respectively the baseload (*B*) (power always on in *phase III*), the net load (*L*) (maximum active power during the fridge cycle minus the baseload) and the gap (*G*) in power consumption that the fridge shows in every cycle between its maximum power and the instant just before decreasing its power to the minimum.

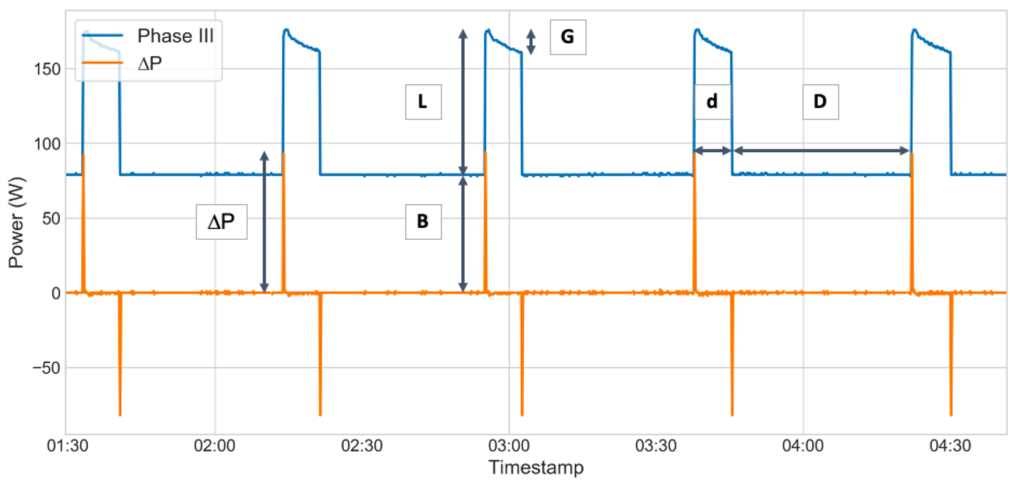


Figure 3.9 Typical fridge load during the night and representation of its main characteristics: ΔP, L, B, G, d, D

The most important feature is the ΔP or load differential which indicates the positive or negative changes in the active power, hence if the device is changing its state (for example increasing or decreasing the power consumption or completely turning on or off). The power change observed in *Figure* 3.9 (orange line), can be expressed with the following *Equation (8*) [44]:

$$\Delta P = P_{t2} - P_{t1} \qquad (8)$$

Where P_{t1} is the steady-state active power at time *t1* and P_{t2} is the steady-state active power at time *t2*. In fact, every appliance has its range of power in which it operates. For the fridge, it depends whether or not it operates with the freezer and it usually ranges from 200 to 400 W (fridge plus freezer) [45]. The highest peak demand can be attributed to the defrosting phase, while the mean consumption of a fridge can be considered around 200 W.

Furthermore, besides the characteristics related to the power consumption, it is also very important to consider the fridge cycle in terms of time and frequency. In *Figure* 3.9, *d* is representing the time when the compressor is in operation to cool down the fridge environment, while *D* is the remaining time before the next operative state. It is possible to associate these two values to the concept of Duty Cycle, or more simply how long the device is on (*d*) within a full cycle, expressed by the following *Equation (9*) [46]:

$$Duty\ Cycle\ (\%) = \frac{d}{d + D} \qquad (9)$$

From an efficiency point of view, the aim is to keep the Duty Cycle value as low as possible, which in practice translates into avoiding opening the fridge's door often and for a long period of time (smaller *d* and larger *D*). In this case study, the Duty Cycle was found to fluctuate between around 17% and 36%.

During the day though, due to its continuous usage, the fridge's signature is not as regular as throughout the night. In *Figure* 3.10 an experiment is shown where the fridge's door has been kept open for ten minutes between 15.40 and 15.50 in a September's afternoon.

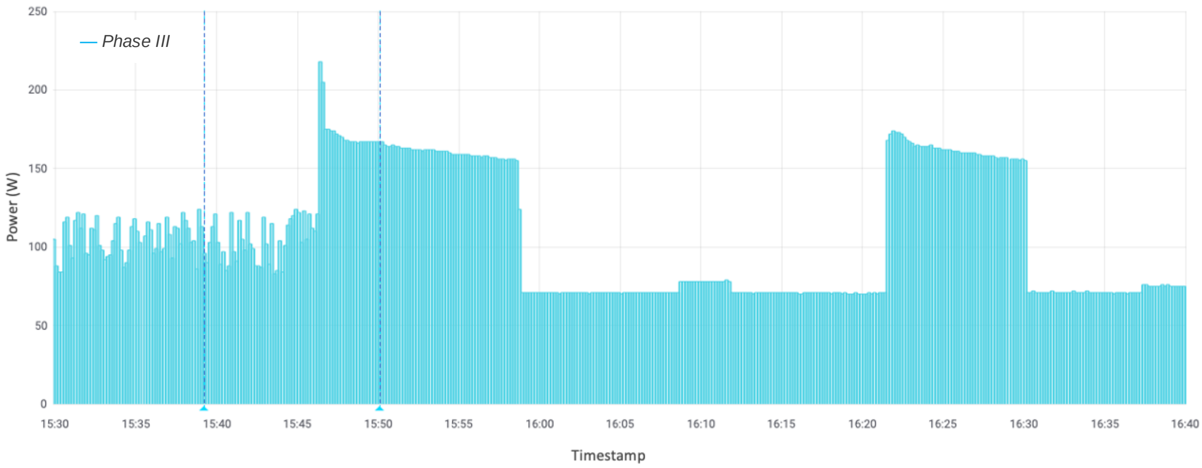


Figure 3.10 Fridge's door kept open for 10 minutes between 15.40 and 15.50 in a September's afternoon. The consumption is shown for phase III of the electrical system. Obtained from Grafana.

Observing the ΔP profile of the fridge and neglecting the noise of the measurements before the event, most likely caused by the laptop computer, the signature of the fridge is still visible and clearly discernible. Comparing the several characteristics just presented in *Figure* 3.9, it can be noticed that the parameters are still evident. The profile presents a ΔP with a sharp increase (*L*), the gap in power (*G*) between the maximum power and the mean power during the on cycle (*d*), in this case of around 1,03 W increase per minute, and finally a negative ΔP. This translates into a noticeable increase of that specific Duty Cycle up to 36% compared to the average of 22% of the cycles of the same day.

From this analysis, it can be inferred that the characteristics most affected by the door opening event are a gap of power *G* during the positive ΔP, and the duration of the cooling cycle *d*. For this reason, their variation over a certain range percentage can be considered as the trigger to detect when the fridge's door was kept open for a long time or for the recognition of an inefficient device with a very high and constant Duty Cycle.

3.3.5 Load Identification: Fridge

The signature recognition method that was developed in this chapter does not use any learning algorithm. Instead, it was inspired by some expert heuristics' methods for event detection, as described in chapter 2.2.2, in which some threshold ranges of power variation are pre-defined for a specific appliance. In a similar way, the fridge's signature was identified after the initialization of its main features as it can be seen in *Figure* 3.11*. Phase III* represents the total electricity consumption on the circuit and the orange line is the predicted load of the fridge throughout that day. Despite the missing availability of a measurement that could precisely show the consumption of this device, it is evident that the prediction of the load signature was rather accurate for the detection of the fridge.

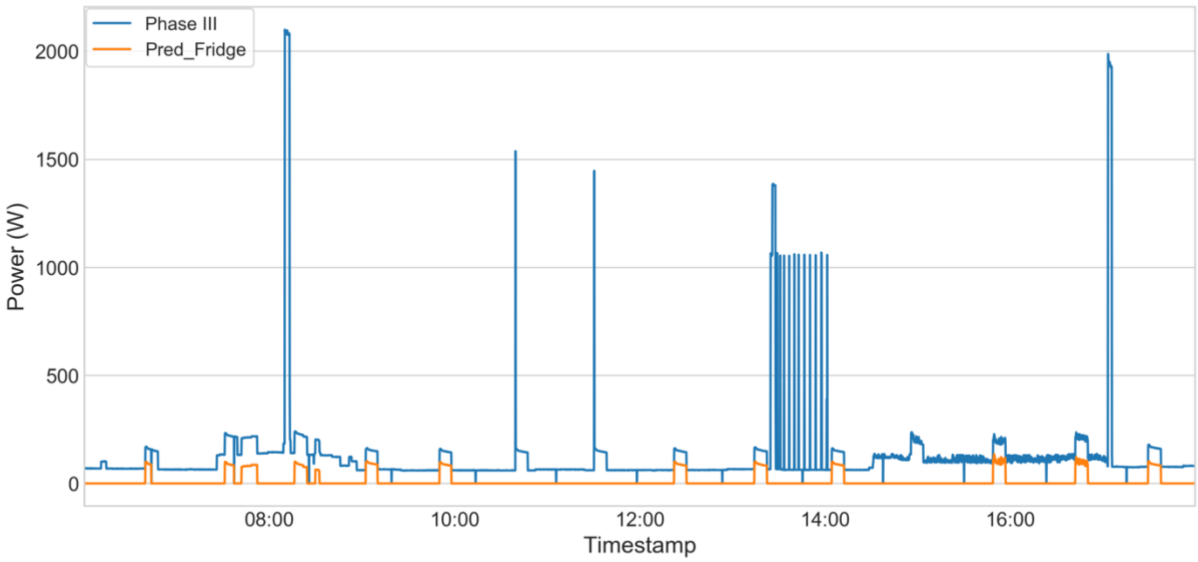


Figure 3.11 Fridge disaggregation during a typical day in September

3.3.6 Results and Performance Metrics

The algorithm developed is able to recognize the occurring ΔPs, to detect an event when it happens and to classify it according to the power differences. However, this classification does not recognize all the different appliances since the algorithm was modelled over the specific characteristics of the fridge.

In *Figure* 3.12*,* the disaggregation of the power of the different loads (lower graph) is shown in comparison to the aggregate consumption of *phase III* (upper graph) during the week between the 28th of September and the 4th of October 2020.

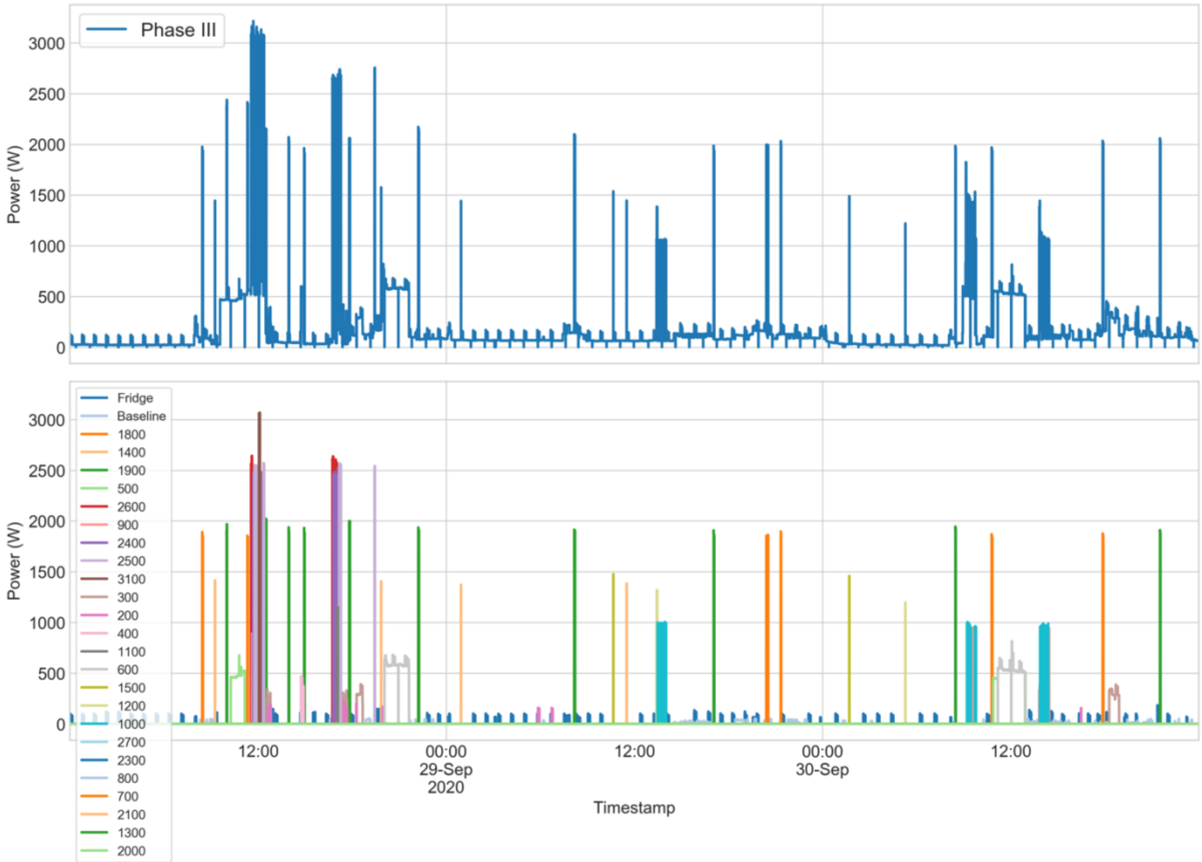


Figure 3.12 Signature Recognition approach for a week in September: aggregated consumption for phase III and loads power disaggregation

In the lower graph, every different colour is representative of a change in ΔP of the same power amount indicated by the legend. Since the fridge is the only load that has been studied, with its characteristics and ranges of ΔP defined, it is the only profile that was recognized and classified as Fridge (dark blue line) together with the baseline. In the same way, attributing a range of typical characteristics to other appliances such as the boiler or the coffee machine, the code could be able to classify also other devices.

It is also worth mentioning that since this method does not use a learning algorithm, the code developed does not learn by itself about the different appliances and does not improve its performances with the larger availability of data. Thus, its accuracy decreases the larger the provided dataset is, due to the higher probability of error. Also, every appliance has to be initialized with a specific number of variables within certain ranges (power and time) to be recognized. The multi-state and continuously variable appliances represent a limitation for this algorithm due to the need of tracking the device state variations over time.

However, for the case study of the fridge, the signature recognition algorithm proved to be able to detect unusual behaviours and could be used as a trigger to alert the inhabitants of possible misbehaviours such as a door which is not perfectly closed, and which would alter the occurring Duty Cycles.

3.3.6.1 Performance Metrics

Due to the unavailability of the individual consumption data for each appliance, especially for the fridge, most of the common metrics for event-based approaches described in chapter 2.3.3 cannot be used in this analysis because there is no possibility of comparison with a ground truth.

Despite this, it was possible to calculate the RMSE (*Equation (4))* comparing the sum of the ΔPs obtained with the aggregated consumption in *phase III*. *Figure* 3.13 shows the sum of all the recognized loads compared to the actual electricity consumed in *phase III* and the error profile given by the difference between the former two profiles in the upper part of the graph (average error 99,6 W per timestamp).

