

Article

Energy Disaggregation Using Two-Stage Fusion of Binary Device Detectors

Pascal A. Schirmer ^{1,*} , Iosif Mporas ¹ and Akbar Sheikh-Akbari ² 

¹ Communications and Intelligent Systems Group, School of Engineering and Computer Science, University of Hertfordshire, Hatfield AL10 9AB, UK; i.mporas@herts.ac.uk

² School of Built Environment, Engineering and Computing, Leeds Beckett University, Leeds LS1 3HE, UK; a.sheikh-akbari@leedsbeckett.ac.uk

* Correspondence: p.schirmer@herts.ac.uk

Received: 26 March 2020; Accepted: 21 April 2020; Published: 1 May 2020



Abstract: A data-driven methodology to improve the energy disaggregation accuracy during Non-Intrusive Load Monitoring is proposed. In detail, the method uses a two-stage classification scheme, with the first stage consisting of classification models processing the aggregated signal in parallel and each of them producing a binary device detection score, and the second stage consisting of fusion regression models for estimating the power consumption for each of the electrical appliances. The accuracy of the proposed approach was tested on three datasets—ECO (Electricity Consumption & Occupancy), REDD (Reference Energy Disaggregation Data Set), and iAWE (Indian Dataset for Ambient Water and Energy)—which are available online, using four different classifiers. The presented approach improves the estimation accuracy by up to 4.1% with respect to a basic energy disaggregation architecture, while the improvement on device level was up to 10.1%. Analysis on device level showed significant improvement of power consumption estimation accuracy especially for continuous and nonlinear appliances across all evaluated datasets.

Keywords: energy disaggregation; non-intrusive load monitoring; regression fusion

1. Introduction

Between 25% and 40% of the global energy consumption and the corresponding amount of carbon dioxide emissions comes from residential buildings [1–4]. It is estimated that in the next two decades the average number of electrical devices used in houses is going to rise [4]. In parallel, climate change and urbanization are affecting the energy load of urban buildings, with the energy load demand growing two times faster than the expansion of urbanization [5] have shown that roughly 20% of households consumed energy is due to faulty equipment or poor operational strategies [6–8]. Therefore, to detect faulty device operation and improve operation strategies, optimization techniques in terms of device detection and load scheduling have been developed to find optimal and suboptimal operational strategies [9]. Additionally, significant progress in smart grids, smart systems, and smart devices was made in the last few decades, considering optimized energy generation and distribution [9,10]. Accordingly, energy management and the deployment of Information and Communication Technologies (ICT) in residential buildings increased as well, in order to reduce households' energy consumption without decreasing living quality levels or violating consumers personality rights and privacy [11,12]. In general, the amount of information gathered is increased progressively with respect to consumer behavior. Especially, usage of energy is monitored to reduce overall energy consumption and peak loads, while improvement of the well-being of consumers is tried to be achieved as well [13].

Studies have shown that for achieving significant decrease in energy consumption smart energy management, smart grids, fine-grained energy monitoring, as well as load forecasting on household

level are indispensable [14,15]. However, nowadays energy monitoring is mostly done via an aggregated measure of energy on monthly bills and does not offer detailed information about energy monitoring. Therefore, to accurately measure energy consumption, smart meters are utilized usually measuring with sampling frequency equal to 1 Hz or more. Smart meters are devices used to measure energy consumption of electrical appliances, based on voltage and current measurements. The energy consumption is calculated at periods of time which usually are every 1 second or more frequently, e.g., up to 30,000 samples per second [16]. The more frequently energy consumption is calculated the more detailed is the captured information of energy consumption; however, increasing the sampling frequency will linearly increase the data to be stored, processed, or transmitted which in turn increases the hardware cost exponentially [17]. Therefore, most recent studies focus on low sampling frequency data, as the majority of commercial smart meters collect data usually at 0.1 Hz or up to 1 Hz to minimize the hardware cost of smart meters and to address the transmission and data-storage capacity limitations [18,19]. Energy saving enhancement can be achieved on device level by detecting faulty device operation and inefficient operating strategies [7]. Knowledge about the appliances' consumption can lead to a reduction of total consumption through increased awareness of energy consumption [20]. Recent studies have shown that households are usually bad at estimating individual power consumption (e.g., overrating small appliances consumption and under-rating the amount of energy for heating) [21]. This means that the energy consumption must be either measured on device level, which disadvantageously results in increased cost due to wiring issues and data acquisition [19], or that the aggregated energy (consumed energy measured centrally for each household) must be split to appliance level automatically, which is called energy disaggregation. Energy disaggregation as defined in [22] is the Non-Intrusive Load Monitoring (NILM) determining the consumption of energy from each individual appliance of a house, performed by processing of measurements of the current and voltage of the overall household's load. The term non-intrusive is used to point out the distinction to Intrusive Load Monitoring (ILM) methods utilizing several measurements and smart meters and set the focus on determining the per device consumption. In other words, NILM is extracting electrical energy consumption at appliance-level based on one central measurement, thus to identify the onsets t_{on} (switch-on times) and t_{off} (switch-off times) of appliances from the aggregated energy signal in order to find the corresponding consumptions per appliance [23].

Several methods to solve the NILM energy disaggregation challenge can be found in the bibliography. These methods are briefly classified in methods using Source Separation (SS) algorithms and in approaches that do not use SS algorithms. Common for all NILM approaches is that they use measurements of the aggregated energy consumption of a household with a sampling frequency f_s in the order of a sample per second up to few tens of kHz [16]. NILM methods may use macroscopic signal parameters (e.g., active/reactive power [24,25]) or microscopic ones (e.g., transient energy and harmonics [26–28]), depending on the sampling rate f_s , to split the aggregated signal in appliance level [29]. Appliance identification methods not using SS algorithms are based mainly on supervised methods and the extraction of features, which will be used either for training a Machine Learning (ML) algorithm (e.g., Support Vector Machines (SVM) [30], Artificial Neural Network (ANN) [31], Decision Tree (DT) [32], K-Nearest Neighbours (KNN) [33]), or defining a set of rules or thresholds [28]. As regards appliance identification methods using SS algorithms, they are based on single-channel source separation and solve the task with optimization criteria. Approaches using source separation extract the power consumption characteristic pattern of every appliance from the aggregated signal using an optimization algorithm with constraints [19,34,35]. Commonly reported SS algorithms in the NILM task are Independent Component Analysis (ICA) [36], Non-Negative Matrix Factorization (NMF) [37], and Sparse Component Analysis (SCA) [38]. Source Separation-based NILM approaches are unsupervised; however, a priori knowledge is needed as only the aggregated signal measurements are used, thus making them semi-supervised [19], in contrast to the NILM approaches without using SS algorithms, which are supervised. Furthermore, cutting edge technology in machine learning has led to a number of recently proposed in the literature deep learning approaches using big datasets,

like the Almanac of Minutely Power Dataset (AMPds) [39]. Methodologies using Convolutional Neural Networks (CNNs) [40–42], Recurrent Neural Networks (RNNs) [43,44] and Long Short-Term Memory (LSTM) architectures [44,45], denoising autoencoders (dAEs) [46], and Gated Recurrent Units (GRUs) [40] can be found in the bibliography. Furthermore, additional questions regarding consumer privacy and real-time capability arise with the high frequent measurements of energy consumption, and have been discussed in [47,48] for security relevant issues and in [17] and [49] for low cost disaggregation and real-time capability.

There is still no established approach for solving the NILM problem and literature reports multiple solutions with and without source separation. There are numerous electrical devices which have steady state behavior [22] and are typically modeled as finite state machines [22,50] as well as electrical devices with non-steady behavior, which have nonlinear and/or continuous characteristics [51,52]. The identification of such appliances when working in parallel or showing strong time-dependent behaviors [53] is still an unsolved problem, especially for nonlinear and continuous devices. In this paper a two-stage fusion approach is proposed aiming at representing different device combinations and their time varying behavior more accurately. The proposed methodology is based on supervised learning and utilizes low frequency data as well as steady-state features, similar as in [54–56].

The remaining of this article is organized as follows. Section 2 presents the proposed two-stage fusion methodology. In Sections 3 and 4, the experimental set-up and the experimental results are given, respectively. Finally, in Section 5 conclusions are provided.

