NILM-based approach for energy efficiency assessment of household appliances

Fernando D. Garcia 1, Wesley A. Souza 2, Ivando S. Diniz 1 and Fernando P. Marafão 1*

*Correspondence: fernando.marafao@unesp.br
1 São Paulo State University (Unesp), Institute of Science and Technology, Avenida Três de Março, 511, Alto da Boa Vista, 18087-180 Sorocaba, Brazil
Full list of author information is available at the end of the article

Abstract
This paper presents a novel Non-Intrusive Load Monitoring (NILM) approach focusing on the Energy Efficiency (EE) assessment of residential appliances. This approach (NILMEE) is able to identify the individual consumption of several household devices, providing proper information for evaluating energy efficiency and pointing out the operational issues or labelling mismatches of appliances, while recommending better practices for energy usage in specific consumer installations. The proposed approach was developed and evaluated by embedding the NILM engine on an electronic power meter, which performs a microscopic analysis on measured voltages and currents and provides the load disaggregation using the Conservative Power Theory for the feature extraction, K-Nearest Neighbours for the appliance classification, and the Power Signature Blob for the energy disaggregation. The disaggregation algorithm performance evaluation is carried out using NILMTK. Results show that NILM transcends the regular energy usage calculation, serving as a tool that enables the diagnosis of household appliances using the energy efficiency indexes provided by labels and standards.

Keywords: Load disaggregation, Energy management, Energy efficiency, Smart meter, Non-intrusive load monitoring

Introduction
Worldwide, energy efficiency policies can bring significant advantages to energy consumers and suppliers, leading to social, environmental, and economic benefits. Energy Efficiency (EE) may be an important resource to increase energy supply, as the first response to bigger demands and economic development (WEC 2013; IEA 2016). Therefore, the growing concern about improving EE has led to the creation of many new technologies, such as those related to energy management to control peak energy demand, better use of home appliances and the development of increasingly more efficient home appliances. Trotta (2018). These may be effective actions to optimise modern smart grid operations, supporting both consumers and electric utilities toward the energy efficiency perspective. Moreover, considering that residential consumers may represent over 30% of the energy consumption in some countries, household appliances may
interfere with the overall energy efficiency. Special attention has been paid to heating and cooling devices, leading to energy-consuming loads in residential installations (Pino-Meijas et al. 2017).

Thus, improvements in EE in residential installations may be essential for increasing the overall efficiency of the power grid and, traditionally, it basically relies on providing detailed information on the energy usage and efficiency of individual household appliances (Jin et al. 2017). The most common approaches are based on energy labelling initiatives, which serve not only as guides for consumers but also to stimulate innovative product development for complying with lower rates of energy consumption (Issock et al. 2018). Energy efficiency labels help consumers assess the appliance’s classification under particular operating conditions, such as those defined in standards and national programs (Wong and Krüger 2017; Merini et al. 2019). However, recent studies show that the labelling approach may not be enough for achieving expected goals in different countries (EC 2019) and further initiatives are required.

Strategies for increasing energy efficiency may be based on Non-Intrusive Load Monitoring (NILM) (Hart 1992), which was introduced decades ago, however the reduced computational power may have limited its practical applications (Makonin 2012). Considering modern computing engines and tools, NILM has attracted widespread attention, and the recent efforts focus on the proper use of artificial intelligence to create solutions for home and industrial energy management systems, demand-side management, and identifying appliances (Hart 1992; Aboulian et al. 2019; García et al. 2017). NILM represents a powerful tool to disaggregate the energy consumption of electrical installations, which can be embedded in modern energy meters (Souza et al. 2019; Zhang et al. 2018), providing information for assessing energy consumption in terms of household appliances.

Considering the load consumption disaggregation and based on the increasing energy-awareness of individual equipment, consumers may adapt consumption behaviours, replace equipment or install management systems focusing on energy/money savings (Baets et al. 2017; Mack et al. 2019), either on residential, commercial or industrial scenarios (Sadeghianpourhamami et al. 2017; Henriet et al. 2018; Stankovic et al. 2016). Recent NILM studies have been based on different attribute extraction methods, accuracy evaluations, and load disaggregation results ranging from 70% to 98% (Souza et al. 2019; Sadeghianpourhamami et al. 2017; Esa et al. 2016; Wong et al. 2013; Abubakar et al. 2015; Le and Kim 2018; Aladesanmi and Folly 2015; Klemenjak 2018; Bao et al. 2018; Gopinath et al. 2020).

However, most studies in the NILM literature are focused only on the analysis of the energy disaggregation concept and do not take into account the correlation with energy efficiency (Sadeghianpourhamami et al. 2017; Ruano et al. 2019). Few studies in the literature indicate the use of the NILM approach for EE assessment in the household scenario (Kong et al. 2020) and some of them focus on indicating high-consumption loads (Doherty and Trenbath 2019; Kong et al. 2020; Souza et al. 2020), feedback and recommendations to the consumer (Carrie Armel et al. 2013; Berbakov et al. 2019; Kong et al. 2020), and anomaly detection regarding the default appliance’s power signature (Rashid et al. 2019). However, there are no studies that compare the appliance EE label indexes with short and long-time use of operation. Hence, with this gap in mind, this paper introduces a novel NILM-based approach called NILMEE (NILM for EE), which can
disaggregate the energy demand of appliances of interest, while analysing the EE of each appliance in terms of a national labelling program.