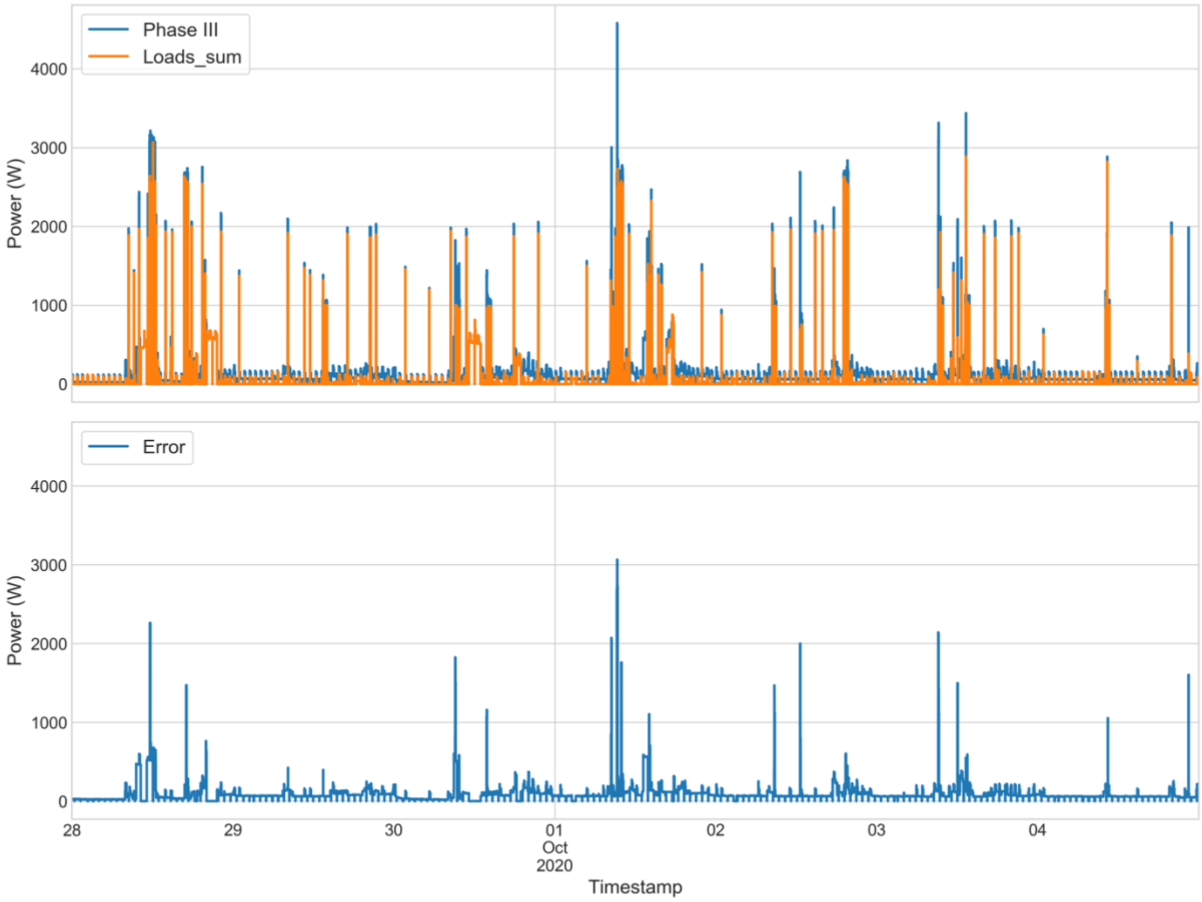


Figure 3.13 Comparison between the actual aggregated consumption (Phase III) and the sum of the predicted loads and the error given by the difference of the two

The RMSE calculated for the overall accuracy of the loads recognition is 180, a value which is comparable to the results obtained using the NILMTK approach for the REDD dataset with both CO and FHMM algorithms presented in *Figure* 3.5*.* As previously explained in chapter 2.3.3, this value alone cannot be perfectly representative for the accuracy of this methodology since it is a not normalized measure which does not indicate the number of true positive recognitions.

3.4 Challenges and Limitations

Despite having been studied for decades, energy disaggregation still shows limited implementation due to a series of challenges that need to be overcome to allow a broader employment of this technology. In this chapter, the main limitations of NILM and the challenges encountered for the development of this thesis will be further analysed and classified into three categories: data acquisition, load identification and performance evaluation.

3.4.1 Data acquisition

As anticipated in chapter 2.3.1, the creation of a dataset is a very complex and time-consuming process which makes of data acquisition one of the most problematic steps of energy disaggregation. The quality and the kind of data used to perform the algorithms, are not only influencing the results of the analysis but the whole NILM process and the evaluation phase. Depending on the sampling frequency of the collected data, the type of power measurements obtained, the timeframe, the location, or if the whole-house consumption and the individual loads data are both collected or not, there are a multitude of different possibilities to perform disaggregation.

At the data collection phase, the noises, network malfunctions, data compression and the difference in the type of power meters installed in each household can generate mismatching data, reduced accuracy and loss of raw data. Missing data are also to be intended as those measurements from single loads or sub-metered data which are not possible to obtain. This is a key issue for the NILM technology. As a matter of fact, some devices cannot be monitored using plugs at the meter level because they are not always connected to the same plug such as the vacuum cleaners or others that do not even have a plug like ceiling lights. One answer could be to use more often the circuit level data in which those appliances are included. However, if the aim is to monitor the consumption of specific loads this won’t be an effective solution [21].

The same problem was encountered in this thesis for the signature approach method where the individual loads were not available. In this case, knowing the circuit level data and the breakdown of the appliances connected to each of the three phases was a great simplification which allowed the recognition of the fridge load. Another drawback was the impossibility of using the data already available within the Watch-E server because of its too low resolution (five minutes frequency).

At the hardware level, the data transmission is mainly composed of wireless connections and needs to avoid data losses and guaranteeing security for privacy issues. According to the type of product, the installation of the auxiliary hardware at the smart meter level, most often plug-in, could require the assistance of a specialized operator. The system implemented for the signature approach was of easy installation, but it took many months to obtain a stable connection that allowed a continuous data transmission from the house.

3.4.2 Load Identification

Once data is collected, the challenge lies on the load identification strategies to disaggregate the total electricity consumption. In chapter 2.2.4, the difficulty of recognizing specific loads for particular situations was described. For example, this happens when there are many devices of the same kind in one household (most commonly fridges and televisions), or when switch events occur concurrently. Besides, when different types of appliances, in particular low-power ones, simultaneously draw similar power, an overlap of load features occurs which makes them difficult to be discerned. Another problem that must be faced is related to the diverse appliances' categories, especially for multi-state and continuously variable devices for which the precision of load disaggregation is still not satisfactory. In fact, there is not a uniform load signature model for each appliance type due to the variances of their category, specifications and manufacturer, thus affecting the performance of the algorithms and making the signatures incompatible [21].

In case of supervised learning approaches, the labelling of training data requires human effort and intervention. In order to create a tailored signature dataset to train the algorithms, every appliance has to be manually switched on or off to be identified creating a time-consuming process and fallible results. At the same time, off-line training of the algorithms from a generalized signature dataset would be limited by the devices used as samples because of the impossibility to include all the model characteristics of every appliance inside this dataset. Therefore, the NILM approach has scalability problems since with off-line training not all the appliances in a household can be recognized [10][20].

The signature algorithm proposed in this thesis did not include a learning approach. While this simplified the code and facilitated its replicability, it also limited the potential of the results. The impossibility of improving the recognition with time and the manual labelling of the specific characteristics for every device are some of the consequences. Moreover, a precise classification of multi-state and continuously variable appliances is not possible.

3.4.3 Performance Evaluation

The performance of the NILM algorithms is strictly dependent on external factors. The presence of standardized frameworks that promote unified interfaces, data formats, and performance metrics influences the possibility of comparison, analysis and further research of these algorithms.

If the introduction of the NILMTK, as presented in chapter 2.3, helped to evaluate energy disaggregation algorithms in a more conventional manner, there is still a long way before a safe growth environment for this technology can be ensured. The lack of a clear understanding on how traditional metrics behave with NILM and the unbalanced comparison and incompatibility between results obtained using different sets of metrics, algorithms and datasets, constitute strong limitations. Moreover, there are no specific metrics tailored for this application. One solution could be defining new metrics based on the concept of cost-sensitive learning where the wrong event detections from less consuming appliances is not as relevant as the same mistakes done for a larger consumption device. Another solution could be the development of an online platform to evaluate the performances of NILM algorithms, guaranteeing that the different approaches are assessed under the same conditions [20].

The lack of individual loads in the signature recognition approach prevented the obtained results to be adequately compared with the NILMTK results since most of the performance metrics adopted are based on the comparison of the disaggregated loads with the ground truth measurements of the single devices. Furthermore, it was also not possible to implement the database that has been obtained from the Dutch household within the toolkit due to a mismatch with the data format.

4. Economic and Social Impact

In this section of the thesis, energy disaggregation is analysed from an economic and social perspective. After the technical analysis of NILM which emphasised both the advantages and the limitations of the technology, this chapter explains how energy disaggregation can be implemented and for which applications. The business case and the market potential of NILM are introduced through concrete examples of companies that have already developed this technology. Afterwards, the social impact is assessed describing the different type of feedbacks that can be provided to increase the consumers level of engagement and evaluating the energy savings obtainable from energy disaggregation.

4.1 Methodology

The socio-economic analysis was performed by consulting literature reviews and market reports. Additional insights were gathered by taking part in the 5th International Workshop on Non-Intrusive Load Monitoring (NILM) held in November of 2020 in an online format.

The literatures used can be classified into different sources of documentation:

- companies' white papers and webpages
- journals on products and customer satisfaction
- experimental papers on the efficiency of energy feedbacks
- systematic reviews on the energy feedback studies

The methodology also included the installation of a commercially available NILM product in the Watch-E office to evaluate its performance and validate the thesis's hypothesis on the efficiency of the feedbacks provided to the users.

Due to the early stages in the implementation of this technology, only little information is present regarding the potential business models that could be adopted to develop energy disaggregation whose market is considered to be part of the larger Home Energy Management Systems (HEMS) sector. A study was conducted to evaluate the impact of NILM in the American, European and Dutch markets considering the possible economic and energy savings.

There are many uncertainties regarding the percentage of obtainable savings from disaggregated energy feedbacks. A review of different studies was conducted to specify the different types of feedbacks that can be received by the consumers and their respective level of customer engagement. Thereby, the most common biases that usually influence the efficiency of the customers response to these feedbacks will be emphasized according to the obtained results of energy savings.

Finally, the NILM Workshop 2020 organized by a growing NILM research community brought together professors, researchers, practitioners and students in order to offer useful insights about the current state of the art of the technology, newly proposed applications and other innovations in the sector from all around the world.

4.2 Economic Impact

The evaluation of the potential economic impact of an emerging technology requires the analysis of different factors. For energy disaggregation there are many different applications, stakeholders and markets that need to be considered. This large variety is one of the main advantages of NILM since it allows the technology to be versatile in its employment.

In this section, the findings have been divided into four subchapters. First, the different NILM applications are distributed according to the possible customer segmentation. Thereafter, the product implementation is divided based on how the appliance specific data is acquired, and base on where the NILM algorithm is performed. Afterwards, the proposed business models are described providing examples of ongoing projects and companies already employing this technology. In the final part of the economic assessment, the market potential section will identify a context regarding costs and benefits of NILM according to specific geographic areas.

4.2.1 Customer segmentation and NILM applications

There are numerous potential use-cases for energy disaggregation and consequent benefits of appliance specific over whole-house data applications. These applications can be classified according to the two customer segments: electricity consumers (residential, commercial and industrial), and utilities and service companies. Currently, the NILM technology is mainly used in the residential sector while applications in the industrial and commercial sectors are still limited. This is due to the high disaggregation complexity caused by the presence of many appliances of the same type and by events switching simultaneously. Thus, in the industrial and commercial sector, the disaggregation mostly aims at solving fault detection problems. Therefore, in this thesis, the focus lays on residential consumers applications. Furthermore, energy disaggregation offers numerous other benefits to the policy makers, researchers and developer who share a common concern for energy saving and energy management.

The two customer segments of electricity consumers, and utilities and service companies will be hereunder further analysed followed by other potential NILM applications in the research and development (R&D) and in the policy making sectors.

4.2.1.1 Electricity Consumers

The first type of potential customers are tenants and homeowners who are using electricity without having a clear idea on which appliances are the most consuming ones and responsible of their high energy bills. These people can benefit from appliance specific data to optimize their energy usage and reduce costs [47].

The most commonly mentioned application for energy disaggregation is the possibility of offering these consumers itemized energy bills showing the individual appliances consumption. This would not only lead to a higher energy literacy but also to an increase in the awareness of the distribution of energy consumption per appliance. An additional service is to provide personalized energy saving recommendations on the kind of actions in the everyday life that will easily help reducing the electricity consumed based on the household usage profiling. One example is to compare the current energy tariff plan with the tariffs available in the market and propose either recommendations on the savings obtainable by shifting the usage of some appliances from times where the price is higher to off-peak times or changing energy suppliers and tariff plans [48].

Another employment is to detect anomalies in the electricity consumption. NILM is able to recognize if an appliance's profile is not correctly shaped, identifying thus malfunctioning or inefficient devices. The fridge case study in chapter 3.3.4 can serve as an example for this application. The higher duty cycle percentage indicates that the fridge's door is left open. In the same way, a constantly high duty cycle shows that the frosting cycle is unusual which would indicate that either the seal is damaged, or the appliance is old and inefficient. The possibility of detecting problems with device performances is a useful tool to provide the customers with cost-effective analysis to replace an inefficient appliance, to send push notifications when a device was left on unintentionally or is not shifting into stand-by mode.

Finally, energy disaggregation could be able to monitor the occupancy in a building. This a large research field for the assistance of elderly and people with health problems who could be remotely controlled by checking if their routines and typical behaviours are respected during the day. Alerts could be sent to allow faster aid in case of unforeseen illness or accidents, or reminders could be provided to the patients to perform specific activities such as taking medicines. Furthermore, the occupancy state could be useful to control and optimize smart heating and cooling schedules in a household. However, these applications lead to great privacy and security concerns that must be carefully assessed before a potential implementation.

4.2.1.2 Utilities and Service Companies

This customer segment refers to both public and private utility and service employments. The term utility refers to a company responsible for generation and/or distribution of energy, while a service company is the one delivering retrofits, maintenance or security services. The utilities receive regulatory incentives in order to provide cost-effective solutions to maintain the stability in the grid and regulate the peak consumption so it will not exceed the grid's capacity. Those tasks require coordinated actions which are difficult to obtain with their large and variegated number of energy consumers [47].

A NILM application with a high economic potential for utilities targets Demand Response (DR). This mechanism involves modifying the consumption of electrical loads (consumers demand) in response to the needs of the electricity grid, in particular for very congested grids and for times of high peak consumption [49]. For this purpose, appliances can be classified as non-deferrable appliances whose usage cannot be delayed in time such as fridge and lighting, and deferrable appliances whose operation can be shifted throughout the day such as washing machine, and heating/cooling systems. The deferrable devices can provide flexibility in the electricity demand which can be used to adjust it to the supply. The peak in the demand curve can be smoothed, reducing the total cost of the energy system, and helping the integration of intermittent renewable energy sources in the grid mix. In this case, utilities would use energy disaggregation to identify the total amount of deferrable loads operating in a defined moment and send notifications to the consumers asking them to postpone the usage of some devices during peak times [50].

Another advantage of energy disaggregation for grid operators could be increasing the accuracy of energy demand predictions by matching the individual appliances' profiles with demographic and meteorological information instead of using only historical consumption data. A more precise load forecast would improve the operational planning of power system infrastructures, allow a better analysis of the energy flow in electric networks and evaluation of energy purchasing, and increase the accuracy of energy models [48].

Moreover, utilities and service companies have to deal with the expensive issue of segmenting their diversified customers in order to personalize energy efficiency recommendations. Engaging clients with online platforms by showing them disaggregated energy information could help utilities to cheaply obtain more data to perform customer segmentation. Thereby, other means of gathering customer information, such as call centres and surveys could be avoided. At the same time, these companies would gain trust from the consumers who would feel more involved into the billing and energy consumption process [48]. Thus, utilities could target both individual consumers and whole communities with more specific recommendations and offers. In a similar way, service companies could more easily identify the needs of their customers and increase the relevance and efficiency of their services.

4.2.1.3 Other applications: R&D and Policy Making

Other segments that would benefit from energy disaggregation applications are found in the research and development (R&D), and in the governmental and policy sectors. Many product developers are willing to incorporate NILM technology into already available products or new products. In fact, many companies are entering this market trying to deliver innovative and enhanced features in the automotive, home energy management, security and comfort fields [47]. Differently, governments necessitate tools that easily and cheaply could allow a global understanding of the public behaviours before defining new and effective energy policies.

Manufactures and product developer are aiming to design advanced appliances and high-performance and high-quality equipment for the households. This would enhance the motivation towards the replacement of low efficiency devices and the adoption of new technologies for energy efficiency and pollutant reduction [51].