2. Two-Stage Fusion Methodology

The NILM energy disaggregation task can be described as the problem of estimation of the power consumption of each electrical appliance using the measurements acquired from one central smart meter, within time windows (frames or epochs). In detail, given a set of $M - 1$ known appliances each consuming power p_m , with $1 \leq m \leq M$, the aggregated power P_{agg} measured by the central smart meter will be

$$P_{agg} = f(p_1, p_2, \dots, p_{M-1}, p_g) = \sum_{m=1}^{M-1} p_m + p_g = \sum_{m=1}^M p_m \quad (1)$$

where $p_g = p_M$ is a “ghost” power consumption, which is usually consumed by one or more unknown appliances. In NILM, the aim is to calculate estimations $\hat{P} = \{\hat{p}_m, 1 \leq m \leq M\}$ of the power consumption of each electrical appliance m using an estimation method f^{-1} with minimal estimation error and $\hat{p}_M = \hat{p}_g$, i.e.,

$$\begin{aligned} \hat{P} &= \{\hat{p}_1, \hat{p}_2, \dots, \hat{p}_{M-1}, \hat{p}_g\} = f^{-1}(P_{agg}) \\ \text{s.t. } \underset{f^{-1}}{\operatorname{argmin}} \left\{ (P_{agg} - \hat{P})^2 \right\} &= \underset{f^{-1}}{\operatorname{argmin}} \left\{ \left(P_{agg} - \sum_{m=1}^M \hat{p}_m \right)^2 \right\} \end{aligned} \quad (2)$$

As Equation (2) is practically impossible to be solved using an analytical solution, most energy disaggregation methodologies are based on segmentation of the aggregated signal into frames and estimation of the power consumption on device level within each frame using a machine learning based model, which can either be one model per device following the “one vs. all” approach [57] or a multi-class device identification model [58]. The architecture of the baseline one-stage NILM approach based on regression estimators of power consumption is presented in Figure 1.

Specifically, the one-stage NILM methodology consists of preprocessing, feature extraction, and a regression model for estimating the appliances power consumption \hat{P} . During preprocessing the aggregated signal is initially filtered, in order to remove peaks as proposed in [59], frame blocked in time frames h_t of length L , and a feature vector $v_t, v_t \in \mathbb{R}^K$, is calculated for each frame h_t , where $1 \leq t \leq T$ and T is the last frame of the aggregated signal. Finally, a regression model is used to estimate power consumption values $\hat{P} = \{\hat{p}_1, \hat{p}_2, \dots, \hat{p}_{M-1}, \hat{p}_g\}$ for each of the M devices. The estimation of each device’s power consumption can be done either using in parallel one regression model per device or using one regression model with M output-estimations.

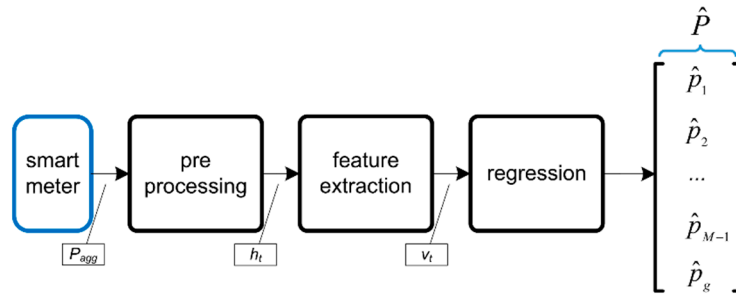


Figure 1. Block diagram of the baseline NILM architecture consisting of preprocessing, feature extraction, and regression estimation of power consumption.

In this work, the one-stage NILM methodology is extended to two stages. In detail, the first stage consisting of classifiers (device detectors) processing the aggregated signal in parallel and each of them producing a binary device-specific detection score, while the second stage consists of regression fusion models for estimating the power consumption of each appliance using as input the stage I results concatenated with the feature vector. The architecture of the proposed two-stage methodology is presented in Figure 2.

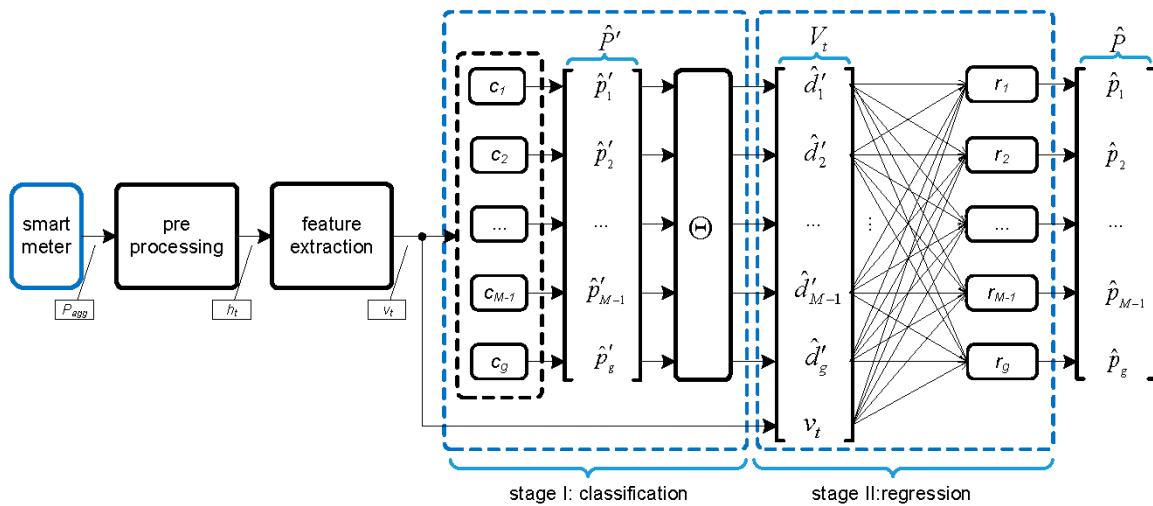


Figure 2. Block diagram of the proposed two-stage energy disaggregation methodology.

In detail, during stage I the feature vectors are initially processed by a set of M classification models $C = \{c_1, c_2, \dots, c_{M-1}, c_g\}$, one for each of the $M - 1$ known devices and one for the unknown ghost-power according to the “one vs. all” approach. The output before the last layer of stage I, $\hat{P}' = \{\hat{p}'_1, \hat{p}'_2, \dots, \hat{p}'_{M-1}, \hat{p}'_g\}$ is the classification score for each of the M devices:

$$\hat{p}'_m = c_m(v_t) \tag{3}$$

where c_m is the classification model for the m th device and v_t is the feature vector as calculated in the feature extraction stage. The predicted class is the one with the highest score \hat{p}'_m . To get the binary decision at the end of stage I, a threshold Θ is applied to transform the initial classification scores \hat{p}'_m to their binary representation, thus labeling if a device is working (1) or not (0):

$$\hat{d}'_m = \begin{cases} 0 & \text{if } \hat{p}'_m < \Theta \\ 1 & \text{if } \hat{p}'_m \geq \Theta \end{cases} \tag{4}$$

Subsequently, the initial binary estimations, $\hat{D}' = \{\hat{d}'_1, \hat{d}'_2, \dots, \hat{d}'_{M-1}, \hat{d}'_g\}$ with $\hat{D}' \in \mathbb{R}^M$, from stage I are concatenated together with the feature vector, v_t to a new feature vector $V_t = \{\hat{D}'|v_t\} \in \mathbb{R}^{(K+M)}$, so as to estimate the power consumptions of the M appliances. Specifically, in the second stage M fusion models, $R = \{r_1, r_2, \dots, r_{M-1}, r_g\}$ with $R \in \mathbb{R}^M$, are receiving as input the new feature vector V_t , giving a numerical estimation (regression) for the appliance power consumption for each of the M devices.

$$\hat{p}_m = r_m(V_t) \quad \text{s.t. } \hat{p}_m \in \{0, \dots, \max(h_t)\} \quad (5)$$

The initial binary estimates of device operation \hat{D}' from the first stage are used from the regression models of the second stage to model any power consumption correlations between the different appliances, i.e., the devices that are likely to work simultaneously within the time frame v_t . Additionally, the restriction on Equation (5) assures that the prediction of power consumption for each single device \hat{p}_m at frame instance t cannot exceed the aggregated power consumption within that frame.

The proposed methodology combines binary device estimates from a first classification stage with a second regression fusion stage, thus any complementary information from the first stage will be captured and learned by the fusion model. Moreover, with the existence of ghost power in the first level, the output of the binary classifiers will be used as a feature for the detection of unknown devices, which offers advantage to the present methodology in real set-up evaluations where unknown devices exist quite often.