The developed strategy was mainly based on the K-Nearest Neighbours (KNN) algorithm (Cover and Hart 1967) to identify the load and the Conservative Power Theory (CPT) (Tenti et al. 2011) as the background for the feature extraction. Moreover, this paper proposes a novel and oversimplified approach for energy disaggregation based on the Power Signature Blob (PSB) introduced by Souza et al. 2019.

This paper differs from previous proposals by introducing the use of NILM for energy efficiency assessment to detect appliances’ consumption deviations while providing maintenance or proper usage recommendations for household consumers. The proposed NILMEE engine was embedded on a smart meter prototype, indicating how such devices would perform the microscopic analysis of household appliances, thus being much more useful than merely for monthly energy measuring or on-off control.

The following sections briefly review the NILM approach, describe the evolution of the PSB concept, show the EE assessment and performance analysis of the NILM purpose, present the NILMEE validation, and depict the results of NILMEE applications in a case study of a real household.

**NILM background and PSB implementation**

Figure 1 depicts a classic example of NILM. This scenario shows the total power drawn versus time, in which different appliances are operating. In this perspective, each appliance has a specific power signature.

In 1985, Hart (1985) conducted a pioneering study on load disaggregation defined as Nonintrusive Appliance Load Monitoring (NALM) – NILM is a derivative term from NALM. This approach was designed to react according to plug and unplug events; i.e., when the appliance is turned on or off. This behaviour enables the design of Finite State Machines (FSM) to track appliance states based on active and reactive power variations (Hart 1992). However, some types of appliances operate in more than one state, and others operate continuously between distinct states, making it difficult to model an FSM accurately (Azaza and Wallin 2017). Recent research efforts have been focused on affordable alternatives for FSM or artificial intelligence (Souza et al. 2019; Sadeghianpourhamami...
As the NILM engine depends directly on electrical quantities, some research efforts have given attention to the NILM hardware (Aboulian et al. 2019; Le and Kim 2018; Klemenjak 2018), usually called Smart Meter or Cognitive Meter (Garcia et al. 2017; Souza et al. 2019; Makonin et al. 2013), or Energy Box (Donnal et al. 2016; Zhuang et al. 2018). Bringing processing to the edge-device, embedded NILM solutions (such as the off-line NILM) enable pre-processing of electrical quantities at the acquisition time (Hamid et al. 2017; Huchtkoetter and Reinhardt 2019). Thus, recent studies investigated the use of Power Theories to develop distinct mechanisms to the load identification: CPT and its application in Souza et al. 2018, 2019, and Fryze’s Power Theory and its application in Teshome et al. 2016.

In order to accurately determine the appliances’ behaviour and load disaggregation, the acquisition of voltage and current signals with standardised data resolution and the sampling rate is necessary and this typically represents a key point for defining minimum hardware requirements to attend the desired NILM performance. Thus, referring to the sample rate, three classes of hardware can support NILM implementation (Esa et al. 2016): macroscopic (Low-Frequency), microscopic (High-frequency), or frequency-independent (event analysis, probability indexes, among others) (Esa et al. 2016; Rottondi et al. 2019; Huchtkoetter and Reinhardt 2019; Bouhouras et al. 2017; Haq and Jacobsen 2018).

Considering the NILM as the critical process to boost energy efficiency actions, the next section presents an improved and embedded version of the PSB – a NILM process that relies on CPT features and in a supervised machine learning algorithm to identify and disaggregate energy usage of appliances.

### Power signature blob

The PSB consists of six stages as depicted in Fig. 2.

(a) and (b) **Data acquisition and processing**

This stage collects voltage and current samples and performs the power term calculation based on CPT, as defined in Tenti et al. 2011.

To provide a more consistent power signature for the edge detection algorithm, a voltage-independent normalised power model $P_{\text{norm}}(t)$ was used, as presented in Eq. 1. The “normalised power” is used to adjust the $P(t)$ by a constant scale, approximating this...
value to the one measured if the utility provides a constant grid nominal voltage $V_s$ (Hart 1992).

$$P_{\text{norm}}(t) = \left( \frac{V_s}{V(t)} \right)^{\beta} P(t)$$  \hspace{1cm} (1)

(c) **Power edge detection**

Power edge detection plays an essential role in NILM event-based implementation (Hart 1992), as it defines the moment to obtain the features used for appliance classification. In general, three appliances' models need to be identified by the edge detection algorithm as pointed out below:

I Binary: Appliances that only have two states: off or on, consuming constant power such as an electric resistance oven.

II Finite State Machine: Appliances that have multiple distinct states mainly characterised to be a composed appliance, such as a refrigerator with automatic defrost.

III Continuously Variable: Appliances that during operation are controlled by electronic devices that vary energy consumption, such as a split air conditioner.

Differently from the FSM used for event detection in Souza et al. 2019, this work uses the windowing concept for the power signal segregation through a shifting set of samples representing stationary and transient states, as presented in Fig. 3.

The RMS void current calculated from the CPT is used to identify the transient states. By definition, the void current parcel presents a residual value resulting from the difference of the total current with the active and reactive parcels. Moreover, the void current indicates the existence of non-linear or harmonic components in the load (Tenti et al. 2011). Thus, the void current becomes a potential event indicator for considering that a transient state causes an instantaneous disturbance on the electric current.

First, the absolute difference of RMS void current greater than a threshold indicates a transient point. The threshold value prevents false events caused by the intrinsic noises of the acquisition system. The time interval between consecutive points where the RMS void current difference is above the threshold defines the transient state region. Finally, if
the amplitude of the RMS active power in the transient state region is above the minimum threshold for appliance detection, the transient should be processed as a valid ON/OFF event.