Furthermore, NILM could be of fundamental importance for building simulation models that aim at predicting energy usage to design energy efficient constructions (in particular for lighting, heating and cooling systems). Architects would benefit from the analysis of energy consumption behaviours in existing buildings for the design of future projects and to understand the discrepancies between predicted and actual energy consumption in a building. Being aware of the disaggregated energy consumption could help investigating different measures for improved energy efficiency, or recognizing inefficiencies in a commissioned building [49].

Another group that would benefit from NILM applications are governments which spend billions every year for energy efficiency programs [51]. These efficiency programs could be defined, evaluated and diversified based on disaggregated data which would offer more insights on their effectiveness. Policy makers could also benefit from energy disaggregation information to enhance a correct allocation of funds and other regulatory elements such as power market tariffs [52].

In conclusion, a greater amount of disaggregated data collected would help to analyse energy usage behaviour, to forecast the demand response mechanism potential, to better manage energy in grids, communities and buildings, and to have a greater understanding of electricity loads and their flexibility.

4.2.2 Product Implementation

There are several approaches to create a business model for the services offered by energy disaggregation. However, it is firstly important to understand the possible ways of implementing a NILM product. In this subchapter it will be clarified how the appliance specific data are obtained and where the NILM algorithms can be performed, considering the main resources that are needed before implementing the technology in the market.

4.2.2.1 Hardware and Software solutions: how to obtain the disaggregated data

Generally, the NILM technology can use two type of solutions to acquire the disaggregated energy consumption data [16][52]:

- Hardware solutions: these are sensor based and involve the development of both hardware and software. A specific device can be plugged to the smart meter to collect the required sampling frequencies in order to perform whole-home monitoring. These devices are usually composed of current clamps for accurate submetering, a power box which measures current and voltage and calculates power, and a gateway to establish communication and data storage. The disaggregation is then achieved either locally at the site location or in cloud services. This kind of solutions may lead to some technical problems during the installation and some compatibility issues when connected in different housing stocks. Besides, within the category of hardware solutions there are also ILM techniques due to their requirement of plugs installed for each individual device and smart appliances.

- Software-as-a-Service (SaaS) solutions: these are generally cloud based and only involve the development of the software. The whole-home electricity data must be already available at an adequate sampling frequency and can then be disaggregated through NILM algorithms. The computations can be performed either in the cloud where the data are collected or locally at the smart meter. In the latter option, the data could be directly processed inside the meter, allowing higher frequency samplings since the data doesn’t need to be transmitted, and avoiding extra hardware installation. Software solutions eliminate the hardware, installation, maintenance and licensing costs and thus show a high potential for market penetration. However, in order for these solutions to be implemented the gaps in the data requirements must be overcome. Currently, only very low frequency data is available (15 minutes to 1 minute), and a standard communication of data between different networks that the smart meters use is not provided yet.

Figure 4.1 illustrates the general components of an Internet of Things (IoT) technology which can be used to summarize the differences between hardware and SaaS solutions.

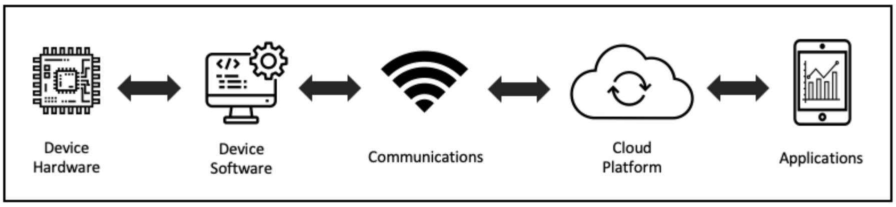


Figure 4.1 IoT technology general components

The hardware solutions require all the components shown in the graph, from the hardware devices to the user applications and interfaces to display the results of energy disaggregation. In contrast, the SaaS model requires solely the two components at the right side of *Figure* 4.1. The cloud platform is used for NILM computations while the applications available over the internet provide the customers with feedbacks and insights about their electricity consumption [53]. As previously explained, for both models it is also possible that the development of a cloud platform is not included, and the algorithms are implemented locally at the site.

Currently, hardware solutions are the most commonly adopted despite the higher complexity that their implementation requires. This is because smart meters that can record high frequency resolution data and that would foster SaaS implementation are very expensive and still not common in residential households. For this reason, hardware solutions are considered for an additional analysis to define where the NILM algorithm can be performed.

4.2.2.2 Hardware solutions: where to perform the NILM algorithm

Figure 4.2 represents the three possible options on where to perform the NILM algorithm and their dependency on the main networks in a household.

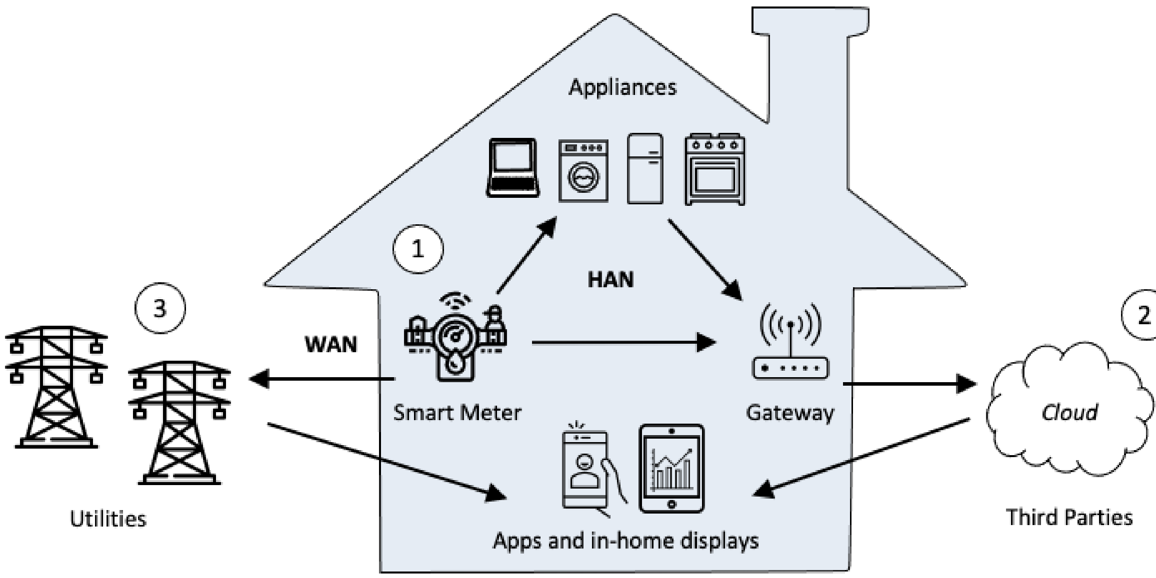


Figure 4.2 Three possible hardware solutions for NILM algorithms implementation

The Home Area Network (HAN), commonly used for HEMS, connects the devices inside the house to provide monitoring and controlling over the energy usage. It receives data at low frequencies (between 1 minute and 1 second) from the smart meter with which it communicates through the Zigbee communication protocol and from the gateway which propagates the Wi-Fi connection. The Wide Area Network (WAN) allows the smart meter to send data at lower frequencies (every 1 hour or 15 minutes) to the utilities in order to extract information and further processing the data. This network allows utilities to perform automated meter readings for billings and monitoring but could be also used for the demand response (DR) mechanism. In this case, the utility could send back the signal to the smart meter, connected to the HAN, to modify the operation of targeted appliances.

As it can be noticed from *Figure* 4.2, the first hardware solution could be implementing a NILM algorithm on the smart meter without the necessity of data transmission outside the household. This system would allow to handle higher sampling frequency (more than 1 kHz) and the data could be directly stored and processed inside the meter without the need of any additional hardware component. The HAN would provide the connection with the appliances and the in-home displays or smartphones to show the results of disaggregation. In this case, the smart meters should be directly designed and commercially distributed with the possibility of providing disaggregation. The second hardware option involves third party companies performing NILM algorithms in the cloud. The HAN would receive the results and deliver the feedbacks and appliance specific information to the consumer displays. Finally, energy disaggregation can be performed on the utility servers without any duplication of data on the cloud or on third parties' servers. However, this would constitute an additional burden for the utilities since they should deal themselves with the complexity of disaggregation algorithms [52].

4.2.3 Business Models

Finally, the NILM business models can be analysed and classified into two different categories according to the customer segment which the company delivering disaggregation is targeting, and to the selling strategy of the products [16]. In fact, a business model can include either selling a product directly to its users (Business–to–Consumer or B2C) or selling it to another business entity who will consequently retail the product to the users who in this case are the energy consumers (Business–to–Business–to–Customers or B2B2C). The business models identified for a NILM product are:

- B2C for energy Consumers (Residential, Commercial and Industrial): delivering a plug-and-play device. For these applications, the energy disaggregation company provides the hardware to be installed in the household or in the office building and delivers the appliance specific information and feedbacks through an online platform (in-home devices or smartphones' apps). In the industrial sector the NILM device can be pared with existing submeters or connected to just few critical circuits.

- B2B2C for Utilities and Service Companies: partnering with utilities that are directly providing and installing the smart meters to the customers. This model also includes the possibility of partnering with building owners and housing corporations to supply the disaggregation devices to the residential consumers. The energy disaggregation is delivered as SaaS for utilities and both as SaaS and hardware solutions for service companies.

The most common approach was long to deliver the disaggregation services directly to the consumers, mainly residential. However, in recent years, utilities demonstrated to be interested in the services provided by NILM algorithms and an increasing number of partnerships have been stipulated with companies offering energy disaggregation.

Additional business proposals for the B2C residential model include the possibility of an online marketplace based on an e-commerce approach. The consumers who identify an inefficient appliance can receive a cost-benefit analysis to replace their device with a more efficient one or customers with the potential of installing PVs and storage systems can be advised on purchase and installation options and directly connect with manufacturers [54]. Other proposals are the home automation, in particular for automatic adjustments of the heating and cooling systems knowing the disaggregated consumption of each system and the occupancy of the inhabitants, and the provision of household insurances for potential risk in the devices utilization [55]. Lastly, another innovative business model proposed is the one by *Verv* which is developing a blockchain-based peer–to–peer energy trading platform using energy disaggregation to understand the electricity production and consumption of households [56].

4.2.3.1 Examples of companies offering NILM

At the present date, there are 42 recorded companies offering NILM technology services [57], 10 more when compared to the energy disaggregation companies listed in 2016 [49]. The fast implementation of this technology in recent years is linked to the great progress in the machine learning and IoT fields but also to the possibility of offering a large number of innovative and diversified energy services to the consumers, increasing their engagement and satisfaction.

One of the leading companies in the NILM industry is *Bidgely*. Founded in the US in 2011, this company is providing energy disaggregation partnering with global utilities and energy retailers. Their NILM technology, protected by 17 patents, is claimed to deliver “true disaggregation” based exclusively on data from each house’s smart meter. Contrarily, “statistical disaggregation” relies on statistic models which are based both on the smart meter data and on geographic and climate information which helps defining the average typical monthly usage of a household. The main differences can be noticed in case of modified scenarios such as going on vacation or the birth of a child where for statistical disaggregation little or no changes can be noticed in the itemized energy bills. True disaggregation instead, is able to clearly recognize these changes providing a more precise energy breakdown. This approach, not being based on statistical data about the location, also offers the opportunity to deliver reports comparing the effective energy consumed to that of other households in the same neighbourhood [55].

Another approach to energy disaggregation is offered by *Opower*, a company acquired by *Oracle* in 2016 which claims to have already helped saving 20 TWh of electricity with their energy efficiency programs partnering with more than 100 utilities worldwide. The design principles of their cloud-based software are based on behavioural change science applied to empower energy consumers through the use of actionable savings tips, and visual cues generated from their energy disaggregation data [58].

The disaggregation business model for utilities has been largely employed by several other companies such as *Uplight* (more than 346 million customers), *Innowatts, Powerly, Smart Impulse, Homepulse, Net2Grid* and *Grid4C.*

However, many disaggregation companies, especially in Europe, are still relatively small and mainly target homeowners directly. An example is *Smappee*, a Belgian company providing hardware solutions to the residential sector. Their NILM technology includes the possibility of managing electric vehicles, solar production, energy storage consumption and helps the consumers to identify loads which are always switched on. On the downside, their solution must be trained by the users to recognize the single appliances by switching them on and off several times before being classified by the algorithm [59]. Other companies which have already a significant number of customers are *Neurio*, *Verdigris*, *Sense* and *Watt-Is.* These companies employ the residential business model for energy disaggregation but are also expanding their business to utilities and enterprises.

Finally, *PlotWatt* is one of the few companies which are offering energy disaggregation for commercial customers. Their cloud-based disaggregation software assists several fast-food chains in helping detecting inefficiencies, malfunctions and broken appliances. Their type of customers indeed, are commonly equipped with the same brand and types of devices which make it easier for NILM algorithms to disaggregate a predefined appliance model [60].

4.2.4 Market Potential

According to a recent study, the global Home Energy Management System (HEMS) market was valued at $7,7 billion in 2019 and it is expected to reach a value of $26,9 billion by 2025 [61]. This sector includes all the technologies for monitoring, managing and controlling the house. It includes lighting, PVs, HVAC and other home appliances; thus, it is not limited to energy disaggregation alone. However, the rapid growth foreseen for this market can be considered as an indication of the potential offered by the NILM technology, justified by the introduction of several disaggregation companies among the 24 innovators in the sector considered in this market potential study [61].

The main market drivers measured in this just mentioned analysis are the rising penetration of smart meters, high energy efficiency investments, increasing of distributed generation and smart grids, elevated electricity prices and growing understanding and adoption of IoT technologies, cloud solutions and data analytics.

Regarding the energy disaggregation field, the US market shows to be the most consolidated due to the high implementation of NILM technologies based on partnerships of disaggregation companies with utilities, as seen with *Uplight*, *Bidglely* and *Innowatts,* already offering NILM services to millions of customers all over the United States. In comparison to the North American market, the European market is still segmented, and companies primarily focus on residential applications. However, due to its competitive energy market, the high electricity prices, the numerous energy efficiency policies and the strict regulation targets, the EU holds a great market potential for energy disaggregation as seen in *Figure* 4.3 [16].

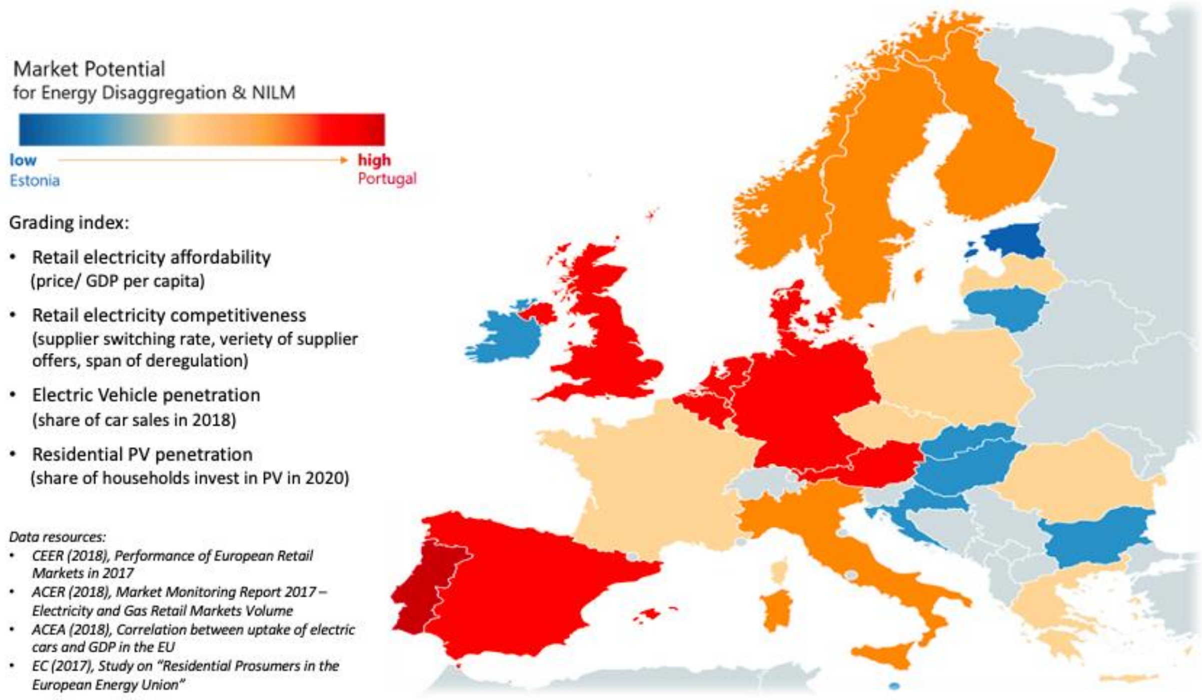


Figure 4.3 EU heat map for energy disaggregation market potential [16]

Figure 4.3 depicts Portugal as the European country with the largest business potential for the NILM technology. In fact, Portugal is among the three countries with the highest electricity prices per income while at the same time the country shows a fast increase of solar panels installations and electric vehicles sales [62]. Other areas with a high potential for energy disaggregation are considered to be the UK, the Netherlands, Denmark, Germany, Belgium, Spain and Austria. Out of those countries, the UK is one of the forerunners for HEMS adoption due to ambitious energy policies and regulations aiming at reducing the country's carbon footprint with energy efficient buildings and smart meter rollouts, making of NILM a very promising technology [63].