3. Experimental Set-up

A detailed description of the databases used to evaluate the one-stage and the proposed two-stage fusion methodology as well as the description of the parameterization of the machine learning algorithms are provided in this section.

3.1. Evaluation Data

To evaluate the proposed methodology presented in Section 2 the data collections Electricity Consumption & Occupancy (ECO) [59], Reference Energy Disaggregation Data Set (REDD) [60], and Indian Dataset for Ambient Water and Energy (iAWE) [61], which are freely and online accessible, were used as they contain low frequency samples from the aggregated data and individual power measurements from each device, respectively. The three databases consist of several datasets with different monitored houses in each. For the present evaluation from the ECO database houses, 1, 2, and 4–6 were used, while the ECO-3 dataset was not used because it does not include the power consumption signals of each appliance but only the aggregated signal. Further, from the REDD database, house 5 was excluded as its measurement duration is significantly shorter than for the rest of the datasets in the REDD database [62]. The datasets used in the present evaluation are shown in Table 1 with column “#App” tabulating the total number of appliances (App) in each dataset and in brackets the number of devices with power consumption above 25 W, with the remaining ones considered as “ghost device”, in alignment with the experimental protocol introduced in [57,58]. The remaining columns of Table 1 are listing the sampling period T_s , the duration T , and the device types included in every dataset. As regards the REDD database, all of it was utilized, ignoring the gaps in the measurements as in [63]. Regarding the ECO and iAWE databases, one week of energy consumption recordings was used in order to the size of training data to be similar with the REDD dataset. Specifically, we used the week from 05/07/2012 until 11/07/2012 for the ECO database and the week from 08/06/2013 until 14/06/2013 for the iAWE database.

Table 1. Overview of the evaluated datasets.

Dataset	#App	Ts	T	App. Type	Appliances
ECO-1	7(6)	1s	7d	One/Multi-State	(1) fridge, (2) dryer, (3) coffee machine, (4) kettle, (5) washing machine, (6) PC, (7) freezer
ECO-2	12(9)	1s	7d	One/Multi-State	(1) tablet, (2) dishwasher, (3) air exhaust, (4) fridge, (5) entertainment, (6) freezer, (7) kettle, (8) lamp, (9) laptop, (10) Stove, (11) TV, (12) Stereo
ECO-4	8(8)	1s	7d	One/Multi-State/ Nonlinear	(1) fridge, (2) kitchen appliances, (3) lamp, (4) stereo/laptop, (5) freezer, (6) tablet, (7) entertainment, (8) microwave
ECO-5	8(6)	1s	7d	One/Multi-State/ Nonlinear	(1) tablet, (2) coffee machine, (3) kettle, (4) microwave, (5) fridge, (6) entertainment, (7) PC, router/printer, (8) fountain
ECO-6	7(6)	1s	7d	One/Multi-State/ Nonlinear	(1) lamp, (2) laptop/printer, (3) routers, (4) coffee machine, (5) entertainment, (6) fridge, (7) kettle
REDD-1	18(17)	3s	All	One/Multi-State/ Continuous	(1) oven, (2) oven, (3) refrigerator, (4) dishwasher, (5) kitchen-outlets, (6) kitchen-outlets, (7) lighting, (8) washer-dryer, (9) microwave, (10) bathroom, (11) electric- heat, (12) stove, (13) kitchen-outlets, (14) kitchen-outlets, (15) lighting, (16) lighting, (17) Washer-dryer, (18) Washer-dryer
REDD-2	9(10)	3s	All	One/Multi-State	(1) kitchen-outlets, (2) lighting, (3) stove, (4) microwave, (5) washer-dryer, (6) kitchen-outlets, (7) refrigerator, (8) dishwasher, (9) disposal
REDD-3	20(18)	3s	All	One/Multi-State/ Nonlinear	(1) outlets-unknown, (2) outlets-unknown, (3) lighting, (4) electronics, (5) refrigerator, (6) disposal, (7) dishwasher, (8) furnace, (9) lighting, (10) outlets-unknown, (11) washer-dryer, (12) washer-dryer, (13) lighting, (14) microwave, (15) lighting, (16) smoke-alarms, (17) lighting, (18) bathroom, (19) kitchen-outlets, (20) kitchen-outlets
REDD-4	18(16)	3s	All	One/Multi-State/ Nonlinear	(1) lighting, (2) furnace, (3) kitchen-outlets, (4) outlets-unknown, (5) washer-dryer, (6) stove, (7) air-conditioning, (8) air-conditioning, (9) miscellaneous, (10) smoke-alarms, (11) lighting, (12) kitchen-outlets, (13) dishwasher, (14) bathroom, (15) bathroom, (16) lighting, (17) lighting, (18) air-conditioning
REDD-6	15(14)	3s	All	One/Multi-State/ Nonlinear	(1) kitchen-outlets, (2) washer-dryer, (3) stove, (4) electronics, (5) bathroom, (6) refrigerator, (7) dishwasher, (8) outlets-unknown, (9) outlets-unknown, (10) electric- heat, (11) kitchen-outlets, (12) lighting, (13) air-conditioning, (14) air-conditioning, (15) air-conditioning
iAWE	10(9)	1s	7d	One/Multi-State/ Nonlinear/ Continuous	(1) fridge, (2) air-condition, (3) air-conditioning, (4) washing machine, (5) laptop, (6) iron, (7) kitchen, (8) TV, (9) waterfilter, (10) watermotor

These weeks were chosen with the intention of having as many appliances as possible in the selected time interval of the aggregated signal. Except this, in [59,64], where the ECO and the iAWE databases were also used the selected time interval has not been specified. The classification of device types is based on their operation as described in [65,66], i.e., one-state electrical appliances have only on/off status (for example resistive lamps, kettles or fridges without significant power spikes),

multi-state devices have a number of discrete power consumption states (e.g., washing machines with numerous washing cycles), nonlinear devices (e.g., electronics) and electrical appliances with continuous power consumption pattern, which are controlled by power electronics (e.g., air condition) and usually have an exponential decay signature. The device signatures may present an amplitude peak in the beginning of the signature, as, for example, in the case of refrigerators. An example power signature for each of the four device categories was extracted from the REDD databases and is illustrated in Figure 3.

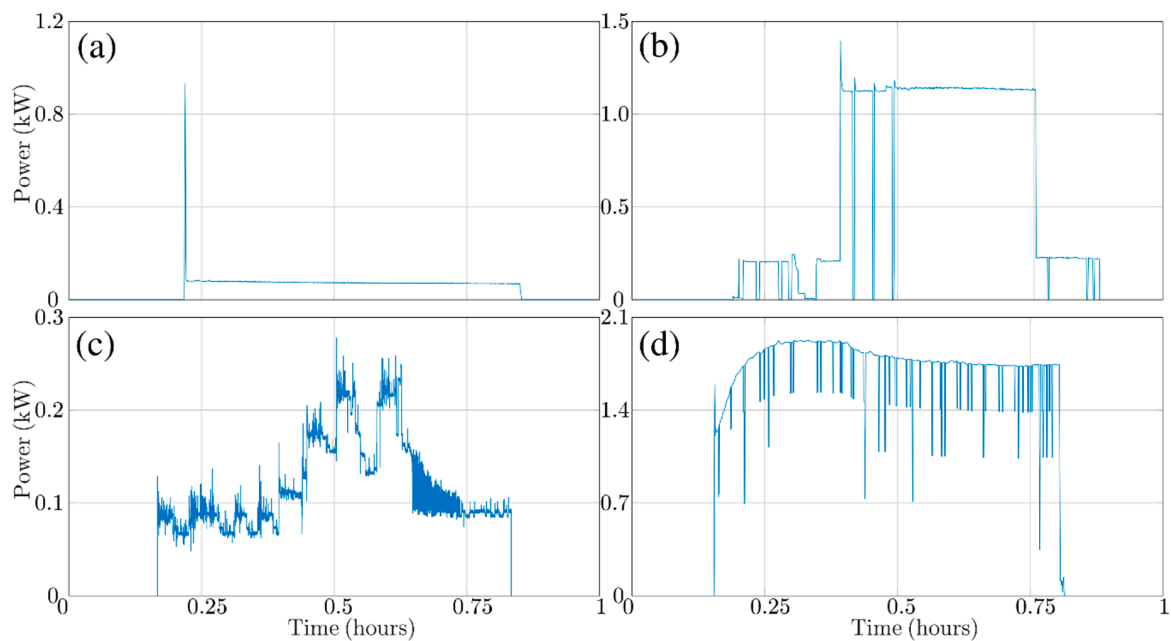


Figure 3. Examples of appliance signatures for (a) one-state appliance with significant peak (refrigerator), (b) multi-state appliance without significant peak (dishwasher), (c) nonlinear appliance (laptop), and (d) continuous appliance with decay (air-conditioning) from the REDD database.

