

Article

MORED: A Moroccan Buildings' Electricity Consumption Dataset

Mohamed Aymane Ahajjam ^{1,2} , Daniel Bonilla Licea ³, Chaimaa Essayeh ¹, Mounir Ghogho ^{1,4,*} and Abdellatif Kobbane ²

¹ TICLab, International University of Rabat, Rabat 11100, Morocco; aymane.ahajjam@uir.ac.ma (M.A.A.); chaimaa.essayeh@uir.ac.ma (C.E.)

² ENSIAS, Mohammed V University in Rabat, Rabat 10000, Morocco; abdellatif.kobbane@um5.ac.ma

³ Faculty of Electrical Engineering, Czech Technical University, 166 36 Prague, Czech Republic; daniel.bonilla-licea@uir.ac.ma

⁴ School of EEE, University of Leeds, Leeds LS2 9JT, UK

* Correspondence: mounir.ghogho@uir.ac.ma

Received: 3 November 2020; Accepted: 10 December 2020; Published: 21 December 2020



Abstract: This paper consists of two parts: an overview of existing open datasets of electricity consumption and a description of the Moroccan Buildings' Electricity Consumption Dataset, a first of its kind, coined as MORED. The new dataset comprises electricity consumption data of various Moroccan premises. Unlike existing datasets, MORED provides three main data components: whole premises (WP) electricity consumption, individual load (IL) ground-truth consumption, and fully labeled IL signatures, from affluent and disadvantaged neighborhoods. The WP consumption data were acquired at low rates (1/5 or 1/10 samples/s) from 12 households; the IL ground-truth data were acquired at similar rates from five households for extended durations; and IL signature data were acquired at high and low rates (50 k and 4 samples/s) from 37 different residential and industrial loads. In addition, the dataset encompasses non-intrusive load monitoring (NILM) metadata.

Keywords: non-intrusive load monitoring; open dataset; electricity disaggregation

1. Introduction

Public datasets are published to help advance research fields by overstepping the laborious tasks of data acquisition and data management, and providing researchers with valuable data. One of these research fields where such datasets are useful is energy disaggregation, where researchers have been collecting and publishing datasets of electricity consumption for almost a decade now. In this paper, we focus on this kind of dataset.

Energy disaggregation revolves around the identification of the energy consumption of individual loads from the aggregated energy consumption records. It is considered a key element, particularly for low-cost and simple energy monitoring in buildings that are experiencing a growing and highly variable demand for energy [1]. Energy disaggregation gives rise to new opportunities and new applications in the energy management of microgrids [2], in smart homes [3], and in smart cities [4–6].

The lack of energy consumption datasets has long hindered the field's progression. The last decade has seen a collective effort towards alleviating this issue. Since the publication of REDD [7] in 2011, several datasets, mainly intended for energy disaggregation purposes, have been published every year. These datasets can be classified into event-based datasets (EB) and event-less datasets (EL). The main difference between EB and EL datasets is that the former type provides events that are additional information specifying equipment's state changes happening in the consumption

measurements. In general, such datasets provide different consumption data: whole premises (WP) aggregate consumption, individual circuit consumption (IC), and individual load consumption (IL). It is important to note that:

- WP measurements reflect the global consumption of premises. They are taken from the main switchboard.
- IL measurements reflect the consumption of individual loads. There are two types:
 - Signatures (or traces), which are unique consumption patterns of loads [8], monitored individually for short periods of time independently of other measurements (as in HFED [9] or Tracebase [10]).
 - Ground-truth consumption data (also called plug-level data), which are the consumption of loads measured at individual plugs of monitored premises for longer durations and are often taken in parallel with the acquisition of the WP consumption (as in BLUED [11], Dataport [12], or BLOND [13], etc.).
- IC measurements reflect the circuit-level energy measurements from the premises' electrical mains. IC consumption data can describe the consumption of a combination of loads or those of individual loads, in which case IC measurements become similar to IL measurements.