4.2.4.1 Costs and Benefits

The market rollout of energy disaggregation is hindered by the initial installation costs and the uncertainties regarding the performance of the technology and consequently the effectiveness on the energy savings.

There are a limited number of studies which propose a cost analysis for the implementation of energy disaggregation. At the moment, the installation of a NILM hardware solution for residential customers who are already equipped with a smart meter is approximated between \$200 and \$500 [64]. As it will be further analysed in chapter 4.3.3, the expected percentage of energy savings is difficult to predict and it is strongly influenced by the kind of feedback that is sent to the consumers and other numerous biases. According to the results obtained in this study, when savings are obtained with direct and personalized appliance specific feedbacks, an average of 5% energy savings per household can be expected. This scenario is set as reference to estimate the potential cost benefits of the technology.

The analysis that was conducted in this thesis compares three different markets: United States, Europe and the Netherlands. The percentage of smart meters rollout for each of those markets is expected to be 80%, 77% and 95% respectively by 2024 [8][65][66]. The same percentages have been assumed as the correspondent potential for NILM technology adoption. The total cost of an energy disaggregation hardware solution is overestimated to be 500€ per household. Additional information on the calculations and the results are reported in *APPENDIX II* (Data sources: [67][66][68][69][70]).

In the US, if 80% of the households would benefit from energy disaggregation this could result in 11 TWh of electricity saved nationally each year corresponding to approximately 1 billion euros cost savings. In Europe, where there is a higher number of total households and the price of electricity is more than double the price of electricity in the US, the savings would reach 45 TWh of electricity which leads to 9,6 billion euros annual savings. In the case of the Netherlands, the implementation of energy disaggregation would lead to 1 TWh energy savings, corresponding to 179 million euros savings.

However, if the high investment for the NILM hardware is done by the individual residential consumers, the savings that are obtained are definitely small, even when assuming the rather optimistic value of 5% energy savings used for the analysis. Despite these energy savings, the return of investment for a Dutch family would be almost 20 years, for a European family it would be less than 12 years, and, in the US, the disaggregation investment would be covered in 9 years.

In conclusion, despite the energy savings obtained from the implementation of NILM under optimistic assumption, the price of 500€ for the hardware costs is still too high for a residential consumer in Europe to independently adopt the technology. Consequently, the role of utilities, housing corporations and other large organizations would be fundamental for the widespread utilization of energy disaggregation technology, driven by the potential benefits mentioned in chapter 4.2.1 and by the strict targets imposed by the EU to reduce their energy consumption.

4.3 Social Impact

The implementation of a NILM technology able to provide an accurate and detailed appliance specific breakdown of the energy consumption to consumers is alone not enough to achieve quantifiable energy efficiency improvements. The energy disaggregation data obtained needs to be analysed and converted into actionable feedbacks for the consumers in order to achieve greater energy savings [71]. This transition can be accomplished by incorporating the understanding of the psychological and behavioural characteristics of energy consumers to the technological challenges.

This section focuses on how energy disaggregation, and in general energy feedbacks, can have an impact on promoting energy efficiency among residential electricity consumers. Firstly, the concept of environmental psychology and some pioneering behavioural theories are introduced to describe how customer engagement and awareness on the topic can be increased. Subsequently, different typologies and characteristics of feedbacks are explained and classified by effectiveness based on literature reviews. Finally, the impact of appliance specific and whole house feedbacks is compared and quantified in terms of potential energy savings obtainable per household.

4.3.1 Behavioural Change and Customer Engagement

The sustainable utilization of electricity and the reduction of its consumption, defined as energy conservation, constitute an especially difficult challenge for the residential sector. In fact, due to the intangible and ephemeral nature of electricity, energy conservation is not perceived as a coherent field of action by households' inhabitants. Nowadays, most of the behaviours performed every day are connected to the consumption of electricity, from the usage of appliances in the house to the charging of mobile phones [72]. For the consumers, these diversified sets of actions and behaviours are difficult to be associated quantifiably with their final energy bills. Thus, it is difficult to feel a sense of control over the modification of their consumption. For this reason, customer engagement is a key resource to increase the general awareness on the issue of energy conservation and transform electricity consumers into active agents able to understand their own energy utilization and adjust their behaviours accordingly.

Recognized in the late 1960s, *Environmental Psychology* is an interdisciplinary field which studies the interaction between individuals and their surroundings. In recent decades, when environmental problems such as climate change and pollution have been largely attributed to anthropogenic actions, this social science provides different methodologies to modify human's behaviours in order to reverse the effects of these problems while preserving people's needs and comfort [73]. In the energy disaggregation framework, some of these theories can support the adoption of NILM technologies and the integration of appliance specific data with effective approaches to provide tailored feedbacks. By doing so, consumers are supported to naturally shift to more sustainable behaviours.

There are three aspects of behavioural change that should be taken into account in order to maximise the positive impact on energy conservation. The first is the elasticity of the change which is the extent of energy reduction derived from the adoption of the behaviour. Second is the plasticity or probability of the behaviour to be implemented. Third and last is the timescale with which the changes may be accomplished. These factors are often in discordance since for example the easiest interventions to be adopted may also be the ones which seems to lead to a lower impact on the environment such as reducing the temperature of the thermostat or unplugging appliances. Contrarily, the proposed substitution of a fridge or the retrofit of a household would require a larger timescale and lower plasticity of adoption by the consumers and thus, encountering higher resistance to be changed. It has been demonstrated that directly targeting interventions with high elasticity may not be the most effective starting point [12]. Even though all the measures are important to be adopted, it seems simpler to begin with actions which are most easily implementable, and which nonetheless lead to reduced energy consumption and constant environmentally friendly habits.

In order to understand which are the different phases that an energy consumer has to face for a successful behavioural change, the *heuristic model of environmentally relevant behaviour* is presented in a modified version which integrates the area of feedbacks [72], together with the *Transtheoretical Model of Change (TMC),* cited in several publications on energy efficiency [74][75]. The following *Figure* 4.4 is a graphical integration of the two theories aimed at the comprehension of the process which leads to building new sustainable habits, through the support of relevant feedbacks.

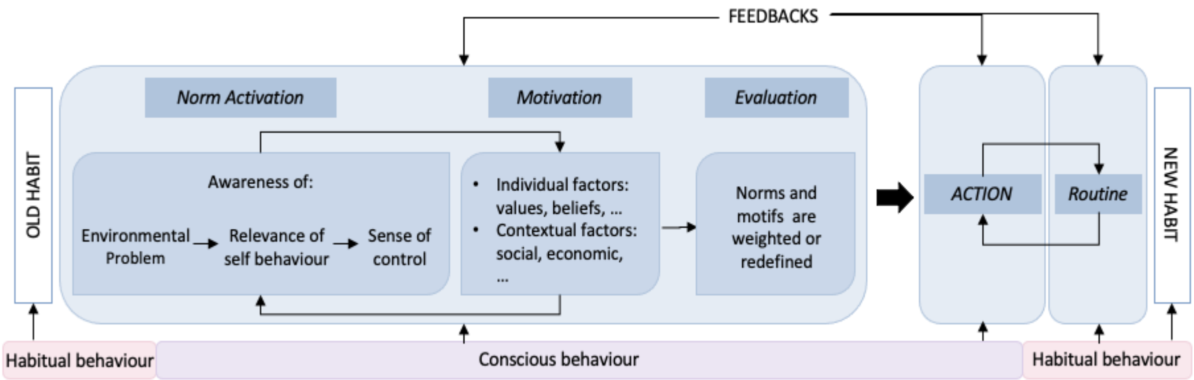


Figure 4.4 Revised scheme of two environmental psychology models: Heuristic model of environmentally relevant behaviour and Transtheoretical model of change (TMC)

According to a simplified version of TMC tailored for the context of residential energy consumption [75], there are three stages which an individual faces when engaging with the process of behavioural change: raise awareness (left light blue cluster), inform complex actions and maintain sustainable routines. All these steps should be continuously supported by a feedback mechanism which would act as a reinforcement during the behavioural process development, identifying the constraints that people face, and offering information at a time and place which is close to the decision by the consumers (real-time feedbacks). The possible feedback systems and their characteristics will be further analysed and explained in the following subchapter 4.3.2.

The first stage of raising awareness among energy consumers is the most complicated one. Studies have demonstrated that there is a high misconception of energy utilization in the households especially when people are asked to estimate their consumption or to disaggregate their bills identifying the most consuming devices [74]. In *Figure* 4.4*,* this stage is represented by the three blocks of the *heuristic model of environmentally relevant behaviour*: norm activation, motivation and evaluation. In order to replace an old habit, a large quantity of information and feedbacks are needed for a consumer to start questioning his own norms and beliefs in a process called norm activation. The person must realize that there is an environmental problem, understand that his or her behaviour is relevant in the contribution to the problem itself, and become conscious that he has the possibility of influencing his behaviour and consequently the problem. For example, this sense of control over the energy consumption in a household can be significantly increased with appliance specific information which would allow the consumers to have a breakdown of their electricity utilization and adopt direct measures to reduce their consumption of one specific device [72].

After the consumer has realized that he or she can contribute to the problem the norm activation phase is completed. Afterwards, he needs to have a motivation to be convinced to act. The drivers for motivation can be categorized into individual and contextual factors. The individual factors are constituted by self-identities, beliefs, norms and personal values. Contextual factors are economic, cultural, institutional aspects that may vary across different social groups but are also influenced by the access to technology and information which may affect the consumers' motivation. Among these drivers, personal values play one of the most important roles, representing the goals for which people strive for in life. These can be classified into four types: hedonic values when people focus on their pleasures and on what make them feel good, egoistic values when people focus on their resources and their self-interest, altruistic values when people focus on the consequences or benefits for others, and finally biospheric values when people focus on the consequences for nature and the environment [12]. In every human, all of these values coexist even though some are more pronounced than others. The values prioritized by an energy customer are thus important to define the type of motivator, and consequently feedback, that would appeal most to the consumer and stimulate action. The drivers could vary from obtaining monetary savings rather than saving the environment or helping the future generations or their communities to live a better life. However, it may happen that sometimes norms and motifs are in conflict with each-others and thus, the person needs to weight and redefine all these aspects in what is called the evaluation process before taking the final decision leading to the action [72].

At this point, the conscious behaviour constituted by the awareness stage and the action itself has been taken. Nevertheless, for the behavioural change to be effective, the person needs to pursue a routine and keep sustaining the change until it is transformed into a habitual behaviour and thus, a new sustainable habit (*Figure* 4.4) [74]. In this final step too, feedbacks represent a key role to support the achievements obtained by the consumers or helping them redefine their decisions in case the expected results are not obtained.

There are numerous other environmental psychology theories which are applied in the energy conservation sector and aim at understanding why people engage in pro environmental behaviours. Among the most dominant models there are the *Theory of Panned Behaviour* (TPB), the *Value – Beliefs – Norm theory of environmentalism* (VBN) and the *Norm Activation Model* (NAM). While the former supports the idea that individuals make rational choices with the intention of obtaining maximum benefits with minimum costs, the VBN and the NAM assume that people conserve energy when moved by moral obligations [76]. Those obligations are triggered by environmental concerns and values in case of the VBN, and by the normative and moral dimension in the NAM [77].

In conclusion, environmental psychology and its interdisciplinarity result as an important tool to identify the necessary steps to induce behavioural change. This would then help in the selection of the most effective strategies to engage the customers to reduce their electricity consumption and increase their awareness on energy conservation.

4.3.2 Energy Feedbacks

The process of providing people with information about their actions or behaviours is known as feedback. This well-studied mechanism has been proved to be an effective strategy, if adequately delivered, to increase the awareness of energy consumers and consequently lead to energy savings. As emphasized in the previous subchapter 4.3.1, the feedback can serve as a trigger during the process of behavioural change. Additionally, the feedback can be exploited in three different ways: as a spark emphasising the problem, as a facilitator helping people to reach the targeted behaviour and as a signal reminding individuals to perform that behaviour [78].

In this context, the expression eco-feedback refers to a technology that aims at providing strategies and feedbacks on individual or group of behaviours in order to reduce the environmental impact and influence energy savings. Based on the TPB and NAM environmental psychology theories, this technology is specifically designed to strategically present information to energy consumers considering the complexity of human behaviour and leveraging the availability of sensing technology and interactive displays [79]. In order to be more effective, the eco-feedback technology includes different interventions that can be used individually or in combination, acting as strategies to increase the people's motivation towards pro environmental behaviours. These can be classified as antecedent interventions, occurring prior to the action during the awareness stage in *Figure* 4.4, and as consequent interventions happening after the action occurred [80].

The antecedent interventions are information, goal setting, modelling and commitment [79][80]. The first and most widely diffused way of intervention is simply presenting information. Besides providing general knowledge, information can also include justifications to explain the reasons why those behaviours should be adopted, and instructions to teach how to perform that behaviour. However, this type of intervention alone results in marginal effects and rarely leads to action. Studies suggest that its effectiveness depends on the individual's values and other motivations. Thus, information should be coupled with other kinds of interventions [80]. The second type of antecedent strategy is goal setting which consists in defining an objective to be achieved in a future scenario. These objectives can be self-set or directly assigned to individuals or groups of people in order to stimulate them to reach the imposed energy savings targets. Nevertheless, the new goals can be in conflict with pre-existing ones and thus, this strategy is very often assisted by energy programs and other strategies such as information and feedbacks to give people the possibility of checking their own progress. Modelling is another antecedent intervention, and it consists of observing peers as models to perform and demonstrate sustainable behaviours. This is a tool which is based on the concept of learning in a social environment where people apprehend how others behave in certain contexts and are motivated by their actions. Nowadays, with the popularity of social media and television shows it is easier to deliver this knowledge to a broader audience and show the appropriate behaviour for energy conservation. Finally, commitment is the promise to behave in a specific way, stimulating the consistency, the attitude and the social influence of an individual. This kind of intervention has long lasting impact on the behaviour, even when the pledge is no longer active.

The second type of eco-feedback strategies are consequent interventions [79][80]. Those are used post-behaviour to affect future decisions, to help the establishment of a routine or to evaluate its effectiveness. There are only two kinds of consequent interventions. The first is the implementation of rewards and penalties, often connected to the goal setting strategy. Some usual techniques are the gamification which involves small rewards such as points when the goal has been reached, and tax reduction for long term investments such as solar panels installation. However, this is a short-term strategy to instigate behavioural change and it has been shown that the electricity usage often returns back to the baseline after the reward is removed, differently from commitments. The last consequent intervention is the feedback itself. In this part of the analysis the feedback is to be more specifically considered as the response to the different behaviours and actions that energy consumers perform quotidianly as it is further clarified and classified hereunder.