As can be seen from Table 1, the evaluated datasets vary in terms of number of appliances, monitoring durations, as well as appliance type, and therefore are accurately representing the various characteristics of nowadays households [59,60]. All evaluated datasets have a low sampling rate in the order of seconds and only the active power samples of the aggregated signal are utilized, offering a good trade-off between computational load and real-time operation [64].

3.2. Parameterization and Feature Selection

At the preprocessing of the aggregated signal, a median filter of five samples was used for smoothing as proposed in [59], and afterwards the preprocessed signal was segmented into overlapping frames of length equal to $L = 10$ samples and time shift between successive frames equal to 5 samples. The optimal number of samples per frame was determined through grid search on a bootstrap dataset with ideal aggregated data (without ghost power), consisting of one dataset out of each database (ECO-2, REDD-2 and iAWE) similar as in [67,68].

All devices with constant power consumptions of less than 25 W were removed from the datasets and added to the ghost-power, while the aggregated data was not modified, which ensures that the training as well as the testing was done with real measurements of the aggregated data and not with an artificial dataset created through summing consumptions of all appliances [69]. The set of binary classifiers C was trained, one for every device m and separately for each dataset according to the “one vs. all” approach, whereas the threshold was set equally to $\Theta = 25$ W for all appliances. During the training phase the set of features, v_t , was determined from a time window of active power samples

h_t and the Min/Max, Mean, Energy, RMS, Percentiles25/75, Median, Zero Crossing rate, Peak2Rms, Range, Standard Deviation, Skewness, Kurtosis, and Variance values were extracted according to their statistical importance determined by the ReliefF algorithm [70] resulting to a $K = 15$ dimensional feature vector similar as in [71,72]. Specifically, Mean, Energy, RMS were used to model steady-state behavior, while Min/Max, Percentile75/25, Median, Zero Crossing rate, Peak2Rms, Range, Standard Deviation, Skewness, Kurtosis, and Variance was used to model for the transient behavior and the variation within the frames [73]. As all databases are sampled with relatively low sampling frequencies the feature vector only contains steady-state features.

Similarly, the regression fusion models were trained using the intermediate binary scores from the first stage, \hat{D}' , as well as the original feature vector v_t . In detail \hat{D}' and v_t were concatenated into a single feature vector and used to train the set of fusion regression models R , one for each of the M devices. Both the one-stage architecture (Figure 1) and proposed two-stage fusion architecture (Figure 2) were trained with the first half of each dataset and tested on the second half of each dataset, thus without overlap between training and test subsets.

For building the models of the one-stage and two-stage architecture Deep Neural Networks (DNNs), K-Nearest-Neighbors (KNNs), Decision Trees (DTs) in a Random Forest (RF) implementation, and Support Vector Machines (SVM) were used. Short description and free parameters of the evaluated classifiers are tabulated in Table 2. The values of the adjustable parameters of the evaluated regression algorithms were fine-tuned empirically by performing grid search on a bootstrap subset of the training data composed of the ECO-1/2/4/5/6 database which didn't include any ghost power. The performance was evaluated in terms of appliance power estimation accuracy (E_{ACC}), as proposed in [60] and defined in Equation (6).

$$E_{ACC} = 1 - \frac{\sum_{t=1}^T \sum_{m=1}^M |\hat{p}_m^t - p_m^t|}{2 \sum_{t=1}^T \sum_{m=1}^M |p_m^t|} \quad (6)$$

where \hat{p}_m is the estimated power, p_m the ground-truth power consumption of the m th device, T denotes the total number of frames, and M is the number of electrical appliances including the ghost power. The free parameters optimization of the regression models with respect to the power estimation accuracy E_{ACC} at the end of the one-stage architecture, \hat{p}_m , are shown in Table 2.

Table 2. Parameterization results E_{ACC} (%) for four different classifiers, namely, Deep Neural Networks (DNNs), Random Forest (RFs), K-Nearest-Neighbors (KNNs), and Support Vector Machines (SVMs).

Deep Neural Network (DNN)						
Nodes/Layers	4	8	16	32	64	128
1	80.4	87.5	87.9	83.7	86.4	81.7
2	70.1	86.4	86.9	87.5	82.7	83.6
3	80.4	86.7	87.9	88.7	88.4	84.2
4	75.4	87.9	87	87.2	85.3	83.7
Random Forest (RF)						
Trees	8	16	32	64	128	256
	85.5	85.3	85.5	85.4	85.4	85.4
K-Nearest-Neighbours (KNN)						
K	1	2	3	4	5	6
	82.2	82.7	82.7	83.1	83.3	82.4
Support Vector Machine (SVM)						
Kernel	Linear	Gaussian	Rbf	Pol-2	Pol-3	Pol-4
	55.0	72.3	76.3	59.2	63.6	67.8

As shown from Table 2, the optimized parameters (in bold) of the regression models are a DNN model with 3 hidden layers and 32 sigmoid nodes per layer, a KNN with $K = 5$ nearest neighbors, a RF with 32 trees per forest and a SVM with Radial Basis Function (RBF) as kernel with parameters gamma equal to 12.8 and C equal to 1.45. The DNN model achieved accuracy equal to 88.7% and outperformed all other evaluated regression models on the bootstrap subset of the training data.

4. Experimental Results

The NILM methodology described in Section 2 was tested based on the experimental protocol presented in Section 3 using the parameter optimization results of Table 2. To evaluate NILM accuracy on electrical appliance level, Equation (6) was modified by removing the sum across the M appliances, thus resulting to

$$E_{ACC}^i = 1 - \frac{\sum_{t=1}^T |\hat{p}_m^t - p_m^t|}{2 \sum_{t=1}^T |p_m^t|} \quad (7)$$

The experimental results in terms of E_{ACC} (%) for all evaluated datasets, all evaluated classification algorithms and for both the one-stage and proposed two-stage architecture are tabulated in Table 3. The best performing energy disaggregation scores per dataset are indicated in bold for both one- and two-stage results.

Table 3. Performance of energy disaggregation in terms of E_{ACC} (%) for different datasets using the one-stage (I) and the proposed two-stage fusion methodology (II).

Dataset	DNN		RF		KNN		SVM	
	I	II	I	II	I	II	I	II
ECO-1	74.5	76.2	78.4	79.4	76.0	77.7	67.0	67.0
ECO-2	85.5	87.5	86.3	89.3	85.4	86.4	78.5	80.5
ECO-4	83.8	84.6	83.8	86.9	82.1	82.2	81.5	81.5
ECO-5	88.3	90.3	89.2	90.2	88.1	89.1	88.4	89.4
ECO-6	78.4	80.1	84.6	86.1	83.7	84.2	71.9	74.6
REDD-1	71.3	73.1	78.0	79.0	74.9	75.3	66.3	66.3
REDD-2	74.9	79.0	85.3	87.3	84.4	84.4	81.1	81.1
REDD-3	67.6	69.6	70.6	71.7	69.2	69.9	66.3	66.3
REDD-4	73.9	75.3	74.5	75.1	72.6	73.5	72.5	73.3
REDD-6	79.9	81.3	81.6	82.7	79.3	79.5	70.8	70.8
iAWE	64.7	66.0	67.2	69.2	66.9	67.9	77.4	80.8

As shown in Table 3, the best performing classifier amongst all tested datasets, when using the one-stage architecture, is RF outperforming all other classifiers except for the case of *iAWE* dataset where the SVM classifier achieves significant higher performance in terms of energy disaggregation. Furthermore, the results in Table 3 show that the two-stage fusion methodology improves the overall E_{ACC} performance across all evaluated datasets. In terms of average improvement per dataset E_{ACC} increases between 0.6% and 4.1% depending on the dataset and the classifier. The most significant improvements in terms of relative performance were observed when using DNN as classifier where performance was improved by 4.1% (REDD-2 dataset). The improvement in terms of absolute E_{ACC} values, i.e., the average increase in estimation accuracy when considering the best experiment for the first stage as the baseline performance, ranges between 0.6% and 3.4% when using SVM and RF as classifiers and the results were statistically significant when comparing their accuracy scores on frame level of the one-stage and the two-stage fusion architectures. In detail, RF outperformed SVM in ten out of eleven datasets with exception of the *iAWE* database, which is probably due to the significant

higher proportion of continuous appliances which is in line with results in literature reporting high accuracies for SVM in case of appliances with strong time varying behavior [73,74]. The evaluation results demonstrate the validity of the proposed method as it has offered improved performance when tested in several and highly dissimilar (with respect to the sampling rate f_s , the number and the type of devices) datasets as presented in Section 3 and shown in Table 1.