Such datasets reflect the energy consumption in residential or commercial premises; therefore, they have proven to be useful in energy management applications, including load disaggregation and forecasting, energy efficiency, photovoltaic systems sizing, etc. In addition, their public release has highlighted several issues regarding the collection and utilization of the datasets [14]. Apart from the obvious differences in granularity and formats in which data are provided, building datasets raise challenges in the acquisition phase (e.g., difficulty in acquiring IL data from all plugs available in a commercial building, and difficulties in accessing plugs of some loads), or in the events labeling process (e.g., difficulty to log all events as they happen), resulting in short durations of the monitoring of WP consumption with events, or even in missing or mismatching data between the consumption and the events list. In addition, the lack of a general consensus over suitable metrics to assess researchers' use of datasets may hinder achieving higher disaggregation performance [15].

Nevertheless, there has been a boost of interest in the field. For instance, Batra et al in [16] and then in [17], provided the Non-Intrusive Load Monitoring Toolkit (NILMTK) (v0.1 then v0.2), which is a Python-based open source toolkit; in addition to containing multiple open datasets, the toolkit facilitates their processing (e.g., parsing, processing, and diagnosing datasets) and the assessment of disaggregation algorithms. Further, Kelly and Knottenbelt in [18] provided NILM Metadata, a hierarchical metadata schema specific to energy disaggregation with controlled vocabularies to represent appliances, meters, buildings, and datasets. More recently, Pereira in [19] proposed a common data model with three main data entities deemed essential in any dataset for energy disaggregation tasks: (i) consumption data embracing raw waveforms (i.e., current and voltage) and processed waveforms (i.e., power metrics); (ii) ground-truth data which can report the appliance activities (i.e., power events information), the user activities (e.g., doing laundry), and the consumption of individual appliances (or individual circuit consumption); (iii) data annotations, which can be either metadata or comments. The author also proposed a file format (i.e., Waveform Audio File Format) that offers several features (e.g., waveform data and annotations stored in a single compact file favoring coherent and structured datasets, optimized files with minimal overheads favoring faster dataset manipulation) that are most beneficial for creating and manipulating such datasets.

The contribution of this paper is twofold: (i) it provides an overview of prior open datasets of electricity consumption, and (ii) it introduces a new electricity consumption dataset, coined MORED, and the methodology used for its acquisition (see Figure 1). Although not limited to these fields, MORED is intended for energy-related research fields such as load disaggregation and energy forecasting. The novelty of this dataset resides in several parts:

- MORED is the first dataset collected from Moroccan households.
- Unlike many other datasets providing a single type of consumption measurement, MORED provides three kinds: WP, WP, and IL ground-truth at low rates, and IL signatures at different sampling rates.
- MORED contains event lists describing on/off transitions in WP electricity consumption of monitored IL (i.e., IL ground-truth), and various state transitions in IL signatures.
- MORED WP electricity consumption was acquired following a process relying on different parameters, such as the neighborhood's socio-economic status and the household's number of residents, in order to capture a representative sample of the Moroccan urban context.
- MORED IL consumption was acquired in both residential and industrial settings.
- MORED includes NILM metadata, providing more details about the monitored loads and premises.