In summary, the power edge detection stage results in a sequence of positive or negative peaks, as depicted in Fig. 4. This step triggers the feature extraction stage for the appliance recognition, as presented in stage (d).

(d) Features extraction
Feature extraction depends on the NILM approach: macroscopic or microscopic feature extraction. In general, macroscopic NILM monitors changes in the power signature, while microscopic NILM uses waveform details to extract features for the load disaggregation (Esa et al. 2016; Bouhouras et al. 2017; Haq and Jacobsen 2018).

In this work, the RMS values of active power, active, reactive, and void current compose the set of features for appliance classification.

The feature extraction occurs for every transient and uses the data of the contiguous steady states. Thus, the features correspond to the absolute difference between the mean value of each attribute in these states.

(e) Appliance classification
The classification uses the features from (d) to track the appliance that generated the event. Regardless of the algorithm, the primary idea of classification is the association of the appliance and its state.

In this study, a supervised algorithm with previous knowledge of the loads identifies the appliances using the four CPT effective attributes — active power, active, reactive and void current — saved in a local database.

By using CPT attributes in a four-dimensional space, each appliance creates a cluster and the KNN associates the event’ attributes with the nearest cluster in this 4-dimensional space.

(f) Appliance energy disaggregation
The total power draw represents the sum of the individual appliance power during its operation. Considering that $P_i$ is a vector that contains all the information of power that

---

1Available on: github.com/FernandoDelunoGarcia/nilmee-energyinformatics
the \( i \)-th appliance consumes and admitting that a small noise could corrupt the signal, then the aggregated power at time \( t \) (Hart 1992) is generalised as presented in Eq. 2a.

\[
P(t) = \sum_{i=1}^{n} (a_i(t) \times P_i(t)) + e(t)
\]  

(2a)

\[
a_i(t) =
\begin{cases} 
1 & \text{if appliance } i \text{ is on at } t; \\
0 & \text{if appliance } i \text{ is off at time } t.
\end{cases}
\]  

(2b)

Considering that the previous steps define the moment (Power Edge Detection) and which appliance (Appliance Classification) generated the event, the average value of active power and the appliance type can determine the consumption approximation. To do this, it is necessary to form groups of events that correspond to the period of appliance operation.

Considering every peak as a new event, this work proposes using a database to store samples of active power during the period that the appliance is on. It means that every appliance will produce a batch of samples (Blob (Souza et al. 2019)) that represent its operation history.

To obtain a more approximated value of appliance energy consumption, the disaggregation algorithm extracts the mean value of each Blob associated with each class to evaluate the average of active power during that period. Blobs that contain more than one activated appliance use the historical value of each one as depicted in Fig. 5.

**Assessment of energy efficiency - beyond NILM**

A number of different regulations have been proposed worldwide, aiming at energy efficiency practices to strike on electrical energy waste (WEC 2016). An example is the development and application of labels that determine the energy efficiency of appliances and manage the development of innovative products, complying with new specifications, thus serving as a guide for consumers (WEC 2016).

Relying on the application of the load disaggregation method and handling the information from the appliances EE label, the proposed NILMEE system can correlate this information to the real consumption, in order to provide users with a proper energy efficiency assessment. Thus, the NILMEE method is able to: calculate and compare real measurements to those indices stated in the appliance EE labels; point to the appliance

![Fig. 5](image-url) Disaggregation using a batch of samples associated with each classified appliance
operating issues; propose ways of improving the appliance’s performance, among other 
EE recommendations.

Therefore, in order to show how NILM can be merged to the EE labelling program, the 
next section discusses some information from the most important program on energy 
efficiency in Brazil, the National Electricity Conservation Program (PROCEL) (Eletrobras 
2019).

Energy efficiency of household appliances

PROCEL is based on the Brazilian Labelling Program (PBE), which aims to provide 
information on the performance of products and appliances, mainly concerning energy 
efficiency (Eletrobras 2016). PBE labels provide a classification that can go from letter A 
for the most efficient equipment to letter E for the least efficient ones (Eletrobras 2016), 
as depicted in Fig. 6.

Taking into account the PBE and PROCEL, the following sections describe important 
aspects and formulations used for proposing the NILMEE approach and useful for pro-
viding further recommendations to consumers focusing on two high energy demanding 
household appliances: refrigerators and microwaves.

Refrigerator

Refrigerators are among the main targets for labelling programs in the world (WEC 2016). 
In this study, top-freezer refrigerators are nominated as regular refrigerators, and they 
consist of resistive and inductive elements from the equivalent circuit point of view. There 
are basically two operating conditions: the automatic freezer defrost stage that can be

![Fig. 6 Example of a PBE label](image)
represented by a resistive element; and the refrigeration stage that can be represented by a resistive-inductive device (compressor).

The efficiency index of refrigerators represents the ratio between energy usage measured in the laboratory (C) and regular energy usage (Cp) (INMETRO 2006).

In Brazil, the Cp was established by statistical analysis taking into account the universe of products labelled in 2001 (INMETRO 2006). PROCEL performed the labelling of these appliances, and the tests are carried out according to ISO 7371, 8187, 5155 and 8561 (CLASP 2003). Moreover, the maximum levels of consumption (C/Cp) for refrigerators and freezers were defined by the Interministerial Ordinance 326 of 2011 (MCT/MME 2011).