4.3.2.1 Types of Feedbacks

As discussed until now, the behavioural change which leads to energy savings depends on a series of individual and social factors that need to be supported by customized feedbacks and specific interventions in order to move the largest number of consumers to act sustainably. In fact, the effectiveness of the feedbacks largely depends on how, when and to whom those feedbacks are delivered [81]. This type of consequent intervention is used after or during the implementation of the behaviour itself and it is commonly associated to the monthly electricity bill that every energy consumer is used to receive. However, there are many other ways in which it is possible to receive feedbacks about the own electricity consumption. *Table* 4.1 distinguishes between the two main types of feedbacks, indirect and direct feedbacks, and their relative subcategories, while further describing their characteristics in accordance with four dimensions [82][83].

Table 4.1 Feedbacks classification: indirect and direct

	INDIRECT FEEDBACK		**DIRECT FEEDBACK**	
	Standard Billing	**Enhanced Billing**	**In-Home-Display (IHD)**	**Web Connected Devices**
Frequency	Monthly or Bimonthly	Monthly or Personalized	Always-on	Always-on
Medium	Paper bill	Paper or Electronic bill	IHD or Web	IHD, Web or Smart Meter
Type of Information	• Energy consumption • Historic comparison	• Energy consumption • Historic and Social comparison • Energy Efficiency advice	• Energy consumption • Historic and Social comparison • Real-time information	• Energy consumption • Historic and Social comparison • Real-time information • Energy disaggregation
Type of Interaction	Passive	Passive	Active	Active and two-way communication

The indirect feedbacks are the ones received after the electricity has been consumed. Those are traditionally delivered each month by energy retailers and generally show the total electricity consumption of the household and the historic records of the previous years. Contrarily to the standard billing which provides the customers with non-engaging information, the enhanced billing relies on third parties to deliver extra services such as statistical data, energy efficiency advices and social comparisons with similar households. Using this method, the bills could also include some antecedent interventions like goal setting and information. However, both the indirect feedbacks rely on a passive interaction with the consumers. The direct feedbacks on the other hand, are delivered by software platforms, interfaces, or smart hardware located in the house such as in-home displays (IHD), allowing the users to actively engage with real-time information on their energy consumption. In particular, the web connected devices feedbacks rely on the connection of the smart meter to the internet as explained in chapter 4.2.2, enabling energy disaggregation technologies and allowing the consumers to control their smart devices and receive customized feedbacks, in a bidirectional communication [81][82].

Besides the four categories (frequency, medium, type of information, and type of interaction) considered for the classification of direct and indirect feedbacks, there are also many other dimensions that should be considered when developing a feedback system, depending on the aim the system. In addition to the measurements already presented in *Table* 4.1, the following paragraphs include information on further possible dimensions some of which have been grouped together. The six categories considered are [72][74][78]:

- Frequency and Duration: those indicate respectively how often and for how long the feedback is given. The studies clearly show that frequent feedbacks would increase the people's consciousness about the consequences of their behaviours and improve the connection between actions and effects. However, if delivered too frequently those could result in being ignored since for many people energy is not a priority issue. Duration on the other side, is a controversial factor. While it is true that users are more engaged in the initial period of the feedback introduction due to its novelty, their attention fluctuates over time resulting in strong reductions for medium term studies. Only few theories suggest that in the long-term duration, feedbacks would contribute to habit formation and persistent energy savings.

- Measurement: is the unit of measure by which the feedback is given. The information can be provided in terms of energy unit (kWh), economic unit (euro), or carbon emissions (tons of CO_2). These measurements are dependent on the values that the targeted groups are striving for, and thus, their relevance may vary according to the customer segments. Nevertheless, in almost all the studies the feedbacks provided contain all the information, especially the energy and cost ones, and consequently their effectiveness cannot be valued separately.

- Granularity: is the level of details presented by the feedback. These can be either whole-house or appliance specific electricity consumption. As it will be specifically analysed in the subchapter 4.3.3, the few studies conducted on the percentage of actual energy savings obtainable from energy disaggregation are controversial. However, the high level of details provided by appliance specific data could contribute to the awareness process and to increase the users' sense of control.

- Comparison: is referred to a comparison with past usage (historic) or with peers (normative). The historic comparison is a very useful tool to engage users to set personal goals and norms. In the same way, a comparison with other households could stimulate a sense of ambition and competition. For this case too, studies are divided in their results, not being able to demonstrate a strong effect of normative comparisons over energy savings.

- Medium and Design: those are referred to the medium through which the feedback is provided and the design of the interface with which the users interact. The most effective feedbacks have been proved to be the computerized and interactive ones which allow the consumers to choose the type of information they are most interested in. At the same time, the graphic should be simple, clear but not simplistic and composed by a mix of numerical, analogue and symbolic design.

- Type and Level of Interaction: those refers to the interaction of the users with the feedback (active or passive), and to the social interaction. As demonstrated from *Table* 4.1, active interactions increase the user's engagement with the feedbacks. In a similar way, sharing goal achievements, results of gamification progress, or energy savings, either on social media or with other users could trigger the consumers to participate and engage with the energy conservation.

These feedback characteristics can be arranged in very different ways and combined with other interventions in order to obtain the highest possible energy savings for the targeted consumers. Even though there is not a universal most effective feedback that could help everyone to save the same amount of energy, several studies identified those dimensions that showed to be of fundamental importance for the process of behavioural change.

Studies identified that the best hypothetical feedback should be a direct feedback system connected to the internet, delivered through an interactive display located in the household, such as an in-home display, where the users can actively access very frequently, on a daily basis. This display should be positioned in an area of the house accessible to every member of the family and provide a clear and appealing interface which should show the current electricity consumption, its comparison with historic data and with the consumption of other similar households, in addition to monetary and economic measures. Besides, the combination with antecedent interventions and some other common information such as the weather forecasts or the possibility of controlling the thermostat would attract the consumers to use the IHD more often. Contrarily, redundant and confusing information could lead the users to a reduced engagement. In addition, energy disaggregation constitutes another important dimension that should be present in a feedback mechanism in order to reduce the disconnection between consumer's actions and their consequent effects, while better identifying their patterns of consumptions and directly modifying their utilization to diminish the energy consumption [82].

Finally, a special mention is to be attributed to the rebound effect. This is the phenomenon according to which greater energy savings obtained through behavioural change would trigger additional energy consumption so that the net effective savings over time could be uncertain [74]. Since it was detected in the results of many studies, the rebound effect is a very challenging aspect influencing the feedbacks' net effectiveness.

4.3.3 Energy Savings

Numerous studies have been conducted over the years to understand how the different types of feedbacks and interventions affect human behaviour in the context of energy conservation. After having analysed the substantial variety of adoptable strategies, it is evident that the results obtained from these studies are as well variegated and not always representative of plausible scenarios. In fact, most of the analyses are carried out for small sample size and similar targeted groups in a short-term scenario of only a few weeks or months. Many biases are also present in the studies such as for example when people undergoing the experiment are aware of the aim of the study and thus, have the tendency to behave more sustainably, called the Hawthorne effect [84].

In this optic, systematic reviews try to find robust results across multiple studies which match defined criteria to answer a specific research question. Those reviews usually include meta-analyses which combine the different statistical results from the selected studies into a single statistical analysis to have a more accurate and realistic outcome. Most reviews are not directly aimed at the evaluation of energy disaggregation but instead target all aspects of feedbacks. However, one of these meta-analysis which assessed several studies conducted from 1995 to 2010, defines energy disaggregation as the “holy grail of energy efficiency”, attributing more than 12% of annual energy savings to direct feedbacks provided real-time with appliance specific information [52]. Another well esteemed review from 2008 compared different projects to identify the most effective feedback dimension in terms of energy savings. The energy savings resulted to usually range from 5% to 12%. In this case too, the most efficient feedbacks were those that provided frequent and detailed, appliance specific information [72]. Similarly, a recent meta-analysis on 70 studies collected over the last 50 years for the evaluation of the energy feedback systems was conducted by the European Union in 2019. This review suggested that an appropriate feedback mechanism could realistically contribute to reduce between 5% to 10% of a household’s energy consumption. Direct feedbacks with the two-way communication interaction (category including energy disaggregation) are recommended and considered more effective than indirect ones. However, the study did not directly mention the effectiveness of appliance specific feedbacks. It is worth mentioning that the lower percentage of energy savings compared to previous studies has been partially attributed to an increasing general awareness on energy efficiency by the population throughout the years [81].

In the context of savings achievable through energy disaggregation alone, only fewer studies are available. To date, the only systematic review on the effectiveness of domestic, disaggregated electricity feedback examined twelve studies that were conducted up to 2015 [84]. It investigates the possible energy savings on the whole population considering the results of the different studies to be positively biased and weighted by the number of participants in each study. The authors take into account whether the studies considered the Hawthorne effect or not, the medium through which the disaggregation was delivered (IHD or websites), and the presence of 16% “energy enthusiasts” in the population who are most willing to engage with energy conservation. The remaining percentage of the total consumers are rather pessimistically considered to save no energy at all, justifying this negative approximation to be balanced by the positive biases. The result of reduced energy consumption through disaggregation feedbacks ranges from 0,7% to 4,5% for the total population [84].

Additionally, this systematic review considers the “fine-grained” or true disaggregation obtained with high resolution data to not reduce the consumption to an extent which would justify the engineering efforts when compared to the statistically calculated “coarse-grained” disaggregation (explained in chapter 4.2.3). In fact, one study conducted by *Bidgely* in 2013 and included in the review, reported a 6% of average energy savings through their true disaggregation, cloud-based NILM solution with favourable reactions and high engagement by the 150 participants in California [85]. This result was not evaluated high enough compared to the estimated 4,5% for the fine-grained disaggregation to be considered necessary, even though no specific studies have ever compared the two technologies directly.

Finally, the review questions if the energy disaggregation feedbacks enable greater saving than aggregated ones. Despite the studies may suggest that aggregate feedbacks are more effective, it is important to analyse the differences in the medium that was used for both cases. For example, in one study conducted in 2015, appliance specific information was offered for three weeks through a website obtaining only 0,5% energy savings [80]. During the same period, real-time aggregated feedbacks presented on an IHD managed to save 8,1% of the electricity consumption. This difference can be associated to the greater attractivity and easiness of access of the interactive display over the website, and even more by the higher frequency of the visits, eight visits a day against only one visit a day for the webpage [80]. Overall, the comparison is not equal for the aggregated feedbacks to be considered as more effective than the disaggregated ones. Yet, the analysis suggested that users who were receiving appliance specific feedbacks had higher perception of their energy consumption. For this reason, the authors supported the idea that disaggregation would lead to more significative reductions.

There are two additional studies on the effectiveness of energy disaggregation that were not included in the mentioned systematic reviews and that in the context of this thesis are worth analysing. The first one differs from all the other studies because it presents a long-term analysis on real-time NILM and eco-feedback system for 52 weeks including twelve families in Portugal during 2011 [86]. Despite the increased awareness and better understanding of individual appliances consumption, the study reported no significant energy savings. This result was justified by the response-relapse effect according to which after the initial phase of novelty and curiosity the participants go back to their normal behaviours, in this case study decreasing their attention on the feedback interactive display by 90%. However, it is important to specify that the monitor was located in a hardly accessible area of the house where only one person, usually the father, was in charge of checking the consumption and that the disaggregation necessitated manual labelling. This mechanism started only a few weeks after the experiment begun, with the support of technicians because the energy consumers considered it to be a too complex system and they didn’t install it spontaneously [86].

The second and most recent study on energy disaggregation feedbacks was conducted with over 700 people in Germany for six months during 2017 [87]. In order to avoid biases, the purpose of the study was not mentioned to the participants who were provided with a NILM system, accessible via mobile application, which was able to distinguish only five of the most consuming devices, the always-on loads and the remaining “other appliances”. These participants were divided into four target groups: one only receiving disaggregated feedbacks, one receiving financial incentives over their energy savings, one being compared with other consumers in the study, and the last group as well compared to other participants, but also receiving monetary rewards based on ranking. The study showed that the monetary and comparison incentives did not affect the effectiveness of the disaggregated feedbacks. However, it demonstrated that the overall electricity consumption was reduced by 5% with disaggregated feedbacks compared to a fifth control group who received aggregated feedbacks. This confirms the results obtained from the systematic review. Another important insight from this study is that the highest reductions were obtained during the two parts of the day when people are most at home, the early morning and the evening, and highest savings were recorded for the most energy intensive appliances, reaching savings of 10% to 15% [87].

In conclusion, the studies supported the hypothesis that customized and tailored feedbacks lead consumers to increased electricity savings. In particular, energy disaggregation used as a tool to reduce electricity consumption via feedbacks resulted to have a positive effect on the sense of control and awareness of the consumers and showed potential energy savings between 0,7% and 5%. However, there are still doubts about the worthiness of fine-grained disaggregation compared to coarse-grained disaggregation and about whether the effectiveness of a feedbacks is more affected by its dimensions such as medium, frequency and design rather than the greatest technological effort of increased granularity.

4.4 Challenges and Limitations

The economic and the social impacts of the NILM technology proved to have a great potential on energy consumers and other stakeholders if wisely implemented. Nevertheless, both the analyses encountered numerous challenges and limitations that are hampering on one side the experimentations on the achievable energy savings, and on the other, the possible economic deployment of the technology. The main limitations are described in this chapter subdivided according to economic and social impact.

4.4.1 Economic Impact

The numerous and variegated possible applications of the NILM technology facilitate the potential deployment of energy disaggregation and bring together different customer segments and stakeholders. However, the deployment of energy disaggregation is still hampered by the technical challenges of the hardware and SaaS solutions. In fact, to employ energy disaggregation both the solutions still necessitate of an expensive hardware installed inside the household that provide high resolution data from the smart meter. In addition, these solutions, if not implemented inside the household, need to assure the data transmission on the cloud with appropriate protocols and personal data security procedures.

Moreover, the prices of this innovative technology are still too high and the results too inaccurate for the energy consumers to personally invest on energy disaggregation. For this reason, at this state of the art, there should be more incentives and motivations from utilities and service companies to deliver NILM technology and foster its implementation. Using the B2B2C business model could be convenient for utilities which already have smart meters installed in numerous households. This could increase the trust of their customers who could obtain extra incentives, besides helping the prices of energy disaggregation technology to be reduced.

Finally, there is no availability of a market potential study on the NILM technology alone. The one performed in this thesis has many limitations linked to the missing data on the numbers of smart meters already installed in each of the three nations analysed in the study. Furthermore, while the 5% of achievable energy savings used, as seen in the social analysis, may be positively biased, the estimated cost of the technology at 500€ is a very high approximation of the NILM price, and thus providing a strong negative bias to the results.

4.4.2 Social Impact

The analysis on the social impact of energy disaggregation initially focused on the understanding of how the process of the consumers' behavioural change can be influenced to act sustainably and save energy. In this context, the challenge lays in the multitude of different values and norms that define human beings and consequently, in the need of the feedbacks to be customized and tailored to the specific targeted groups. Moreover, the adoption of new pro environmental habits is a delicate phase which should be sustained by frequent inputs and engaging long-term interventions that would help the customers to learn and familiarise with the changes.

The main limitation in this field is in the lack of long-term studies comparing several feedback dimensions among large and diversified targeted groups in order to verify the actual effectiveness of the feedback mechanism. In fact, the significant costs associated to the deployment and operation of the studies led to a majority of small sample sizes, very similar target groups and short-term analysis which limited the accuracy of the results. It is important to test the effectiveness and cost benefit of the feedbacks on different targeted groups, both with low and high electricity consumption, or according to the demographic area. In fact, there is a lack of comparative studies between national and international experiments that could help with the definition of national energy efficiency programs to help the adoption of a structured feedback system. Besides, due to the lack of long-term studies there is little statistical information about the response-relapse effect on the decreasing utilization of the feedbacks over time, and the rebound effect according to which the total energy savings are cancelled in time by other triggered energy consumptions. An additional issue for these studies is the inevitable presence of biases caused for example by the small sample size of the experiments, the Hawthorne effect, and the publication biases of negative results less willing to be published.