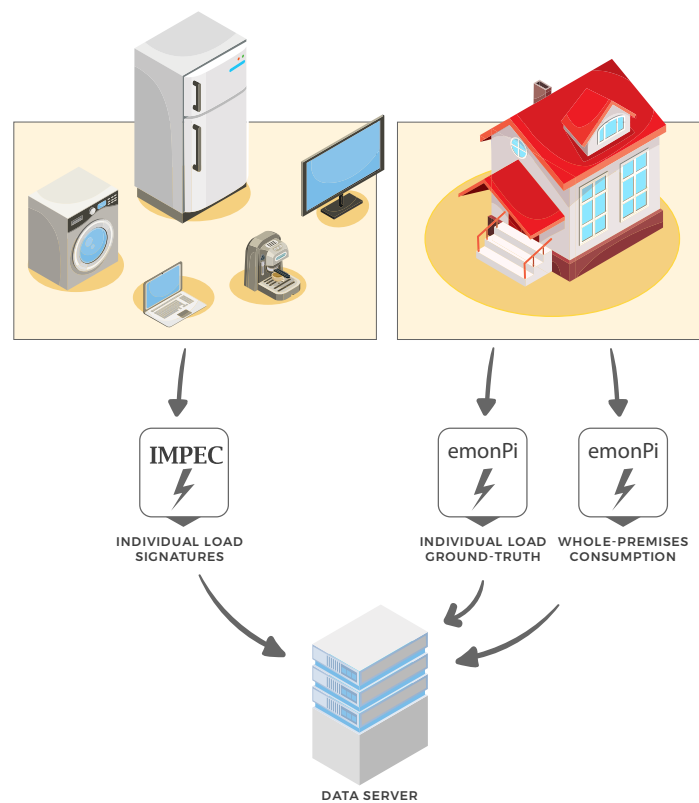


Figure 1. MORED acquisition architecture of whole premises (WP) and individual load (IL) electricity consumption.

This paper is organized in two parts: Section 2 presents a detailed survey of datasets for energy disaggregation published up to August 2019, highlighting their categories, features, and advantages. Section 3 introduces the proposed dataset, MORED.

2. Previous Work

To the best of our knowledge, there are 28 open datasets for energy disaggregation up to August 2019. Table 1 presents the nomenclature of the electrical quantities used in the remainder of the paper, while Table 2 outlines all known datasets. We begin by presenting a summary of EL datasets and then that of EB datasets.

Table 1. Electrical quantities nomenclature.

Symbols	Electrical Quantity
$i(t)$	Instantaneous current
$v(t)$	Instantaneous voltage
I	RMS current
V	RMS voltage
ϕ	Phase difference
f	Fundamental frequency
\bar{I}	Average current
P	Active power
Q	Reactive power
S	Apparent power
Pt	Active energy
Qt	Reactive energy
St	Apparent energy
PF	Power factor
DPF	Displacement power factor

2.1. EL Datasets

In EL datasets, researchers provide consumption measurements without information about loads' state transitions. A quick glance at Table 2 shows that most datasets belong to this category, because the acquisition of the associated measurements is easier and less time-consuming [14].

The following is an exhaustive description of the EL datasets:

- **Reference Energy Disaggregation Dataset (REDD) [7]:** It provides single-phase electric consumption data from several residential settings from the US. REDD comprises $i(t)$ and $v(t)$ sampled at 15k[S/s]; \bar{S} at WP levels extracted at a 1[S/s] rate and at IC levels extracted at a 1/3[S/s] rate. All data were acquired with 24-bit resolution Analog to Digital Converters (ADC). A total of six houses were monitored for WP data, and a total of eight distinct classes of appliances were monitored for IL data, over a total period of 119 days. All file measurements are provided as generic data (DAT) files. The official website of REDD is: <http://redd.csail.mit.edu>; it can be downloaded upon requesting the logging credentials by email.
- **Household Electricity Use Study (HES) [20]:** This study has targeted 251 different UK households (terraced houses, semi-detached houses, detached houses, bungalows, and flats) with different residents (single pensioner, single non-pensioner, multiple pensioners, household with children, and multiple persons with no dependent children) for different periods (26 households were monitored for a whole year, and the rest were monitored for a month). IL measurements were made of roughly 19 distinct classes of appliances extracted at 1/120[S/s] to 1/300[S/s] rates, in addition to acquisitions of outdoor and indoor temperatures sampled at a 1/120[S/s] rate and then averaged over 10 min intervals. In addition, all household residents were asked to fill in forms to describe their usual usage of several of the monitored appliances (washing machines, washer dryers). The official website of the study, although not containing the HES dataset, is: <http://randd.defra.gov.uk/Default.aspx?Menu=Menu&Module=More&Location=Nonone&ProjectID=17359>; a software portal was said to be under development to provide access to the dataset, yet no updates have been made and the data still are not available.
- **The Tracebase dataset [10]:** It provides IL power measurements data (called traces) of single-phase electrical appliances from more than 10 houses and office spaces in Germany. It comprises \bar{P} extracted at 1[S/s] and 1/8[S/s] rates, acquired from 40 different appliance classes over a total of 1883 days. All file measurements are provided in Comma-Separated Values (CSV) format. The official website of the Tracebase dataset is: <https://github.com/areinhardt/tracebase>.