Thus, having the disaggregated power of the refrigerator, the energy efficiency indexes can be evaluated. Moreover, the record of the refrigerator activity log enables appliance diagnosis such as the duration of the refrigeration cycles.

Taking this information from the proposed NILM engine, several procedures can be created to notify the consumer about the days when energy consumption was higher than the average, as well as to provide suggestions for reducing the energy consumption.

Thus, a supervisory system may notify consumers about appliances’ unusual operation or consumption, sending alerts or suggestions such as:

- Thermostat adjustment may be required (e.g., according to the season of the year or stored food);
- Possible issue with door rubber sealing - evaluation required;
- Avoid opening the door unnecessarily or for a long period of time;
- Avoid hot food in the refrigerator;
- Appliance consumption differs from original EE label.

**Microwave**

Microwaves are home appliances with a maximum power that is normally higher than 1000W. According to PROCEL, the useful power ($P_u$) of a microwave is determined in laboratory experiments (INMETRO 2012) and the Eq. 3 shows the calculation of the energy efficiency.

$$n = 100 \times \frac{P_u \left(t_{total} - t_{aq}\right)}{W_{in}}$$

where:

- $n$ : energy efficiency, expressed as a percentage;
$Pu$ : microwave output (useful) power in Watts;
$t_{total}$ : total test time, in seconds;
$t_{aq}$ : magnetron warm-up time, in seconds;
$W_{in}$ : energy consumption during power output test, in Wh (Watt hour).

Considering the disaggregated power of the microwave, magnetron warm-up time can be evaluated and then the energy efficiency index can be calculated. Besides, registering microwave activity enables us to diagnose the appliance such as the average of active power and time usage.

Thus, the supervisory system may notify consumers about appliances’ unusual operation or consumption, sending alerts or suggestions like:

- To obtain maximum performance and prevent overheating, do not obstruct the air outlet of the microwave;
- Always keep the internal surface of the oven clean for more efficient cooking.
- Take frozen food out of the freezer before use to avoid using a microwave to defrost the product;
- Appliance consumption differs from original EE label.

The following section outlines the steps towards load disaggregation and energy efficiency analysis for the equipment of interest, resulting in the implementation of NILMEE.
Experimental setup

A metering infrastructure presented in Fig. 7 was created to validate the proposed method. A cognitive meter was placed at the PCC for the overall load monitoring and smart plugs were used for connecting each load of interest, in order to be used as ground truth, i.e., as a real parameter for evaluating the algorithm performance. Note that NILMEE is implemented in the cognitive meter and this infrastructure was built to create a measurement dataset used in the NILMTK (Batra et al. 2014) – a framework used to analyse and validate the NILMEE.

 Acquisition of the measurement data

The cognitive meter architecture is depicted in Fig. 8. The processing unit was based on a “BeagleBone Black (BBB) rev. C”, which contains the application processor AM335x 1GHz ARM® Cortex-A8, with 512MB DDR3 RAM and 4GB 8-bit eMMC flash memories. In addition, the processor contains two programmable real-time units (PRU).

PRU is a 32-bit microcontroller that was used for data acquisition, because it shares a memory region with the application processor. This configuration enables the NILMEE algorithm to run on the main core, while the PRUs control the data acquisition from analog front-end (AFE), which is composed of a delta-sigma ADC (ADS131E08S) with simultaneous sampling of 8-channels with 24-bit resolution at 15,360 Hz.
Table 1 The description of the refrigerators used in the experimental results

| Refrigerator | Type                  | Capacity                        | Purchasing date |
|--------------|-----------------------|---------------------------------|-----------------|
| R1           | Double Door           | Fresh Food: 329 litres          | September 2017  |
|              | Frost Free Refrigerat.| Freezer: 100 litres             |                 |
| R2           | Double Door           | Fresh Food: 317 litres          | July 2008       |
|              | Frost Free Refrigerat.| Freezer: 86 litres              |                 |
| R3           | Double Door           | Fresh Food: 346 litres          | September 2012  |
|              | Frost Free Refrigerat.| Freezer: 113 litres             |                 |
| R4           | Frost Free Refrigerat.| Fresh Food: 236 litres          | December 2018   |
|              |                       | Freezer: 25 litres              |                 |

The cognitive meter software was configured to perform the signal acquisition of the current and voltage, generating the RMS voltage and current, active and apparent power, and the CPT current components: active, reactive, and void.

Moreover, two smart plugs are used on the plug of each load of interest to provide the ground truth and a personal computer was used as the central node of a wireless sensor network as well as a data concentrator. The cognitive meter sends the measurement data at a 40 Hz rate, while the smart plugs installed on loads of interest provide data at each second. This procedure is performed only to prepare the training data used in NILTMK, having comparative parameters for the proposed NLM performance.