In conclusion, the possible 0,7% to 5% reduction of energy consumption achievable from energy disaggregation feedbacks may be a limitation for the implementation of the NILM technology. However, more studies with equal conditions should be conducted on the potential of this technology to accurately compare the effectiveness of appliance specific over aggregated feedbacks and between true and statistical disaggregation. Still, energy disaggregation may provide consumers with increased sense of control, energy awareness, and knowledge on specific and tailored ways to decrease their own electricity consumption.

5. Conclusion

This thesis provided a detailed evaluation of the Non-Intrusive Load Monitoring (NILM) technology from a technical, economic and social prospective, specifically for the residential use-case. The Netherlands was set as reference scenario for the energy disaggregation modelling aimed at the recognition of the fingerprints of the most common household's appliances.

The different stages of the NILM process were used as a structure to define the state of the art of the numerous algorithms that are being developed with the progresses in the machine learning field. Even though the event-less approaches remain the most computationally intensive methods, those have the potential to achieve the best load identification results through unsupervised learning algorithms. In order to assess their performances, the two reference algorithms, CO and FHMM, were implemented for the REDD and DRED datasets. Overall, the FHMM algorithm performed better on both datasets showing a more precise signature recognition of the top five consuming appliances. The REDD datasets obtained the highest percentage of true positives, while the DRED was found to have the lowest root mean square error (RMSE). This means that the algorithms recognized the highest number of exact events for the REDD dataset, and the highest concentration around accurate events for the DRED dataset. The analysis emphasized the limitations of the NILM technology which is still relying on not standardized performance metrics for the evaluation of the results and the absence of a common framework that could simplify the comparison between different datasets and algorithms.

The second energy disaggregation model was designed to test both the hardware installation and the potential of event-based approaches without using machine learning. The developed algorithm was inspired by expert heuristic methods that analyse the differential of the changes in the active power of the appliances and it tried to detect the occurred events through the extraction of the main features of the devices. The results were obtained for a typical Dutch house during the month of September 2020 and for the specific case study of the fridge. Despite the economic and easily implementable system design, the algorithm, not including a learning approach, was not able to identify multi-state and continuously variable devices. However, the model precisely recognized the fridge signature and its duty cycles, showing its potential to be used as a trigger to send push notification to the consumers for misfunctioning and inefficient devices. Overall, machine learning based approaches are believed to obtain more accurate results and precise appliances classifications.

Furthermore, an economic analysis was performed to assess the business models and market potential of the NILM technology. Two main customer segments were identified as the energy consumers, and the utilities and service companies. There are numerous applications for which energy disaggregation can be employed. Besides promising residential energy savings, the NILM technology has a great potential to foster demand response, grid flexibility services and more tailored energy efficiency programs. The NILM products were classified in hardware and SaaS solutions in accordance to how and where the disaggregation is performed. Currently, the hardware solutions are easier to implement due to the commonly low frequency resolution of the smart meters installed in the households. Even though cloud and SaaS solutions have seen a recent increase in deployments, data transmission and security remain important issues to be considered.

The market potential of energy disaggregation and in general of the HEMS is forecasted to grow strongly in upcoming years. Nevertheless, the NILM solution resulted to be too expensive for the residential consumers to personally invest in this technology, having a return of investment in the Netherlands of 20 years. In order to be widely implemented, utilities could be the facilitators, since they are already in charge of the smart meter installations in the households. Utilities could promote and deploy the NILM technology on national scales and thus, help the disaggregation solutions to reduce the initial costs in a B2B2C business model.

Additionally, the potential impact that energy disaggregation can have on energy conservation of the residential consumers has been discussed. However, providing an accurate and detailed appliance specific breakdown to energy consumers is not enough to achieve quantifiable energy efficiency improvements without the implementation of a customized feedback mechanism. This feedback would provide the consumers with engaging and tailored recommendations that could lead them to create new pro environmental habits and to act sustainably. A direct feedback system, frequently delivered through in-home displays and in combination with interventions such as goal setting, comparisons, information and appliance specific data, proved to increase energy awareness and savings. The numerous studies conducted throughout the years to understand the percentage of energy savings obtainable using these feedbacks are influenced by several biases, small sample sizes of the targeted groups and the short-term experiments. For this reason, it is still not certain that the identified 0,7% to 5% reduction of energy consumption achievable through energy disaggregation feedbacks is enough to be judged more effective than the one obtained from aggregated feedbacks.

Finally, in a context where the electricity consumption is intensifying, the grids are becoming more congested and the buildings need serious energy efficiency refurbishments, the energy disaggregation technology, if widely implemented, could contribute to alleviate these issues. The NILM technology necessitates further studies to obtain more accurate results, yet the conservation effects from smart meters rollout seem to be larger if appliance specific feedbacks can be provided. Besides rising awareness and sense of control on the energy consumption for residential customers, energy disaggregation could constitute a tool to better understand, forecast and manage electricity on larger scales in a future made of smart grids, efficient buildings and sustainable energy consumption.

5.1 Future Work

The comprehensive overview of the NILM technology proposed in this thesis reviled the numerous limitations preventing energy disaggregation from being widely implemented and it is aimed to foster additional studies to fill these gaps.

Firstly, the researchers in the NILM communities should propose a universal framework where it would be possible to perform novel disaggregation algorithms and equally compare the different methods and results. As proved in this thesis, the standard performance metrics are not sufficient for a fair evaluation of the algorithms' outcomes. Besides, the data acquisition phase showed a strong limitation for the technical analysis performed in this thesis. To be accurately assessed, the signature recognition algorithm proposed should be evaluated with a dataset including individual loads measurements of the devices present in the household which in this case was not possible due to a mismatch of data format with other available datasets. Furthermore, as in the case study of the fridge in chapter 3.3, the profiles of other appliances should be analysed, and their features identified in order to validate the potential of this algorithm for the detection of disaggregated loads.

Moreover, the economic analysis showed that at the current price, the NILM technology is not cost-effective for residential consumers and thus, more studies should be performed to assess and validate its economic impact. In particular, similar works could analyse how advantageous would be for utilities to invest on energy disaggregation and what would be the penetration rate of NILM in the three countries considered in this thesis.

In conclusion, there should be more studies where the social and psychological research about how to deliver the information obtained from energy disaggregation to the consumers are integrated with the technical research about NILM. In fact, it is important that more assessment studies could focus on the comparison between fine-grained and coarse-grained disaggregation in terms of energy savings and increased awareness. In this way, the intense engineering efforts to obtain more granularity and precision from energy disaggregation would be justified by the economic and social advantages that this technology could bring to energy consumers.

Bibliography

[1] IEA, “The Critical Role of Buildings,” Paris, 2019. Accessed: Apr. 27, 2020. [Online]. Available: https://www.iea.org/reports/the-critical-role-of-buildings.

[2] European Commission, “In focus: Energy efficiency in buildings,” *European Commission*, 2020. https://ec.europa.eu/info/news/focus-energy-efficiency-buildings-2020-feb-17_en (accessed May 01, 2020).

[3] BPIE, “97% of buildings in the EU need to be upgraded,” *BPIE - Buildings Performance Institute Europe*, 2017. http://bpie.eu/publication/97-of-buildings-in-the-eu-need-to-be-upgraded/.

[4] European Parliament and the Council, “Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings,” *Off. J. Eur. Union*, vol. L, no. 153, pp. 13–35, 2010.

[5] Energiesprong Foundation, “Energiesprong,” *Energiesprong Foundation*. https://energiesprong.org/ (accessed Dec. 19, 2020).

[6] Stroomversnelling, “home - Stroomversnelling.nl,” *Stroomversnelling*. https://stroomversnelling.nl/ (accessed Dec. 19, 2020).

[7] European Commission, “Smart grids and meters,” *European Commission*, 2019. https://ec.europa.eu/energy/node/83 (accessed Dec. 19, 2020).

[8] F. Tounquet and C. Alaton, “Benchmarking smart metering deployment in EU,” *Eur. Comm.*, no. December, 2019, doi: 10.2833/492070.

[9] R. Gold, C. Waters, and D. York, “Leveraging Advanced Metering Infrastructure to Save Energy | ACEEE,” *Energise 2020 Energy Innov. a Sustain. Econ. Pap. Proceedings. New Delhi 2020 Alliance an Energy Effic. Econ.*, no. 202, pp. 349–356, 2020, [Online]. Available: https://aceee.org/research-report/u2001.

[10] A. Zoha, A. Gluhak, M. A. Imran, and S. Rajasegarar, “Non-intrusive Load Monitoring approaches for disaggregated energy sensing: A survey,” *Sensors (Switzerland)*, vol. 12, no. 12, pp. 16838–16866, 2012, doi: 10.3390/s121216838.

[11] Watch-E, “Assist residents to achieve energy performance,” *Stroomversnelling*, 2020. https://stroomversnelling.nl/nieuws-bericht/begeleid-bewoner-om-energieprestaties-te-behalen/.

[12] L. Steg, R. Shwom, and T. Dietz, “What Drives Energy Consumers?,” *IEE Power Energy Mag.*, no. January/February, pp. 20–28, 2018.

[13] J. Revuelta Herrero *et al.*, “Non intrusive load monitoring (NILM): A state of the art,” in *Advances in Intelligent Systems and Computing*, 2017, vol. 619, pp. 125–138, doi: 10.1007/978-3-319-61578-3_12.

[14] A. Ridi, C. Gisler, and J. Hennebert, “A survey on intrusive load monitoring for appliance recognition,” *Proc. - Int. Conf. Pattern Recognit.*, pp. 3702–3707, 2014, doi: 10.1109/ICPR.2014.636.

[15] A. L. de S. Pereira, “Hardware and software platforms to deploy and evaluate non-intrusive load monitoring systems,” Universidade da Madeira, 2016.

[16] Cleantech Group, “Energy Disaggregation and NILM – If You Can Measure It, You Can Improve It,” *Cleantech Group*, 2019. https://www.cleantech.com/energy-disaggregation-and-nilm-if-you-can-measure-it-you-can-improve-it/ (accessed Dec. 19, 2020).

[17] B. Townson, “‘NILM: Vehicle or Destination?,’” *International Workshop on Non-Intrusive Load Monitoring*, 2016. https://www.youtube.com/watch?v=h9qykpYcnP4&feature=youtu.be.

[18] 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, pp. 1–29, 2019, doi: 10.3390/en12112203.

[19] G. W. Hart, “Nonintrusive Appliance Load Monitoring,” *Proc. IEEE*, vol. 80, no. 12, pp. 1870–1891, 1992, doi: 10.1109/5.192069.

[20] L. Pereira and N. Nunes, “Performance evaluation in non-intrusive load monitoring: Datasets, metrics, and tools—A review,” *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 8, no. 6, pp. 1–17, 2018, doi: 10.1002/widm.1265.

[21] N. Linge, X. Liu, Q. Liu, and M. Lu, “Non-intrusive load monitoring and its challenges in a NILM system framework,” *Int. J. High Perform. Comput. Netw.*, vol. 14, no. 1, p. 102, 2019, doi: 10.1504/ijhpcn.2019.10021110.

[22] M. Sun, F. M. Nakoty, Q. Liu, X. Liu, Y. Yang, and T. Shen, “Non-Intrusive Load Monitoring System Framework and Load Disaggregation Algorithms: A Survey,” in *International Conference on Advanced Mechatronic Systems, ICAMechS*, Aug. 2019, vol. 2019-Augus, pp. 284–288, doi: 10.1109/ICAMechS.2019.8861646.

[23] Z. Wang and G. Zheng, “Residential appliances identification and monitoring by a nonintrusive method,” *IEEE Trans. Smart Grid*, vol. 3, no. 1, pp. 80–92, 2012, doi: 10.1109/TSG.2011.2163950.

[24] K. D. Anderson, M. E. Berges, A. Ocneanu, D. Benitez, and J. M. F. Moura, “Event detection for Non Intrusive load monitoring,” *IECON Proc. Ind. Electron. Conf.*, pp. 3312–3317, 2012, doi: 10.1109/IECON.2012.6389367.

[25] Hernández, A. Ruano, J. Ureña, M. G. Ruano, and J. J. Garcia, “Applications of NILM Techniques to Energy Management and Assisted Living,” *IFAC-PapersOnLine*, vol. 52, no. 11, pp. 164–171, 2019, doi: 10.1016/j.ifacol.2019.09.135.

[26] K. S. Barsim and B. Yang, “Toward a semi-supervised non-intrusive load monitoring system for event-based energy disaggregation,” *2015 IEEE Glob. Conf. Signal Inf. Process. Glob. 2015*, pp. 58–62, 2016, doi: 10.1109/GlobalSIP.2015.7418156.

[27] J. Z. Kolter and M. J. Johnson, “REDD : A Public Data Set for Energy Disaggregation Research,” *SustKDD Work.*, no. 1, pp. 1–6, 2011, [Online]. Available: http://users.cis.fiu.edu/~lzhen001/activities/KDD2011Program/workshops/WKS10/doc/SustKDD3.pdf.

[28] A. Uttama, A. Reyes Lua, and R. V. Prasad, “LocED: Location-aware Energy Disaggregation Framework Categories and Subject Descriptors,” *2nd ACM Conf. Embed. Syst. Energy-Efficient Build. - BuildSys ’15*, no. ii, pp. 45–54, 2015, doi: 10.1145/2821650.2821659.

[29] N. Goddard *et al.*, “IDEAL Household Energy Dataset,” *University of Edinburgh. School of Informatics*, 2020. https://datashare.is.ed.ac.uk/handle/10283/3647 (accessed Dec. 19, 2020).

[30] N. Batra *et al.*, “NILMTK: An open source toolkit for non-intrusive load monitoring,” *e-Energy 2014 - Proc. 5th ACM Int. Conf. Futur. Energy Syst.*, pp. 265–276, 2014, doi: 10.1145/2602044.2602051.

[31] N. Batra, J. Kelly, and O. Parson, “NILMTK: Open source NILM toolkit,” *GitHub*. http://nilmtk.github.io/ (accessed Dec. 19, 2020).

[32] J. D. Kelly, “Disaggregation of Domestic Smart Meter Energy Data,” University of London, Imperial College of Science, Technology and Medicine, 2016.

[33] N. Batra *et al.*, “Towards reproducible state-of-the-art energy disaggregation,” pp. 193–202, 2019, doi: 10.1145/3360322.3360844.

[34] J. Kelly and W. Knottenbelt, “Metadata for energy disaggregation,” *Proc. - IEEE 38th Annu. Int. Comput. Softw. Appl. Conf. Work. COMPSACW 2014*, pp. 578–583, 2014, doi: 10.1109/COMPSACW.2014.97.

[35] A. Uttama, A. Reyes Lua, and R. V. Prasad, “DRED Data Set,” *TU Delft*, 2015. http://www.st.ewi.tudelft.nl/akshay/dred/.

[36] J. Liang, S. K. K. Ng, G. Kendall, and J. W. M. Cheng, “Load signature studypart I: Basic concept, structure, and methodology,” *IEEE Trans. Power Deliv.*, vol. 25, no. 2, pp. 551–560, 2010, doi: 10.1109/TPWRD.2009.2033799.

[37] A. Cominola, M. Giuliani, D. Piga, A. Castelletti, and A. E. Rizzoli, “A Hybrid Signature-based Iterative Disaggregation algorithm for Non-Intrusive Load Monitoring,” *Appl. Energy*, vol. 185, no. December 2019, pp. 331–344, 2017, doi: 10.1016/j.apenergy.2016.10.040.

[38] Arne Kaas, “MakeTheMeterSmarter.com – Open Source Disaggregation initiative,” *GitHub*, 2018. https://makethemetersmarter.com/disaggregation/.

[39] J. Kelly, “NILM Metadata 0.2.0 documentation,” *GitHub*, 2014. https://nilm-metadata.readthedocs.io/en/latest/index.html (accessed Dec. 19, 2020).

[40] A. Fiol, “Algorithms for Energy Disaggregation,” p. 47, 2016, [Online]. Available: https://upcommons.upc.edu/bitstream/handle/2117/89937/108724.pdf.