In a next step we performed analysis of energy disaggregation performance on device level for one dataset out of each database. Table 4 tabulates the E_{ACC} on device level for the ECO-2, REDD-2, and iAWE datasets. The choice for the three datasets was made according to the characteristics of the datasets shown in Table 1. Specifically, datasets which have roughly the same number of appliances (<10) and are similar in their collection of appliances thus having appliances of the same type were chosen.

Table 4. Per device performance E_{ACC}^m (%) of the one-stage (I) and the proposed two-stage fusion (II) architecture using the best performance classifier (RF) conducted from the per dataset results. The superior method is given in bold while in the column “category” appliances with significant power spike are marked as “PS”.

Device	Category	ECO-2		REDD-2		iAWE	
		I	II	I	II	I	II
Air exhaust	one-state	98.4	98.4	-	-	-	-
Fridge	one-state (PS)	74.7	79.2	86.1	92.3	48.3	55.6
Entertainment	nonlinear	83.9	91.6	-	-	-	-
Freezer	one-state (PS)	83.6	87.5	-	-	-	-
Lamp/Light	one-state/nonlinear	55.6	55.6	71.8	78.8	-	-
Laptop	nonlinear	59.9	65.6	-	73.7	54.3	59.0
Stove	multi-state	-	-	73.5	-	-	-
TV	nonlinear	84.6	94.7	-	-	59.0	65.5
Stereo	nonlinear	84.5	85.5	-	68.1	-	-
Kitchen	-	-	-	67.8	74.1	-	-
Microwave	one-state	-	-	75.8	89.7	-	-
WM	multi-state	-	-	89.6	79.5	78.8	78.7
DW	multi-state	-	-	79.1	97.5	-	-
Disposal	one-state	-	-	97.5	-	-	-
Iron	one-state	-	-	-	-	91.2	91.2
Air Condition	continuous (PS)	-	-	-	-	45.4	50.3
Watermotor	continuous	-	-	-	87.8	57.4	62.3
Ghost	-	80.5	87.0	84.4	-	80.1	87.6

As can be seen in Table 4, there is a relation between performance improvement and appliance category with one/multi-state devices without significant power peak signature showing no performance improvement and nonlinear and continuous appliances as well as one-state appliances with significant power peak showing significant performance improvement. Depending on the dataset, the performance increase varies up to 0.4% for one/multi-state devices without power spikes, up to 7.4% for devices with power spike, up to 10.1% for nonlinear devices and up to 4.9% for continuous devices respectively. In detail the highest performance increase in the three tested datasets was observed for nonlinear appliances namely the TV (10.1%) and the Entertainment (7.7%) in the ECO-2 dataset. Significant increase in performance was also observed for devices with power spikes (PS) in

their signature, like the Fridge, the Freezer, and the A/C with maximum improvement equal to 7.4%, 3.9%, and 4.9%, respectively. The lowest or no performance improvement was observed for one-state appliances without power spikes, e.g., resistive lamps or disposal.

In order to directly compare the proposed methodology with other approaches proposed in the literature we additionally tested our method on five selected loads from the REDD-2 dataset, namely the refrigerator, lighting, dishwasher, microwave, and furnace. These loads were used in [55] because they carry a large percentage of the overall consumed energy and they have been used in other publications [67,75]. Furthermore, the disaggregation results were evaluated both in a noisy (with ghost data) and a noiseless (with synthetic data) setup as in [75] for both the one-stage and the proposed two-stage fusion architecture. The results are tabulated in Table 5.

Table 5. Performance evaluation E_{ACC} (%) for five selected appliances from the REDD-2 dataset for both one-stage (I) and proposed two-stage fusion (II) methodology.

Device	Category	REDD-2 (noisy)		REDD-2 (noiseless)	
		I	II	I	II
Fridge	one-state	80.2	93.2	87.5	94.2
Light	nonlinear	78.7	81.5	77.9	81.6
Dishwasher	multi-state	87.0	88.7	93.8	94.2
Microwave	one-state	93.1	93.7	95.6	95.8
Furnace	multi-state	82.4	83.9	87.2	87.8
Average	-	90.7	93.4	93.2	95.7

From Table 5 it is seen that the presented two-stage fusion model outperforms the baseline one-stage system in both the noisy and noiseless setup with 93.4% (2.7% improvement) and 95.7% (2.5% improvement), respectively. Moreover, the largest improvements can be observed for the appliances with significant power spikes and nonlinear behavior, i.e., the fridge and the light with 13.0% (6.7%) and 2.8% (3.7%), respectively. For the purpose of comparison with previously published NILM approaches the summary of methods using the same databases and the E_{ACC} performance metric presented in [76] was used. Furthermore, the summary of results of [76] was updated by incorporating very recent results found in the literature utilizing deep learning. However in the latest published deep learning approaches many researchers started utilizing databases with even lower sampling frequency and longer monitoring duration (e.g., AMPds [39] or UK-DALE [77]) as in [41,42,44,78], or utilizing different accuracy metrics (e.g., normalized RMSE in [45]) making direct comparison impossible. The results are tabulated in Table 6.

From Table 6, it is shown that the two-stage fusion methodology achieves higher accuracy than all other published methods evaluated on the REDD datasets 1–4 and 6. As regards the experimental setup using five appliances of the REDD-2 dataset (initially proposed in [55]) the proposed fusion architecture performs better than all reported NILM methods, except the method of Makonin et al. [75] utilizing HMM sparsity which achieved 1.4% higher accuracy than our proposed fusion methodology in the noisy set-up; however, the energy data used in [75] were manually modified to time align data acquired from two different smart meter devices while we have used the original data from the REDD-2 dataset without any modification. Moreover, for the approach presented in [75], the performance on the full REDD dataset with all 18 appliances across all houses (1, 2, 3, 4, and 6) has not been reported in the literature and thus direct comparison with our approach is possible only using the REDD-2 dataset with five devices. Regarding the results presented in [40] are not directly comparable with our approach (which performs 8.8% better) as a modified training/test setup has been used. To compare our performance with the one reported in [45] we calculated the normalized RMSE used in [45]. Our proposed methodology has normalized RMSE equal to 0.24, which is 0.11 better than the score reported in [45]. Considering the results from Tables 3 and 4, the proposed two-stage

fusion methodology demonstrated improvements both in average and per device performance across all evaluated datasets with all evaluated classifiers, demonstrating the validity of the methodology. As regards the effect of different datasets when using the same classifier, the improvement in terms of E_{ACC} varies between 0.6% and 4.1% as can be seen in Table 3. The main reasons are the different number of devices in each dataset and the distribution of appliance types, i.e., how many appliances of a specific type (e.g., one-state or nonlinear) can be found in each dataset. Considering the results in Table 3 in combination with the database categorization in Table 1 it can be seen that datasets with small number of appliances (e.g., ECO-1 or REDD-2) have a slightly higher improvement in estimation accuracy and show improvements of approximately 1.0–4.1%, while datasets with larger number of appliances (e.g., REDD-1 or REDD-3) show improvements of up to 1.6%. Moreover, the datasets including significant number of continuous appliances or nonlinear appliances (e.g., ECO-2 or iAWE) benefit more from the two-stage fusion architecture. Continuous or nonlinear devices may have high correlation with the daily routine of the users/consumers as well as they may have dependencies between them, e.g., the Entertainment appliances which in the general case are interconnected with the TV. For electrical appliances having dependencies with other devices or depending on residents' everyday routine, the a priori information of the devices operating together or following similar everyday routine patterns, e.g., most of the times working or not working at the same time, can boost the estimation of the power consumption of those devices. For such appliances, power consumption estimation can be improved from the proposed two-stage fusion methodology in which estimates of the operation of other devices (identified at the first stage of the proposed architecture) are utilized. In addition, energy consumption estimation for appliances presenting power spikes, i.e., peaks that appear during the switching on of electrical motors, e.g., in fridges or freezers, was found to get improved by the fusion stage of the proposed NILM architecture, given that the existence of a power spike in a frame changes the total amount of energy to be disaggregated. Therefore, it is beneficial having an initial estimate of which appliances are likely to be working (calculated from the first stage in the two-stage architecture), to discriminate power spikes from appliances with constant high-power consumption.