Data analysis with NILMTK

The dataset preparation to integrate with NILMTK consists of the simple data aggregation of each energy meter. The data from the cognitive meter and the smart plugs are merged in an HDF5 (Hierarchical Data Format) file structure supported by NILMTK. The following steps for data analysis are performed with the available dataset:

1. The cognitive meter data is processed to identify stationary and transient states, as described in “(a) and (b) Data acquisition and processing” section
2. The transients’ timestamps are stored for further analysis
3. The stationary states are created between every pair of transient states using the timestamp as a reference
4. The measurements average values are calculated for the stationary states
5. For each transient state, the difference of the measurements average values considering adjacent stationary states is calculated, forming the set of features for the appliance classification stage
6. The individual ground truth measurements are used to associate the appliance label on the transients detected by the main meter
7. The classifier uses labelled events for next appliance classifications

Table 2 Comparison between the nominated and measured energy efficiency indexes of the refrigerators used in the experimental results

|                 | R1   | R2   | R3   | R4   | R4 (seal failure) |
|-----------------|------|------|------|------|-------------------|
| Labelled EE     | 0.779| 0.772| 0.746| 0.853| 0.853             |
| Evaluated EE    | 0.754| 0.645| 0.812| 0.802| 1.203             |
| Labelled kWh/month | 53  | 56   | 58   | 24.3 | 24.3              |
| Evaluated kWh/month | 51.3| 46.8 | 63.15| 22.87| 34.27             |
Table 3  Microwave description

| Microwave | Capacity | Purchasing date |
|-----------|----------|-----------------|
| MW1       | 30 litres | August 2013     |
| MW2       | 30 litres | June 2017       |
| MW3       | 28 litres | January 2008    |
| MW4       | 27 litres | March 2011      |

8 The disaggregation procedure comprises the identification of the “on and off” appliance events. The association of these events composes the disaggregation block used for appliance operation analysis.

In order to demonstrate the capability of the proposed approach for load disaggregation, as well as the appliances diagnosis and energy efficiency evaluation, the next sections present the experimental results of the NILMEE performance and two EE case studies with different appliances.

Experimental results

Following the experimental procedure based on NILMTK, this section describes the procedure for estimating the EE indexes for two loads of major interest: refrigerators and microwaves. The results of the proposed methodology applied on real data is also presented and discussed in terms of energy efficiency evaluation.

Initially, an exploratory analysis of data collected by the cognitive meter was performed to identify the appliances of interests, considering the operation of a single equipment at a time. Later, the NILMEE was validated on a real household scenario.

Energy efficiency assessment on individual appliances

Refrigerators

The refrigerator has a FSM power signature with three possible states. The defrost system (typically a resistive load) is activated in two moments over the 24 h, as illustrated in Fig. 9. The refrigeration cycles are equivalent to the compressor’s operation, and their duration may vary for the effort required to control the internal temperature. The defrosting events occurred with an interval of about 12 hours in the performed experiments. It was also
observed that duration of the refrigeration process is longer than the daily average after the defrost periods.

There are some variables that may also influence refrigerator consumption depicted in Fig. 10, such as the number of times that the refrigerator door was opened, if hot food was inserted into the refrigerator, among others.

Comparing the maximum and minimum temperatures of each day, it was observed that the external temperature may contribute to an increase in energy consumption, as indicated in Fig. 11.

Given the refrigerator EE index and Fig. 10 as the basis for the refrigerator’s daily energy consumption, the monthly test revealed that the average consumption is within the range specified by PROCEL and the manufacturer.

Moreover, it was observed that the number of refrigeration cycles is higher than the average on the days when the refrigerator had the highest consumption records. This increase in cycles has a correlation with temperature, as indicated in the Fig. 11 and with the number of occasions in which the door was opened, causing the loss of the internal cooling temperature.

Four refrigerators are used to enhance the EE assessment with a description of each one presented in Table 1. The EE index and the monthly consumption values reported in the PBE label of each refrigerator were used as parameters to compare them with the collected measurements, as presented in Table 2.

According to the results presented in Table 2, R1, R2, and R4 presented EE and consumption values under the limit established in the PBE label. R3 presented a value over the limit, indicating an EE issue. A probable reason for this issue is the refrigerator’s lifetime, which is over six years without maintenance, and the operation in the position of the highest consumption also needs to be considered. Another EE issue is detected on the

| Table 4 | Energy efficiency indexes found for the four microwaves evaluated in the experiments |
|---------|----------------------------------|
|         | MW1   | MW2   | MW3   | MW4   |
| \( t_{ag} \) (s) | 1.58   | 1.58   | 2.53   | 2.44   |
| EE from label (%) | 56.2   | 60.0   | -      | -      |
| EE from evaluation (%) | 55.9   | 59.75  | 57.64  | 53.97  |
| \( P_{\max} \) from label (W) | 1,350  | 1,350  | 1,400  | 1,450  |
| \( P_{\max} \) from evaluation (W) | 1,300  | 1,335  | 1,302  | 1,395  |

Fig. 13 Example of the disaggregation performed by the NILMEE
Table 5: Results achieved by the state detector to recognise the refrigerator and microwave

|                | TPR  | FPR  | F1-score | Accuracy |
|----------------|------|------|----------|----------|
| Refrigerator   | 0.969| 0.969| 0.969    | 0.982    |
| Microwave      | 0.896| 0.884| 0.890    | 0.999    |
| Micro-measure  | 0.991| 0.991| 0.991    | –        |
| Macro-measure  | 0.963| 0.960| 0.962    | –        |

R4 (door with seal failure) case, leading to the increase of more than 50% of the R4 power consumption.

**Microwaves**

Four microwaves were rated for this practical test. The description of each one is presented in Table 3.

The microwaves have two operation stages as depicted in Fig. 12. The first is the magnetron warm-up stage \( t_{\text{lag}} \), which can last up to two seconds; and the second is the microwave operation at full power according to the time set.

Several tests were carried out to evaluate the warm-up time. The results are based on devices operating during 30 s. The EE index and the maximum power values reported in the PBE label of each refrigerator were used as parameters to compare them with the collected measurements, as presented in Table 4. The EE indexes of MW1 and MW2 are in compliance and close to the limits indicated on the PBE label, and they do not present EE issues. MW3 and MW4 do not have the EE certification label, although MW4 has a low EE index compared with other models and replacement recommendations can be suggested to the consumer.