[41] A. Zoha, A. Gluhak, M. Nati, and M. A. Imran, “Low-power appliance monitoring using Factorial Hidden Markov Models,” *Proc. 2013 IEEE 8th Int. Conf. Intell. Sensors, Sens. Networks Inf. Process. Sens. Futur. ISSNIP 2013*, vol. 1, no. May 2014, pp. 527–532, 2013, doi: 10.1109/ISSNIP.2013.6529845.

[42] Raspberry Pi Foundation, “Raspberry Pi,” *Raspberry Pi Foundation*, 2016. https://www.raspberrypi.org/ (accessed Dec. 19, 2020).

[43] InfluxData, “InfluxDB: Purpose-Built Open Source Time Series Database,” *InfluxData*, 2020. https://www.influxdata.com/ (accessed Dec. 19, 2020).

[44] S. Biansoongnern and B. Plungklang, “Non-Intrusive Appliances Load Monitoring (NILM) for Energy Conservation in Household with Low Sampling Rate,” *Procedia Comput. Sci.*, vol. 86, no. March, pp. 172–175, 2016, doi: 10.1016/j.procs.2016.05.049.

[45] NILM community, “Appliance power demand | NILM,” *NILM Wiki*, 2020. http://wiki.nilm.eu/appliance.html (accessed Dec. 19, 2020).

[46] Qubidia, “How do I measure my refrigerator?,” *Qubidia*, 2020. https://www.qubidia.com/index.php/en/blog/56-how-do-i-measure-my-refrigerator (accessed Dec. 19, 2020).

[47] D. Christensen, “NILM Applications for the Energy-Efficient Home,” *NREL Natl. Renew. Energy Lab.*, no. June 2019, 2012.

[48] B. Najafi, S. Moaveninejad, and F. Rinaldi, “Data Analytics for Energy Disaggregation: Methods and Applications,” *Big Data Appl. Power Syst.*, no. October, pp. 377–408, 2018, doi: 10.1016/B978-0-12-811968-6.00017-6.

[49] D. A. Kelly, “Disaggregating Smart Meter Readings using Device Signatures by Submitted in partial fulfilment of the requirements for the MSc Degree in Computing Science of Imperial College London,” no. January 2011, 2016, doi: 10.13140/RG.2.1.1475.3043.

[50] S. S. Hosseini, K. Agbossou, S. Kelouwani, and A. Cardenas, “Non-intrusive load monitoring through home energy management systems: A comprehensive review,” *Renew. Sustain. Energy Rev.*, vol. 79, no. April, pp. 1266–1274, 2017, doi: 10.1016/j.rser.2017.05.096.

[51] M. Zhuang, M. Shahidehpour, and Z. Li, “An Overview of Non-Intrusive Load Monitoring: Approaches, Business Applications, and Challenges,” *2018 Int. Conf. Power Syst. Technol. POWERCON 2018 - Proc.*, no. November, pp. 4291–4299, 2019, doi: 10.1109/POWERCON.2018.8601534.

[52] K. Carrie Armel, A. Gupta, G. Shrimali, and A. Albert, “Is disaggregation the holy grail of energy efficiency? The case of electricity,” *Energy Policy*, vol. 52, pp. 213–234, 2013, doi: 10.1016/j.enpol.2012.08.062.

[53] D. Elizalde, “IoT Framework for Product Managers,” *Daniel Elizalde*, 2020. https://danielelizalde.com/iot-decision-framework/ (accessed Jun. 01, 2020).

[54] Grid4C, “Solutions Overview,” *Grid4C*, 2020. https://www.grid4c.com/solutions-overview#operationalAnalytics (accessed Dec. 19, 2020).

[55] Bidgely, “Solutions,” *Bidgely*, 2020. https://www.bidgely.com/solutions/ (accessed Dec. 19, 2020).

[56] Verv, “Peer-to-Peer renewable energy trading solution,” *Verv*, 2020. https://verv.energy/research (accessed Dec. 19, 2020).

[57] NILM Wiki, “Companies | NILM,” *NILM Wiki*, 2020. http://wiki.nilm.eu/companies.html (accessed Dec. 19, 2020).

[58] Opower, “What is Opower,” *Oracle*, 2020. https://www.oracle.com/industries/utilities/products/what-is-opower.html (accessed Dec. 19, 2020).

[59] Smappee, “Smappee home,” *Smappee*, 2020. https://www.smappee.com/nl/ (accessed Dec. 19, 2020).

[60] PlotWatt, “Energy Efficiency Business,” *PlotWatt*, 2020. https://www.plotwatt.com/ (accessed Dec. 19, 2020).

[61] Frost & Sullivan, “Global Home Energy Management Systems (HEMS) Market,” *Frost & Sullivan*, 2019. .

[62] European Commission, “Study on energy prices , costs and their impact on industry and households,” 2020. doi: 10.2833/49063.

[63] Mordor Intelligence, “Energy Management Systems Market - Growth, Trends, Forecasts (2020 - 2025),” *Mordor Intelligence*, 2020. .

[64] M. Baechler and H. Hao, “Business Case for Nonintrusive Load Monitoring,” *U.S. Dep. Energy*, no. May, 2016.

[65] European Commission, “Smart grids and meters,” *European Commission*, 2019. https://ec.europa.eu/energy/node/83.

[66] Smart Energy International, “Smart meter penetration in North America will reach 81% by 2024,” *Smart Energy International*, 2019. https://www.smart-energy.com/industry-sectors/smart-meters/smart-meter-penetration-in-north-america-will-reach-81-by-2024/ (accessed Dec. 19, 2020).

[67] U.S. Energy Information Administration (EIA), “How many smart meters are installed in the United States, and who has them?,” *U.S. Energy Information Administration (EIA)*, 2020. https://www.eia.gov/tools/faqs/faq.php?id=108&t=3 (accessed Dec. 18, 2020).

[68] Eurostat, “Electricity price statistics - Statistics Explained,” *Eurostat*, 2020. https://ec.europa.eu/eurostat/statistics-explained/index.php/Electricity_price_statistics (accessed Dec. 18, 2020).

[69] ODYSSEE-MURE, “Electricity consumption per dwelling | Electricity dwelling,” *ODYSSEE-MURE*, 2020. https://www.odyssee-mure.eu/publications/efficiency-by-sector/households/electricity-consumption-dwelling.html (accessed Dec. 20, 2020).

[70] Statistics Netherlands - CBS, “Average annual consumption,” *Statistics Netherlands - CBS*, 2020. https://www.cbs.nl/en-gb/news/2020/10/energy-bill-170-euros-lower-this-year/average-annual-consumption (accessed Dec. 20, 2020).

[71] A. Faustine, N. H. Mvungi, S. Kaijage, and K. Michael, “A survey on non-intrusive load monitoring methodies and techniques for energy disaggregation problem,” *arXiv*, 2017.

[72] C. Fischer, “Feedback on household electricity consumption: A tool for saving energy?,” *Energy Effic.*, vol. 1, no. 1, pp. 79–104, 2008, doi: 10.1007/s12053-008-9009-7.

[73] L. Steg and J. I. M. de Groot, *Environmental Psychology: an introduction*. Wiley-Blackwell, 2018.

[74] L. Stronks, “Raising energy awareness : outlines for concrete techniques for analyzing the energy usage of residential and business consumers based on smart meter data,” Radboud University Nijmegen, 2018.

[75] Y. Riche, J. Dodge, and R. A. Metoyer, “Studying always-on electricity feedback in the home,” *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 3, pp. 1995–1998, 2010, doi: 10.1145/1753326.1753628.

[76] K. Weimer, “Behavior Change Towards a Sustainable Lifestyle-Effects of Interventions and Psychological Factors on Pro-Environmental Behaviors,” Mid Sweden University, 2019.

[77] L. Steg and W. Abrahamse, “How to promote energy savings among households: Theoretical and practical approaches,” *Househ. Energy Econ. Consum. Effic.*, no. January 2013, pp. 179–205, 2012.

[78] B. Karlin, J. F. Zinger, and R. Ford, “The effects of feedback on energy conservation: A meta-analysis,” *Psychol. Bull.*, vol. 141, no. 6, pp. 1205–1227, 2015, doi: 10.1037/a0039650.

[79] J. Froehlich, L. Findlater, and J. Landay, “The design of eco-feedback technology,” *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 3, no. January 2010, pp. 1999–2008, 2010, doi: 10.1145/1753326.1753629.

[80] R. Sokoloski, “Disaggregated Electricity Consumption: Using Appliance-Specific Feedback to Promote Energy Conservation,” California State University San Marcos, 2015.

[81] P. Zangheri, T. Serrenho, and P. Bertoldi, “Energy savings from feedback systems: A meta-studies’ review,” *Energies*, vol. 12, no. 19, 2019, doi: 10.3390/en12193788.

[82] T. Serrenho, P. Zangheri, and P. Bertoldi, *Energy Feedback Systems: Evaluation of Meta-studies on energy savings through feedback*. European Commission, 2015.

[83] K. Ehrhardt-Martinez, K. A. Donnelly, and J. A. Laitner, “Advanced metering initiatives and residential feedback programs: A meta-review for household electricity-saving opportunities,” *Am. Counc. an Energy Effic. Econ.*, 2010.

[84] J. Kelly and W. Knottenbelt, “Does disaggregated electricity feedback reduce domestic electricity consumption? A systematic review of the literature,” no. May, 2016, [Online]. Available: http://arxiv.org/abs/1605.00962.

[85] A. Gupta and P. Chakravarty, “Impact of Energy Disaggregation on Consumer Behavior,” *Bidgely*, 2013.

[86] L. Pereira, F. Quintal, M. Barreto, and N. J. Nunes, “Understanding the limitations of eco-feedback: A one-year long-term study,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7947 LNCS, pp. 237–255, 2013, doi: 10.1007/978-3-642-39146-0_21.

[87] A. Gerster and M. A. Andor, “Disaggregate Consumption Feedback and Energy Conservation,” *CRC TR 224 Discuss. Pap. Ser.*, no. 182, 2020.

APPENDIX I

Energy Disaggregation Modelling

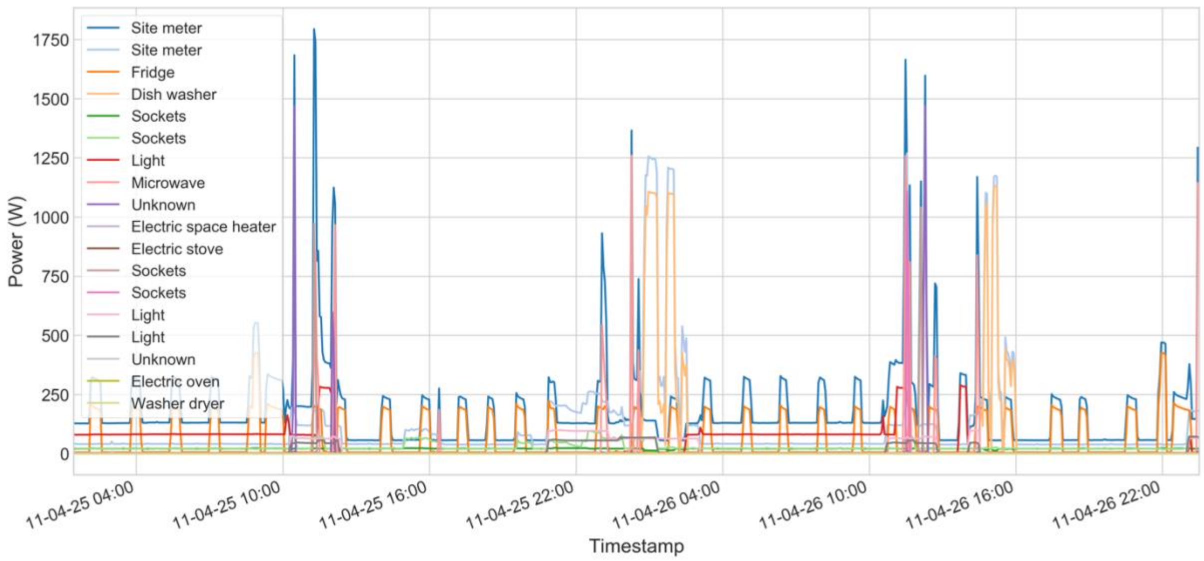


Figure A.1 REDD dataset: Three days aggregated electricity consumption (Site Meters) and individual appliances' consumption between the 25th and the 27th April 2011, Boston, Massachusetts (USA).

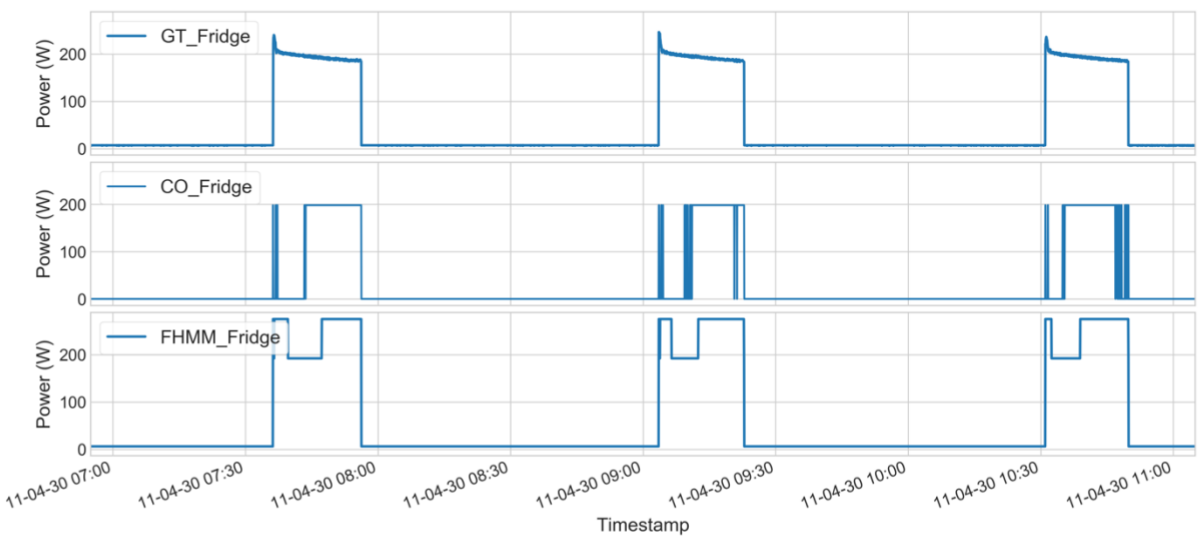


Figure A.2 REDD disaggregation results: 32% of the dataset trained. Fridge's disaggregation with CO and FHMM algorithms compared to the ground truth actual energy consumption during the 30th of April 2011.

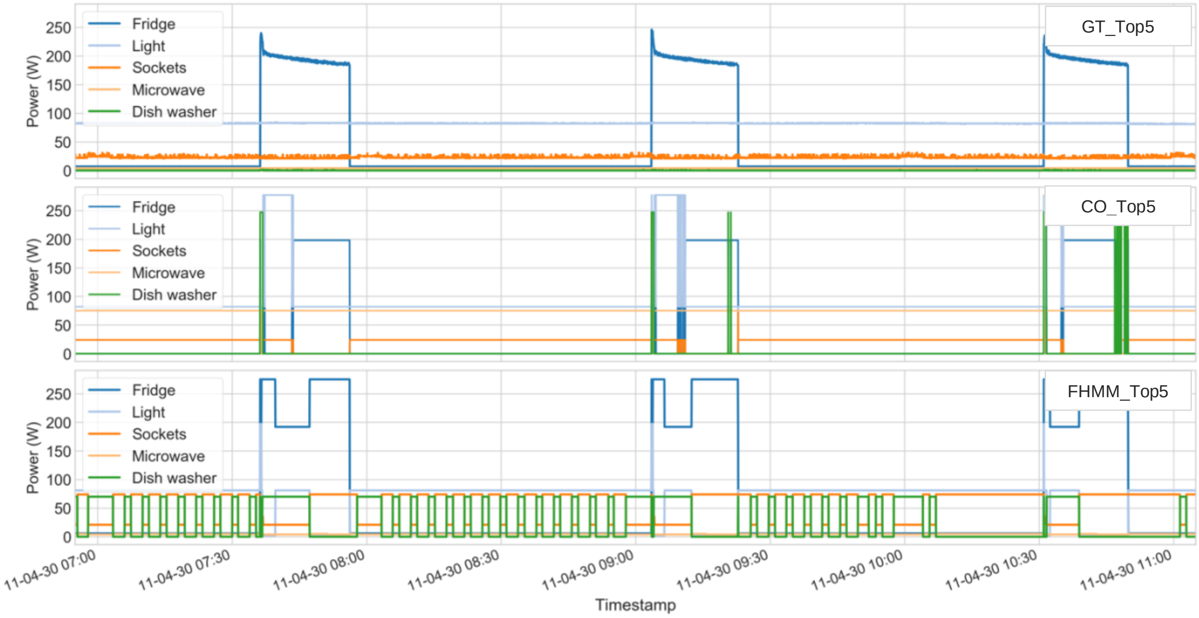


Figure A.3 REDD disaggregation results: 32% of the dataset trained. Most five consuming appliances' disaggregation with CO and FHMM algorithms compared to the ground truth actual energy consumption during the 30th of April 2011.