Table 6. Comparison of E_{ACC} (%) values for recently proposed NILM methodologies (methods marked with an asterisk are not directly comparable because of a dataset transferability set-up used in [40] and a slight change in the accuracy metric in [4,5]).

NILM Method	Publication	Year	Dataset	E_{ACC}	Fusion (E_{ACC})
Powerlets-PED	[79]	2015	REDD-1/2/3/4/6	72.0	79.3
Exact Deep SC	[80]	2017	REDD-1/2/3/4/6	66.1	
Greedy Deep SC	[80]	2017	REDD-1/2/3/4/6	62.6	
Discriminate SC	[81]	2010	REDD-1/2/3/4/6	59.3	
General SC	[81]	2010	REDD-1/2/3/4/6	56.4	
Temporal ML	[82]	2011	REDD-1/2/3/4/6	53.3	93.4
Sparse HMM	[75]	2015	REDD-2 (5 App.)	94.8	
SIQCP	[83]	2018	REDD-2 (5 App.)	86.4	
F-HDP-HSMM	[55]	2013	REDD-2 (5 App.)	84.8	
F-HDP-HMM	[55]	2013	REDD-2 (5 App.)	70.7	
EM-FHMM	[55]	2013	REDD-2 (5 App.)	50.8	92.3 (0.24)
CNN-RNN	[43]	2019	REDD-2 (Fridge)	87.9	
CNN*	[40]	2019	REDD-2 (Fridge)	83.5	
LSTM*	[45]	2015	REDD-2 (Fridge)	0.35	

It was shown in Tables 3–6 that the two-stage fusion methodology improved the estimation accuracy across all datasets. Especially in Table 4, it was shown that the two-stage fusion methodology shows higher performance increase for appliances with power spikes as well as nonlinear and continuous appliances. In Table 5, the results were compared to state-of-the-art literature for five selected appliances for both one-stage and proposed two-stage architecture, while a comparison of average estimation accuracy scores was presented in Table 6, showing the improvement of the method when using the complete dataset.

5. Conclusions

In this paper, a two-stage fusion energy disaggregation approach for non-intrusive load monitoring was presented. The fusion approach combines multiple classifiers producing a binary detection score in the first stage of the architecture, and further uses a fusion of the initial binary estimates to enhance the energy disaggregation accuracy during a second fusion stage. The proposed architecture was evaluated on three different databases using four different classification algorithms and proved to increase the power estimation accuracies for all evaluated databases and classifiers with Random Forests outperforming all other classifiers. Specifically, the proposed two-stage fusion methodology achieved improvement of up to 3.4% among the evaluated datasets and in device level the estimation accuracy was improved by 10.1% when compared to the best performing baseline non-intrusive load monitoring setup. As regards different appliance types, the two-stage methodology significantly improved the power consumption estimation accuracy of continuous and nonlinear devices as well as the power consumption estimation of appliances with high power spikes. The proposed two-stage fusion methodology demonstrated robust performance across several datasets with different characteristics and types of devices as well as estimated well the ghost power produced from unknown devices which is common in households, demonstrating the appropriateness of it in real-life setups. Non-intrusive load monitoring is a difficult task especially when considering nonlinear and continuous appliances. With the evolution of usage of smart meters, large amounts of energy data with duration of several continuous years of recordings is anticipated to be collected in the next years based on which deep learning approaches will be used to develop device identification and energy consumption models. Another future direction is the incorporation of temporal information into the device models to further improve disaggregation accuracy especially in the case of appliances with strongly time varying behavior.

Author Contributions: Conceptualization, P.A.S.; Methodology, I.M.; Writing—original draft preparation, P.A.S.; Writing—review and editing, I.M. and A.S.-A.; Supervision, I.M. and A.S.-A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: This work was partially supported by the UA Doctoral Training Alliance (<https://www.unialliance.ac.uk/>) for Energy in the United Kingdom.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Eurostat. Energy Statistics—An Overview. 2018. Available online: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Energy_statistics_-_an_overview#Final_energy_consumption (accessed on 27 March 2018).
2. Elma, O.; Selmogullar, U.S. A survey of a residential load profile for demand side management systems. In Proceedings of the 5th IEEE International Conference on Smart Energy Grid Engineering (SEGE), Oshawa, ON, Canada, 14–17 August 2017; pp. 85–89.
3. Mostafavi, S.; Cox, R.W. An unsupervised approach in learning load patterns for non-intrusive load monitoring. In Proceedings of the IEEE 14th International Conference on Networking, Sensing and Control (ICNSC 2017), Calabria, Italy, 16–18 May 2017; pp. 631–636.

4. Pérez-Lombard, L.; Ortiz, J.; Pout, C. A review on buildings energy consumption information. *Energy Build.* **2008**, *40*, 394–398. [[CrossRef](#)]
5. Santamouris, M.; Papanikolaou, N.; Livada, I.; Koronakis, I.; Georgakis, C.; Argiriou, A. On the impact of urban climate on the energy consumption of buildings. *Sol. Energy* **2001**, *70*, 201–216. [[CrossRef](#)]
6. Lee, D.; Cheng, C.-C. Energy savings by energy management systems: A review. *Renew. Sustain. Energy Rev.* **2016**, *56*, 760–777. [[CrossRef](#)]
7. Katipamula, S.; Brambley, M. Review article: Methods for fault detection, diagnostics, and prognostics for building systems—A review, Part II. *HVAC&R Res.* **2005**, *11*, 169–187. [[CrossRef](#)]
8. Zeifman, M.; Roth, K. Viterbi algorithm with sparse transitions (VAST) for nonintrusive load monitoring. In Proceedings of the IEEE Symposium on Computational Intelligence Applications in Smart Grid, Paris, France, 11–15 April 2011; pp. 1–8.
9. Ogwumike, C.; Short, M.; Denai, M. Near-optimal scheduling of residential smart home appliances using heuristic approach. In Proceedings of the IEEE International Conference on Industrial Technology (ICIT), Seville, Spain, 17–19 March 2015; pp. 3128–3133.
10. Indragandhi, V.; Logesh, R.; Subramaniaswamy, V.; Varadarajan, V.; Siarry, P.; Uden, L. Multi-objective optimization and energy management in renewable based AC/DC microgrid. *Comput. Electr. Eng.* **2018**, *70*, 179–198. [[CrossRef](#)]
11. Papagiannakopoulou, E.I.; Koukovini, M.N.; Lioudakis, G.V.; Garcia-Alfaro, J.; Kaklamani, D.I.; Venieris, I.S.; Cuppens, F.; Cuppens-Boulahia, N. A privacy-aware access control model for distributed network monitoring. *Comput. Electr. Eng.* **2013**, *39*, 2263–2281. [[CrossRef](#)]
12. McLaughlin, S.; McDaniel, P.; Aiello, W. Protecting consumer privacy from electric load monitoring. In Proceedings of the 18th ACM conference on Computer and Communications Security, Chicago, IL, USA, 17–21 October 2011; p. 87.
13. Buchanan, K.; Banks, N.; Preston, I.; Russo, R. The British public's perception of the UK smart metering initiative: Threats and opportunities. *Energy Policy* **2016**, *91*, 87–97. [[CrossRef](#)]
14. Vrablecová, P.; Bou Ezzeddine, A.; Rozinajová, V.; Šárik, S.; Sangaiah, A.K. Smart grid load forecasting using online support vector regression. *Comput. Electr. Eng.* **2018**, *65*, 102–117. [[CrossRef](#)]
15. Froehlich, J.; Larson, E.; Gupta, S.; Cohn, G.; Reynolds, M.; Patel, S. Disaggregated end-use energy sensing for the smart grid. *IEEE Pervasive Comput.* **2011**, *10*, 28–39. [[CrossRef](#)]
16. Gao, J.; Kara, E.C.; Giri, S.; Berges, M. A feasibility study of automated plug-load identification from high-frequency measurements. In Proceedings of the IEEE Global Conference on Signal and Information Processing (GlobalSIP), Piscataway, NJ, USA, 14–16 December 2015; pp. 220–224.
17. Koutitas, G.C.; Tassioulas, L. Low cost disaggregation of smart meter sensor data. *IEEE Sens. J.* **2016**, *16*, 1665–1673. [[CrossRef](#)]
18. Arghandeh, R.; Zhou, Y. (Eds.) *Big Data Application in Power Systems*; Elsevier: Amsterdam, The Netherlands, 2017; ISBN 9780128119686.
19. Egarter, D.; Bhuvana, V.P.; Elmenreich, W. PALDi: Online load disaggregation via particle filtering. *IEEE Trans. Instrum. Meas.* **2015**, *64*, 467–477. [[CrossRef](#)]
20. Kelly, J.; Knottenbelt, W. Does disaggregated electricity feedback reduce domestic electricity consumption? A systematic review of the literature. In Proceedings of the 3rd International NILM Workshop, Vancouver, BC, Canada, 14–15 May 2016.
21. Du, Y.; Du, L.; Lu, B.; Harley, R.; Habetler, T. A review of identification and monitoring methods for electric loads in commercial and residential buildings. In Proceedings of the IEEE Energy Conversion Congress and Expo, Atlanta, GA, USA, 12–16 September 2010; pp. 4527–4533.
22. Hart, G.W. Nonintrusive appliance load monitoring. *Proc. IEEE* **1992**, *80*, 1870–1891. [[CrossRef](#)]
23. Cominola, A.; Giuliani, M.; Piga, D.; Castelletti, A.; Rizzoli, A.E. A hybrid signature-based iterative disaggregation algorithm for non-intrusive load monitoring. *Appl. Energy* **2017**, *185*, 331–344. [[CrossRef](#)]
24. Schirmer, P.A.; Mporas, I. Energy disaggregation using fractional calculus. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020; pp. 3257–3261.
25. Gisler, C.; Ridi, A.; Zufferey, D.; Khaled, O.A.; Hennebert, J. Appliance consumption signature database and recognition test protocols. In Proceedings of the 8th International Workshop on Systems, Signal Processing and Their Applications (WoSSPA), Algiers, Algeria, 12–15 May 2013; pp. 336–341.