**Load disaggregation validation**

A database with two days was used for training to validate the experimental procedure, and the testing was performed in another dataset of four days with unseen data. The classifier was evaluated regarding the accuracy to determine when the loads of interest are turned on or off. The aim was to detect the loads of interest according to Fig. 13 and to evaluate the load disaggregation performance.

**Classification performance**

The classifier was evaluated using the true positive rate (TPR), false positive rate (FPR), F1-score and Accuracy obtained from the confusion matrix data. These terms are commonly used as performance indicators that use the values of true positives (TP), true negatives (TN), false negatives (FN), and false positives (FP). TP means that instances with a “positive” class are correctly classified as positive. FP means that instances with the “positive” class are incorrectly classified as positive, and should be in the negative class. TN means that some instances with the “negative” class are correctly classified as negative. FN means that some instances with the “negative” class should be classified as positive.

Table 6: Root mean square error of disaggregated data

| Appliance   | RMSE |
|-------------|------|
| Refrigerator| 49.37|
| Microwave   | 41.04|
TPR or recall is the ratio of correctly predicted positive observations for all observations in the actual class. The recall is calculated according to Eq. (4).

\[ TPR = \frac{TP}{TP + FN} \] (4)

FPR usually refers to the expectancy of the false positive ratio over all observations. Moreover, FPR is defined as the probability of falsely rejecting the null hypothesis. The FPR formula is depicted in Eq. (5).

\[ FPR = \frac{FP}{FP + TN} \] (5)

The Accuracy is the ratio of correct predictions by total predictions as depicted in Eq. (6).

\[ Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \] (6)

The F1-score takes false positives and false negatives into account. Intuitively it is not as easy to understand as accuracy, but the F1-score is usually more useful than the accuracy, especially if it has an uneven class distribution. The F-score is calculated according to Eq. (7).

\[ F_{1\text{score}} = \frac{2 \times TP}{2 \times TP + FP + FN} \] (7)

Table 5 presents the performance metrics\(^2\) of the experimental result for loads of interest in this work, where values close to 1 are equivalent to the near-perfect prediction performed through the method. Besides the metrics of each load, Table 5 also shows the macro and micro measures. A macro-measure computes the metric independently for each class and then takes the average, and a micro-measure computes the metric considering the number of samples in each class.

**Disaggregation accuracy**

The disaggregation accuracy was calculated for each load using the root mean square error (RMSE) between the estimated power (\(y\)) and the ground truth (\(\hat{y}\)) of appliance \(n\) for each period \(t\).

\(^2\)These metrics were calculated using NILMTK and the scikit-learn (Pedregosa et al. 2011) library. For an indication on or off state at a given time, both disaggregated and the ground truth data were converted to binary values using a threshold of 50 W.


**Table 7** Warm-up time of the microwave used in the EE evaluation

|    | \( t_{\text{eq}} \) (s) |
|----|-------------------------|
| min| 1.375                   |
| max| 1.625                   |
| avg| 1.571                   |

\[
\text{error}(n) = \sqrt{\frac{1}{T} \sum_{t} \left( y_{(n)}^{(t)} - \hat{y}_{(n)}^{(t)} \right)^2}
\]  

(8)

While the performance metrics assess the correct load event identification, the RMSE indicates the consumption integration error comparing the real consumption of the equipment with the one estimated by the NILM method. The results for the two loads of interest are presented in Table 6. The low RMSE and the F1-score achieved ensures the usage of the proposed system to monitoring appliances of interest.

**Energy efficiency assessment on a real household**

This section presents the NILMEE versatility in a real residential scenario, where several appliances are turned on and off over time. The analysis of the results focusing on the microwave and refrigerator are given in the next sections.

**Microwave analysis**

The microwave EE evaluation was carried out considering the average power value, the warming-up time, and the operating time determined by the NILMEE. Figure 14 depicts two moments when the microwave is turned on.

The warm-up time was determined according to the transient state period until the appliance’s power stabilisation. The total operation time was calculated between the transient stages, when the load is turned on and off. Table 7 shows the warm-up average time of the microwave and Table 8 presents the microwave EE index computed following the Eq. 3.

This analysis shows the success of NILMEE in identifying microwave consumption events and in computing the EE indexes. Information calculated by NILMEE indicates that the household microwave is in accordance with the values indicated on the PBE label.

**Refrigerator analysis**

Two tests were considered to evaluate the refrigerator EE. The first one was designed to check the average daily consumption over four days. Table 9 shows the achieved EE indexes for the first test.

In the second test case, the refrigerator is considered in a regular operation and after a door seal failure simulation. The simulation was carried out in approximately 4 hours, and the monitoring results point out the cooling cycle period increase. In a regular operation, the average time of the refrigeration cycle was approximately 246 seconds. If the door

**Table 8** Energy efficiency index of the microwave used in the evaluation

| Label | Test       |
|-------|------------|
| EE    | 60%        |
| Power | 1350 W     | 1340 W     |
Table 9  Energy efficiency index of the refrigerator used in the evaluation

|        | Label | Test |
|--------|-------|------|
| EE     | 0.772 | 0.645|
| kWh/month | 56    | 46.82|

Seal was violated, this average increased to 414 seconds. Figure 15 shows the increased operating time of a door seal failure situation.

In this test, the daily consumption was increased by 0.24 kWh. Considering that this difference in consumption presented in 4 hours remains linear for the whole day, the daily energy consumption would be greater than that indicated on the refrigerator’s energy efficiency label.