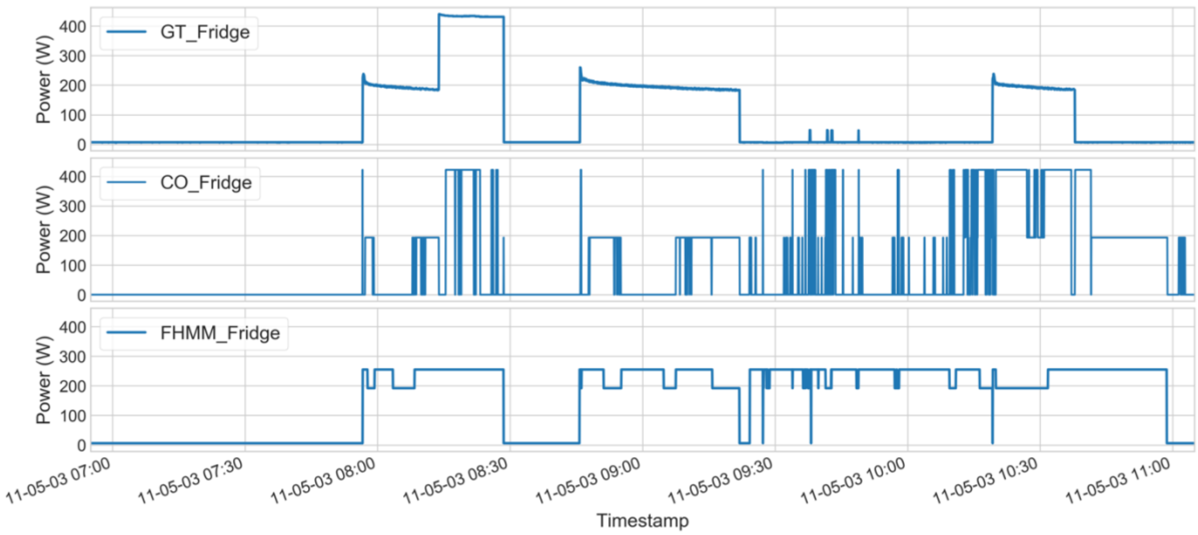


Figure A.4 REDD disaggregation results: 42% of the dataset trained. Fridge's disaggregation with CO and FHMM algorithms compared to the ground truth actual energy consumption during the 3rd of May 2011.

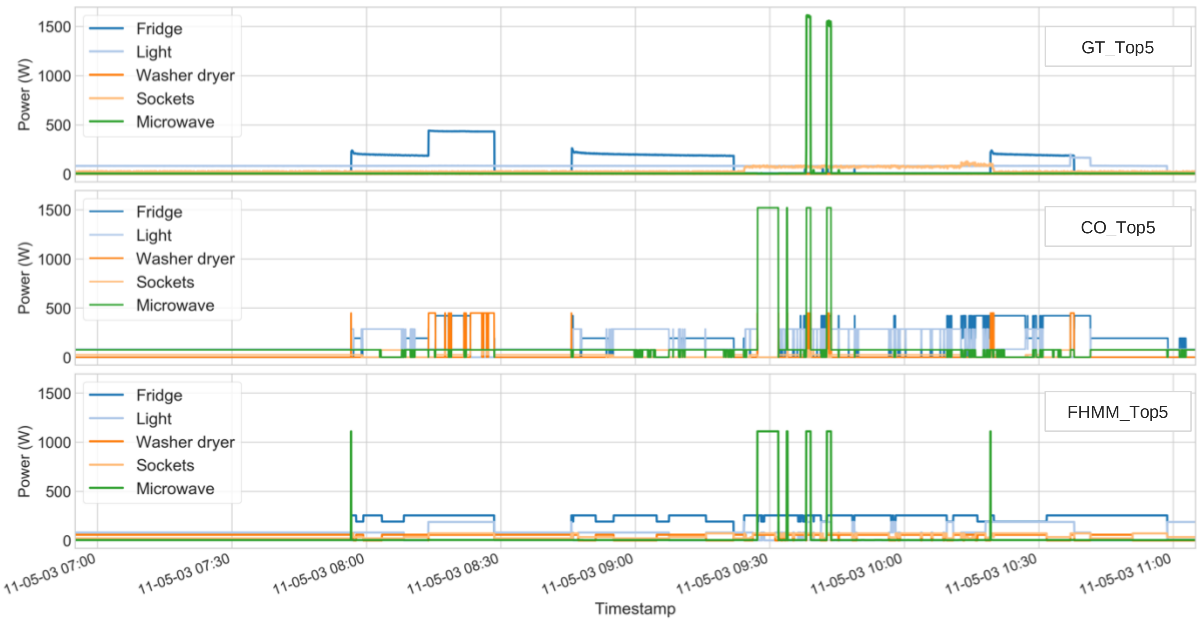


Figure A.5 REDD disaggregation results: 42% of the dataset trained. Most five consuming appliances' disaggregation with CO and FHMM algorithms compared to the ground truth actual energy consumption during the 3rd of May 2011.

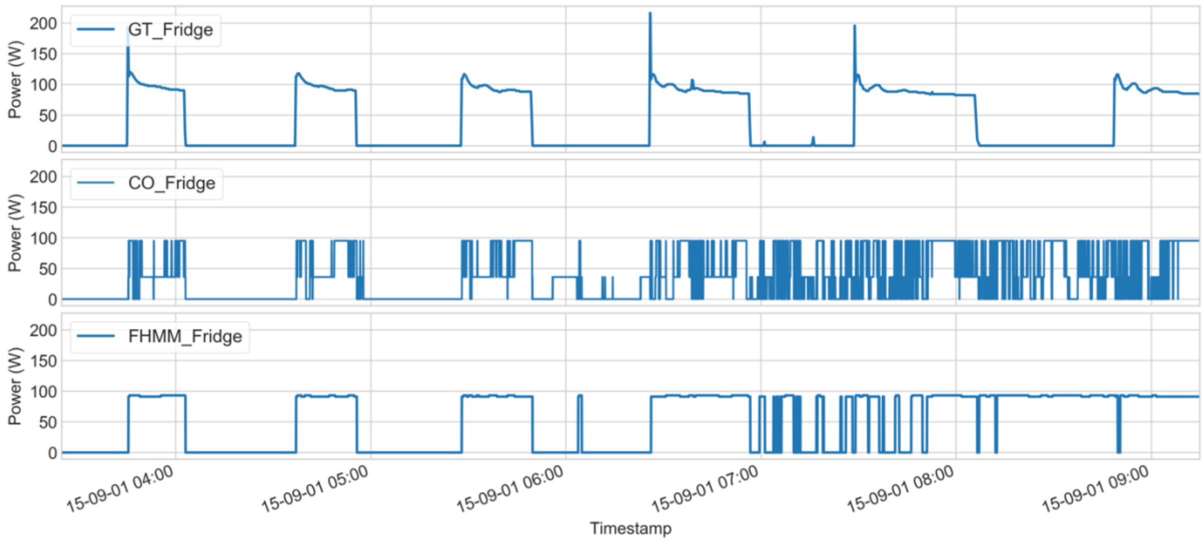


Figure A.6 DRED disaggregation results: 32% of the dataset trained. Fridge's disaggregation with CO and FHMM algorithms compared to the ground truth actual energy consumption during the 1st of September 2015.

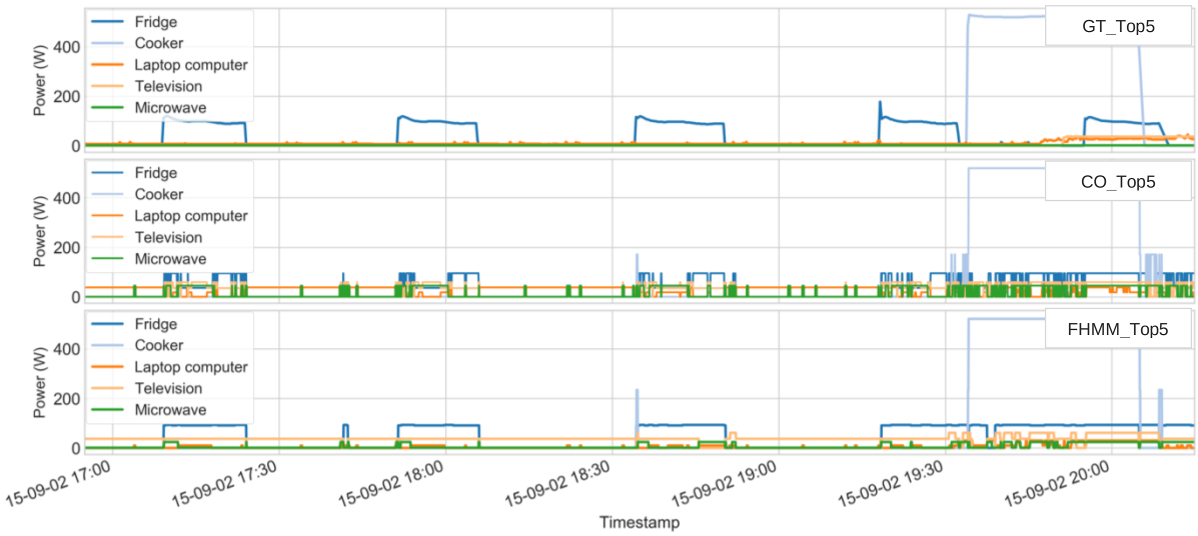


Figure A.7 DRED disaggregation results: 32% of the dataset trained. Most five consuming appliances' disaggregation with CO and FHMM algorithms compared to the ground truth actual energy consumption during the 2nd of September 2015.

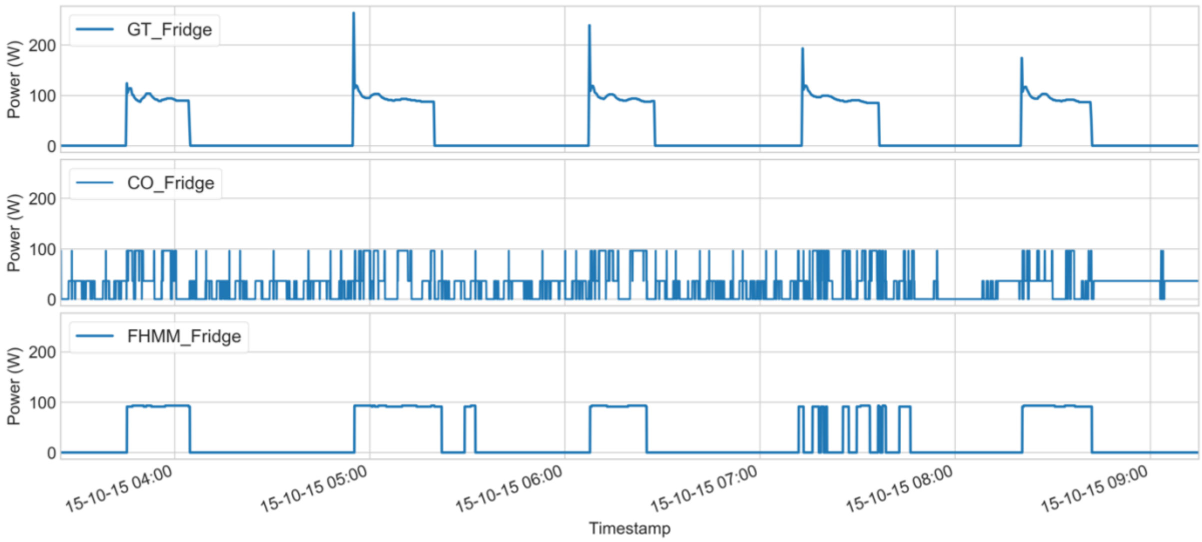


Figure A.8 DRED disaggregation results: 60% of the dataset trained. Fridge's disaggregation with CO and FHMM algorithms compared to the ground truth actual energy consumption during the 15th of October 2015.

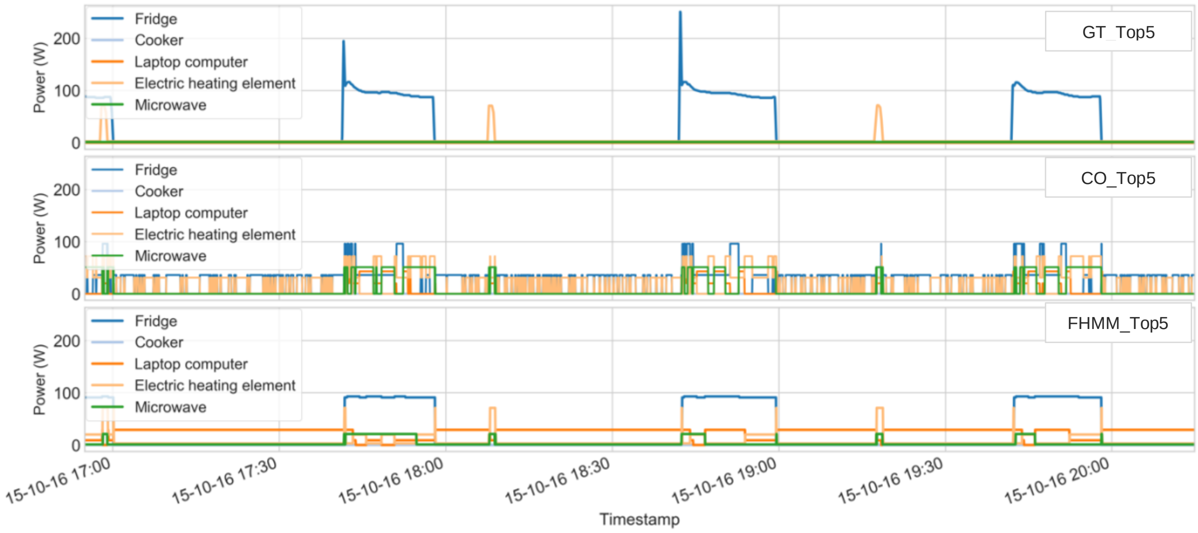


Figure A.9 DRED disaggregation results: 60% of the dataset trained. Most five consuming appliances' disaggregation with CO and FHMM algorithms compared to the ground truth actual energy consumption during the 16th of October 2015.

APPENDIX II

Market Potential

Table B.1 Economic Analysis of the NILM technology in USA, Europe and the Netherlands

	USA	EUROPE	NETHERLANDS
Average electricity consumption per household / year [kWh] [64][69][70]	10.837	4.000	2.450
Energy savings considered from NILM	0,05%	0,05%	0,05%
Electricity saved per household / year [kWh]	542	200	123
Total number of Households	248.750.000	289.610.390	7.500.000
Smart Meter rollout [8][66]	80%	77%	95%
Number of Households with NILM [8][66]	19.900.000	223.000.000	7.125.000
Electricity saved / year [TWh]	11	45	1
Price of Electricity [€][68]	0,1	0,216	0,206
Money saved / year [billion €]	1,078	9,634	0,179
Money saved per household / year [€]	54,19	43,20	25,17
Cost of NILM [€] [64]	500	500	500
Return of Inverstment (ROI) [years]	9	12	20