26. Bouhouras, A.S.; Gkaidatzis, P.A.; Panagiotou, E.; Poulakis, N.; Christoforidis, G.C. A NILM algorithm with enhanced disaggregation scheme under harmonic current vectors. *Energy Build.* **2019**, *183*, 392–407. [[CrossRef](#)]
27. Meziane, M.N.; Abed-Meraim, K. Modeling and estimation of transient current signals. In Proceedings of the 23rd European Signal Processing Conference (EUSIPCO), Nice, France, 31 August–4 September 2015; pp. 1960–1964.
28. Bilski, P.; Winiński, W. The rule-based method for the non-intrusive electrical appliances identification. In Proceedings of the IEEE 8th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Warsaw, Poland, 24–26 September 2015; pp. 220–225.
29. Sadeghianpourhamami, N.; Ruysinck, J.; Deschrijver, D.; Dhaene, T.; Develder, C. Comprehensive feature selection for appliance classification in NILM. *Energy Build.* **2017**, *151*, 98–106. [[CrossRef](#)]
30. Hassan, T.; Javed, F.; Arshad, N. An empirical investigation of V-I trajectory based load signatures for non-intrusive load monitoring. *IEEE Trans. Smart Grid* **2014**, *5*, 870–878. [[CrossRef](#)]
31. Lin, Y.-H.; Tsai, M.-S. An advanced home energy management system facilitated by nonintrusive load monitoring with automated multiobjective power scheduling. *IEEE Trans. Smart Grid* **2015**, *6*, 1839–1851. [[CrossRef](#)]
32. Bilski, P.; Winiński, W. Generalized algorithm for the non-intrusive identification of electrical appliances in the household. In Proceedings of the IEEE 9th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Bucharest, Romania, 21–23 September 2017; pp. 730–735.
33. Kim, Y.; Kong, S.; Ko, R.; Joo, S.-K. Electrical event identification technique for monitoring home appliance load using load signatures. In Proceedings of the IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 10–13 January 2014; pp. 296–297.
34. Matt Wytock, J.; Kolter, Z. Contextually supervised source separation with application to energy disaggregation. In Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, Québec City, QC, Canada, 27–31 July 2014; pp. 486–492.
35. Pathak, N.; Roy, N.; Biswas, A. Iterative signal separation assisted energy disaggregation. In Proceedings of the Sixth International Green and Sustainable Computing Conference, Las Vegas, NV, USA, 14–16 December 2015; pp. 1–8.
36. Semwal, S.; Joshi, D.; Prasad, R.S.; Raveendhra, D. The practicability of ICA in home appliances load profile separation using current signature: A preliminary study. In Proceedings of the International Conference on Power, Energy and Control (ICPEC 2013), Dindigul, India, 6–8 February 2013; pp. 756–759.
37. Figueiredo, M.; Ribeiro, B.; de Almeida, A. Electrical signal source separation via nonnegative tensor factorization using on site measurements in a smart home. *IEEE Trans. Instrum. Meas.* **2014**, *63*, 364–373. [[CrossRef](#)]
38. Piga, D.; Cominola, A.; Giuliani, M.; Castelletti, A.; Rizzoli, A.E. Sparse optimization for automated energy end use disaggregation. *IEEE Trans. Control Syst. Technol.* **2016**, *24*, 1044–1051. [[CrossRef](#)]
39. Makonin, S.; Popowich, F.; Bartram, L.; Gill, B.; Bajic, I.V. AMPds: A public dataset for load disaggregation and eco-feedback research. In Proceedings of the IEEE Electrical Power & Energy Conference, Halifax, NS, Canada, 21–23 August 2013.
40. Murray, D.; Stankovic, L.; Stankovic, V.; Lulic, S.; Sladojevic, S. Transferability of neural network approaches for low-rate energy disaggregation. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Hove, UK, 12–17 May 2019; pp. 8330–8334.
41. Barsim, K.S.; Yang, B. On the Feasibility of Generic Deep Disaggregation for Single-Load Extraction. *arXiv* **2018**, arXiv:1802.02139 2018.
42. Wu, X.; Han, X.; Liang, K.X. Event-based non-intrusive load identification algorithm for residential loads combined with underdetermined decomposition and characteristic filtering. *IET Gener. Transm. Distrib.* **2019**, *13*, 99–107. [[CrossRef](#)]
43. Çavdar, İ.; Faryad, V. New design of a supervised energy disaggregation model based on the deep neural network for a smart grid. *Energies* **2019**, *12*, 1217. [[CrossRef](#)]