Moreover, NILMEE achieved good results in identifying refrigerator events and has been also able to point to operational issues, such as a non-sealed door, based on the comparison of estimated EE indexes and those from PBE labelling program.

Considering the existence of a supervisory system based on NILMEE and embedded on the energy meter, this would boost the interaction with the consumer, pointing out the most appropriate strategies for EE in the home, according to the appliances of interest. Considering the EE issues identified in Tables 2 and 9, NILMEE would return the appliance’s EE condition, and the supervisory system would present recommendations, e.g., in order to improve the refrigerator’s energy efficiency, such as: checking the door seal; avoid placing the refrigerator in sunny places; checking the thermostat regulator position; or consider changing long lifetime appliances. For the case of low EE shown in Table 4, the appliance replacement would be recommended by the supervisory system.

Conclusions

This paper presented a novel NILM-based approach focusing on the assessment of EE in household appliances. Using the so-called NILMEE method, adequate information can be provided to assess the EE, pointing out appliance operational issues or labelling mismatches, and recommending best practices for energy use or maintenance related to individual devices.

The proposed NILMEE technique was based on the PSB method, adopting a microscopic approach for load classification with power features calculated by the CPT. The
robustness of the applied hardware architecture provides the execution of NILM techniques in embedded systems. Moreover, the PSB event detection step was improved by tracking the CPT void current during transients.

Considering a national labelling program, practical results were presented for energy efficiency assessment of two typical household appliances, expanding the possibilities and literature on NILM applications.

As future work, further appliances with PROCEL PBE label will be analysed, such as air conditioners, electric ovens, washing machines, dryers, etc.

Abbreviations
AFE: Analog front-end; ASD: Adjustable speed driver; BBB: BeagleBone black; CPT: Conservative power theory; EE: Energy efficiency; EEC: Energy efficiency coefficient; FN: False negative; FP: False positive; FPR: False positive rate; FSM: Finite state machines; HDF5: Hierarchical data format; KNN: K-nearest neighbors; NILM: Non-intrusive load monitoring; NILMEE: Non-intrusive load monitoring for energy efficiency; NILMTK: Non-intrusive load monitoring toolkit; PBE: Brazilian labelling program; PRU: Programmable real-time unit; PSB: Power signature blob; RMS: Root mean square; RMSE: Root mean square error; TN: True negative; TP: True positive; TPR: True positive rate

Acknowledgements
This research was supported by FAPESP (Sao Paulo Research Foundation) under grants 2012/19375-1 and 2016/08645-9, by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Authors’ contributions
FDG carried out the energy efficiency studies and practical tests, improved the load disaggregation concept, and also structured the first paper draft. WAS participated in the development of the load disaggregation technique, by using the CPT and the state-machine, and he wrote the first version of the manuscript. ISD participated in the development of the energy meter prototype and experimental activities. FPM provided supervision, funding, and other resources for this work and he also contributed to writing the manuscript. All authors contributed to the definition of the proposed NILMEE concept, read and approved the submitted manuscript.

Availability of data and materials
All datasets used for the paper development and methodology evaluation will be available in a public repository (https://gitlab.com/fernando.delunogarcia/psb-knn-dataset).

Competing interests
The authors declare that they have no competing interests.

Author details
1 Sao Paulo State University (Unesp), Institute of Science and Technology, Avenida Três de Maço, 511, Alto da Boa Vista, 18087-180 Sorocaba, Brazil. 2 Department of Computer Science, Federal University of Sao Carlos, Rodovia Joao Leme dos Santos, SP-264, Km 110 - Itinga, 18052-780 Sorocaba, Brazil.

Received: 28 May 2020 Accepted: 29 September 2020

Published online: 19 October 2020

References
Aboulian A, Green DH, Switzer JF, Kane TJ, Bredariol GV, Lindahl P, Donnal JS, Leeb SB (2019) NILM dashboard: A power system monitor for electromechanical equipment diagnostics. IEEE Trans Ind Inform 15(3):1405–1414. https://doi.org/10.1109/TII.2018.2843770
Abubakar I, Khalid SN, Mustafa MW, Shareef H, Mustapha M (2015) An overview of non-intrusive load monitoring methodologies. In: Proceedings of the IEEE Conference on Energy Conversion. pp 54–59. https://doi.org/10.1109/CFCEN.2015.7409813
Aladesanmi EJ, Fally KA (2015) Overview of non-intrusive load monitoring and identification techniques. Proc Symp Control Power Energy Syst 48(30):415–420. https://doi.org/10.1016/j.jspecol.2015.12.414
Azaza M, Wallin F (2017) Finite state machine household’s appliances models for non-intrusive energy estimation. In: Proceedings of the Energy Fedia – International Conference on Applied Energy Vol. 105. pp 2157–2162. https://doi.org/10.1016/j.egypro.2017.03.609
Baets LD, Ruysinck J, Develider C, Dhaene T, Deschrijver D (2017) On the Bayesian optimization and robustness of event detection methods in NILM. Energy Build 145:57–66. https://doi.org/10.1016/j.enbuild.2017.03.061
Bao K, Ibrahimov K, Wagner M, Schmeck H (2018) Enhancing neural non-intrusive load monitoring with generative adversarial networks. Energy Inform 1(1):295–302. https://doi.org/10.1186/s42162-018-0038-y
Batra N, Kelly J, Parson O, Dutta H, Knotenbelt W, Rogers A, Singh A, Srivastava M (2014) NILMTK: an open source toolkit for non-intrusive load monitoring. In: Proceedings of the 5th International Conference on Future Energy Systems. pp 265–276. https://doi.org/10.1145/2602044.2602051
Berbakov L, Batké M, Tomalačević N (2019) Smart energy manager for energy efficient buildings. In: IEEE EUROCON 2019 -18th International Conference on Smart Technologies. pp 1–4. https://doi.org/10.1109/EUROCON.2019.8861616
MCT/MME (2011) Approvance of the goals program for refrigerators and freezers (in Portuguese) - Interministerial ordinance no. 326 of May 26, 2011. Technical report. Ministry of Science, Technology, Innovations and Communications, Brasilia, pp 1–13. https://www.jusbrasil.com.br/diarios/DOU/2011/05/27

Merini I, Molina-García A, Garcia-Cascales MS, Achad M (2019) Energy Efficiency Regulation and Requirements: Comparison Between Morocco and Spain. In: Ezzyni M (ed) Proceedings of the Advanced Intelligent Systems for Sustainable Development. Springer, Cham, pp 197–209

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Couraneau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: Machine learning in Python. J Mach Learn Res 12:2825–2830

Pino-Mejias R, Pérez-Fargallo A, Rubio-Bellido C, Pulido-Arcas JA (2017) Comparison of linear regression and artificial neural networks models to predict heating and cooling energy demand, energy consumption and CO2 emissions. Energy 118:24–36. https://doi.org/10.1016/j.energy.2016.12.022

Rashid H, Singh P, Stankovic V, Stankovic L (2019) Can non-intrusive load monitoring be used for identifying an appliance’s anomalous behaviour? Appl Energy 238:796–805. https://doi.org/10.1016/j.apenergy.2019.01.061

Rottondi C, Derboni M, Piga D, Rizzoli AE (2019) An optimisation-based energy disaggregation algorithm for low frequency smart meter data. Energy Inform 2(1):13. https://doi.org/10.1186/s42162-019-0089-8

Ruano A, Hernandez A, Urena J, Ruano M, Garcia J (2019) NILM techniques for intelligent home energy management and ambient assisted living: A review. Energies 12(11):2203. https://doi.org/10.3390/en12112203

Sadeghianpourhamami N, Ryussinck J, Deichriyjer D, Dhaene T, Develder C (2017) Comprehensive feature selection for appliance classification in NILM. Energy Build 151:98–106. https://doi.org/10.1016/j.enbuild.2017.06.042

Souza WA, Garcia FD, Marafao FP, da Silva LCP, Simoes MG (2019) Load disaggregation using microscopic power features and pattern recognition. Energies 12(14). https://doi.org/10.3390/en12142641

Souza WA, Garcia FD, Moreira AC, Marafao FP, Silva LCP (2020) Automatic consumption management for prepaid electricity meter with NILM. IEEE Lat Am Trans 18(06):1102–1110. https://doi.org/10.1109/TLA.2020.9099748

Souza WA, Marafio FP, Liberado EV, Simoes MG, Da Silva LCP (2018) A nilm dataset for cognitive meters based on conservative power theory and pattern recognition techniques. J Control Autom Electr Syst 29(6):742–755. https://doi.org/10.1007/s40313-018-0417-4

Stankovic L, Stankovic V, Murray D, Liao J (2016) Energy feedback enabled by load disaggregation. In: 1st Energy feedback symposium. pp 78–84. https://strathprints.strath.ac.uk/61936/1/Stankovic_etal_EFS_2016_Energy_feedback_enabled_by_load_disaggregation.pdf

Tenti P, Paredes HKM, Mattavelli P (2011) Conservative power theory, a framework to approach control and accountability issues in smart microgrids. IEEE Trans Power Electron 26(3):664–673. https://doi.org/10.1109/TPEL.2010.2093153

Teshome DF, Huang TD, Lian K (2016) Distinctive load feature extraction based on fryze’s time-domain power theory. IEEE Power Energy Technol Syst J 3(2):60–70. https://doi.org/10.1109/JPETS.2016.2559507

Trotta G (2018) Factors affecting energy-saving behaviours and energy efficiency investments in British households. Energy Policy 114:529–539. https://doi.org/10.1016/j.enpol.2017.12.004

Trotta G (2018) Factors affecting energy-saving behaviours and energy efficiency investments in British households. Energy Policy 114:529–539. https://doi.org/10.1016/j.enpol.2017.12.004

WEC (2013) World Energy Perspective: Energy Efficiency Policies: what works and what does not. Technical report. World Energy Council, London, pp 1–15. https://www.worldenergy.org/assets/downloads/World_Energy_Perpective_Energy-Efficiency-Policies-2013_Full_Report.pdf

WEC (2016) Energy efficiency: A straight path towards energy sustainability. Technical report. World Energy Council, London, pp 1–15. https://www.worldenergy.org/assets/downloads/EnergyEfficiencyAStraightPathFullReport.pdf

Wong IL, Kruger E (2017) Comparing energy efficiency labelling systems in the EU and Brazil: Implications, challenges, barriers and opportunities. Energy Policy 109:310–323. https://doi.org/10.1016/j.enpol.2017.07.005

Zhuang M, Shahidehpour M, Li Z (2018) An overview of non-intrusive load monitoring. Approaches, business applications, and challenges. In: Proceedings of the IEEE International Conference on Power System Technology. pp 4291–4299. https://doi.org/10.1109/POWERCON.2018.8601534