44. He, W.; Chai, Y. An empirical study on energy disaggregation via deep learning. In Proceedings of the 2nd International Conference on Artificial Intelligence and Industrial Engineering (AIIE 2016), Beijing, China, 19 September–20 November 2016.
45. Mauch, L.; Yang, B. A new approach for supervised power disaggregation by using a deep recurrent LSTM network. In Proceedings of the IEEE Global Conference on Signal and Information Processing (GlobalSIP), Orlando, FL, USA, 14–16 December 2015; pp. 63–67.
46. Garcia, F.C.C.; Creayla, C.M.C.; Macabebe, E.Q.B. Development of an intelligent system for smart home energy disaggregation using stacked denoising autoencoders. *Procedia Comput. Sci.* **2017**, *105*, 248–255. [[CrossRef](#)]
47. Li, Z.; Oechtering, T.J.; Skoglund, M. Privacy-preserving energy flow control in smart grids. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Shanghai, China, 20–25 March 2016; pp. 2194–2198.
48. Chin, J.-X.; De Rubira, T.; Hug, G. Privacy-protecting energy management unit through model-distribution predictive control. *IEEE Trans. Smart Grid* **2017**, *8*, 3084–3093. [[CrossRef](#)]
49. Schirmer, P.A.; Mporas, I. Energy disaggregation from low sampling frequency measurements using multi-layer zero crossing rate. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020; pp. 3777–3781.
50. Ridi, A.; Gisler, C.; Hennebert, J. A survey on intrusive load monitoring for appliance recognition. In Proceedings of the 22nd International Conference on Pattern Recognition, Stockholm, Sweden, 24–28 August 2014; pp. 3702–3707.
51. Chang, H.-H.; Lian, K.-L.; Su, Y.-C.; Lee, W.-J. Power-spectrum-based wavelet transform for nonintrusive demand monitoring and load identification. *IEEE Trans. Ind. Appl.* **2014**, *50*, 2081–2089. [[CrossRef](#)]
52. Zhu, Y.; Lu, S. Load profile disaggregation by Blind source separation: A wavelets-assisted independent component analysis approach. In Proceedings of the IEEE PES General Meeting: Conference & Exposition, National Harbor, MD, USA, 27–31 July 2014; pp. 1–5.
53. Schirmer, P.A.; Mporas, I. Integration of temporal contextual information for robust energy disaggregation. In Proceedings of the IEEE 38th International Performance Computing and Communications Conference (IPCCC), London, UK, 29–31 October 2019; pp. 1–6.
54. Harell, A.; Makonin, S.; Bajić, I.V. Wavenilm: A Causal Neural Network for Power Disaggregation from the Complex Power Signal. In Proceedings of the 2019 IEEE International Conference on Acoustics, Speech and Signal Processing, Brighton, UK, 12–17 May 2019; pp. 8335–8339.
55. Johnson, M.J.; Willsky, A.S. Bayesian nonparametric hidden semi-Markov models. *J. Mach. Learn. Res.* **2013**, *14*, 673–701.
56. Makonin, S. Investigating the switch continuity principle assumed in Non-Intrusive Load Monitoring (NILM). In Proceedings of the IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), Vancouver, BC, Canada, 14–18 May 2016; pp. 1–4.
57. Wang, H.; Yang, W. An iterative load disaggregation approach based on appliance consumption pattern. *Appl. Sci.* **2018**, *8*, 542. [[CrossRef](#)]
58. Tabatabaei, S.M.; Dick, S.; Xu, W. Toward non-intrusive load monitoring via multi-label classification. *IEEE Trans. Smart Grid* **2017**, *8*, 26–40. [[CrossRef](#)]
59. Beckel, C.; Kleiminger, W.; Cicchetti, R.; Staake, T.; Santini, S. The ECO data set and the performance of non-intrusive load monitoring algorithms. In Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings, Memphis, TN, USA, 4–6 November 2014; pp. 80–89.
60. Kolter, J.Z.; Johnson, M.J. *REDD: A Public Data Set for Energy Disaggregation Research*; Massachusetts Institute of Technology: Cambridge, MA, USA, 2011.
61. Batra, N.; Gulati, M.; Singh, A.; Srivastava, M.B. It's different. In Proceedings of the 5th ACM Workshop on Embedded Systems for Energy-Efficient Buildings, Rome, Italy, 14–15 November 2013; pp. 1–8.
62. Andrean, V.; Zhao, X.-H.; Teshome, D.F.; Huang, T.-D.; Lian, K.-L. A hybrid method of cascade-filtering and committee decision mechanism for non-intrusive load monitoring. *IEEE Access* **2018**, *6*, 41212–41223. [[CrossRef](#)]
63. Kelly, J.; Batra, N.; Parson, O.; Dutta, H.; Knottenbelt, W.; Rogers, A.; Singh, A.; Srivastava, M. NILMTK v0.2: A non-intrusive load monitoring toolkit for large scale data sets. In Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings, Memphis, TN, USA, 4–6 November 2014; pp. 182–183.

64. van Cutsem, O.; Lilis, G.; Kayal, M. Automatic multi-state load profile identification with application to energy disaggregation. In Proceedings of the 22nd IEEE International Conference on Emerging Technologies and Factory Automation, Limassol, Cyprus, 12–15 September 2017; pp. 1–8.
65. Zoha, A.; Gluhak, A.; Imran, M.A.; Rajasegarar, S. Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey. *Sensors (Basel)* **2012**, *12*, 16838–16866. [CrossRef]
66. Shaw, S.R.; Leeb, S.B.; Norford, L.K.; Cox, R.W. Nonintrusive load monitoring and diagnostics in power systems. *IEEE Trans. Instrum. Meas.* **2008**, *57*, 1445–1454. [CrossRef]
67. Schirmer, P.A.; Mporas, I.; Paraskevas, M. Energy disaggregation using elastic matching algorithms. *Entropy* **2020**, *22*, 71. [CrossRef]
68. Schirmer, P.A.; Mporas, I. Improving energy disaggregation performance using appliance-driven sampling rates. In Proceedings of the 27th European Signal Processing Conference (EUSIPCO), La Coruña, Spain, 2–6 September 2019; pp. 1–5.
69. Pereira, L.; Nunes, N. Performance evaluation in non-intrusive load monitoring: Datasets, metrics, and tools—A review. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2018**, *8*, e1265. [CrossRef]
70. Kononenko, I.; Robnik-Sikonja, M.; Robnik, M.; Pompe, U. *ReliefF for Estimation and Discretization of Attributes in Classification, Regression, and ILP Problems*; Artificial Intelligence: Methodology, Systems, Applications; 1996; pp. 31–40. Available online: <http://lkm.fri.uni-lj.si/rmarko/papers/kononenko96-aimsa.pdf> (accessed on 22 April 2020).
71. Schirmer, P.A.; Mporas, I. Statistical and electrical features evaluation for electrical appliances energy disaggregation. *Sustainability* **2019**, *11*, 3222. [CrossRef]
72. Schirmer, P.A.; Mporas, I.; Paraskevas, M. Evaluation of regression algorithms and features on the energy disaggregation task. In Proceedings of the 10th International Conference on Information, Intelligence, Systems and Applications (IISA), Patras, Greece, 15–17 July 2019; pp. 1–4.
73. Basu, K.; Debusschere, V.; Bacha, S.; Maulik, U.; Bondyopadhyay, S. Nonintrusive load monitoring: A temporal multilabel classification approach. *IEEE Trans. Ind. Inform.* **2015**, *11*, 262–270. [CrossRef]
74. Ma, M.; Lin, W.; Zhang, J.; Wang, P.; Zhou, Y.; Liang, X. Towards energy-awareness smart building: Discover the fingerprint of your electrical appliances. *IEEE Trans. Ind. Inform.* **2017**, *1*. [CrossRef]
75. Makonin, S.; Popowich, F.; Bajic, I.V.; Gill, B.; Bartram, L. Exploiting HMM sparsity to perform online real-time nonintrusive load monitoring. *IEEE Trans. Smart Grid* **2016**, *7*, 2575–2585. [CrossRef]
76. Welikala, S.; Dinesh, C.; Ekanayake, M.P.B.; Godaliyadda, R.I.; Ekanayake, J. Incorporating appliance usage patterns for non-intrusive load monitoring and load forecasting. *IEEE Trans. Smart Grid* **2019**, *10*, 448–461. [CrossRef]
77. Kelly, J.; Knottenbelt, W. The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. *Sci. Data* **2015**, *2*, 150007. [CrossRef]
78. Kaselimi, M.; Doulamis, N.; Doulamis, A.; Voulodimos, A.; Protopapadakis, E. Bayesian-optimized bidirectional LSTM regression model for non-intrusive load monitoring. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Hove, UK, 12–17 May 2019; pp. 2747–2751.
79. Elhamifar, E.; Sastry, S. Energy disaggregation via learning ‘powerlets’ and sparse coding. In Proceedings of the Twenty-Ninth AAAI Conference on Artificial, Austin, TX, USA, 25–30 January 2015; pp. 629–635.
80. Singh, S.; Majumdar, A. Deep sparse coding for non-intrusive load monitoring. *IEEE Trans. Smart Grid* **2017**, *1*. [CrossRef]
81. Kolter, J.Z.; Batra, S.; Ng, A.Y. Energy disaggregation via discriminative sparse coding. In Proceedings of the 23rd International Conference on Neural Information Processing Systems, Kyoto, Japan, 16–21 October 2010.
82. Stankovic, V.; Liao, J.; Stankovic, L. A graph-based signal processing approach for low-rate energy disaggregation. In Proceedings of the IEEE Symposium on Computational Intelligence for Engineering Solutions (CIES), Orlando, FL, USA, 9–12 December 2014; pp. 81–87.
83. Kong, W.; Dong, Z.Y.; Ma, J.; Hill, D.J.; Zhao, J.; Luo, F. An extensible approach for non-intrusive load disaggregation with smart meter data. *IEEE Trans. Smart Grid* **2018**, *9*, 3362–3372. [CrossRef]