- **Individual Household Electric Power Consumption (IHEPC) Dataset [21]:** It provides electric power measurements of a single French household. The dataset comprises \bar{I} , \bar{V} , \bar{P} , \bar{Q} extracted at 1/60[S/s] rates and monitored at WP, and $\bar{P}t$ extracted at the same rates and monitored at IC levels. All data were acquired over a total period of almost 1460 days. All file measurements are provided in standard text (TXT) format. The official website of the IHEPC dataset is: <http://archive.ics.uci.edu/ml/datasets/Individual+household+electric+power+consumption>.
- **UMass Smart* Datasets [22]:** It provides a variety of data describing the energy consumption in numerous US households. There is a 2013 release holding two datasets: Home dataset and Microgrid dataset. A 2017 release brings additional data to the previously released datasets, in addition to the publishing of new datasets: one called Apartment dataset related to energy disaggregation and others related to solar energy and occupancy monitoring.
 - The 2013 release of Microgrid dataset comprises \bar{P} extracted at a 1/60[S/s] rate and acquired for a whole day from 443 homes.
 - The 2017 release of Home dataset is a continuation of the previous one (i.e., the 2013 release of Home dataset). It provides WP/IC environmental and electrical data from more households (seven households) and comprises P extracted every 1/1800[S/s] over a longer period of time (i.e., 1095 days).
 - The 2017 release of apartment dataset comprises \bar{P} extracted at a 1/60[S/s] rate and acquired for a whole day from 114 apartments.

All file measurements are provided in CSV format. The official website of the UMass Smart* datasets is: <http://traces.cs.umass.edu/index.php/Smart/Smart>.

- **Almanac of Minutely Power Dataset (AMPds) [23,24]:** It provides a variety of data describing energy consumption (i.e., electricity, water, and natural gas) of one Canadian household. It has a 2013 release (without environmental and utility billing data) and a 2016 release (i.e., AMPds2). The electrical data comprises $i(t)$, $v(t)$, P , Q , S , f , Pt , Qt , St , PF , and DPF , sampled at 1/60[S/s] rates, at WP and IC levels (of specific appliances or household rooms), over a monitoring period of 365 days in the first release and 730 days in the second. All file measurements are provided in CSV and HDF5 (H5) formats. The official website of AMPds is: <http://ampds.org>.
- **Indian Dataset for Ambient Water and Energy (iAWE) [25]:** It provides a variety of data describing ambient, water and electricity consumption of one Indian household. The electrical data comprises I , V , P , Q , S , f , and PF , extracted at 1[S/s] rates and monitored at WP, IC and IL levels (nine distinct appliance classes) over a monitoring period of 73 days. Additionally, event labels were provided for a single day. All file measurements are provided in CSV format. A preprocessed subsection of the same dataset is provided in H5 format. The official website of the iAWE dataset is: <http://iawe.github.io>.
- **Database of appliance consumption signatures (ACS-F1 and ACS-F2) [26,27]:** It provides IL electricity consumption data taken from hundreds of Swiss houses. There is a 2013 release (called ACS-F1) and a 2014 release (called ACS-F2). Both datasets comprise I , V , P , Q , f and ϕ of 10 classes of appliances (a total of 100 appliances) for the ACS-F1 and of 15 classes of appliances (a total of 225 appliances) extracted at 1/10[S/s] rates over a period of one hour (ACS-F1 had two acquisition sessions of one hour). All file measurements are provided in Extensible Markup Language (XML) and MATLAB (MAT) formats. The official websites of the ACS-F1 and ACS-F2 datasets are: <https://icosys.ch/acs-f1> and <https://icosys.ch/acs-f2>, respectively.
- **Green Electrical ENergy Dataset (GREEND) [28]:** It provides electric power measurements of nine Italian and Austrian households. The dataset comprises only P extracted at 1[S/s] and monitored at WP and IL levels over a total period of almost 356 days. All file measurements are provided in CSV and in H5 formats. The official website of GREEND is: <http://www.andreatonello.com/greend-energy-metering-data-set>.

- **UK Domestic Appliance-Level Electricity (UK-DALE) dataset [29]:** It provides electric power measurements from several households in the UK. Since its first release in 2015, it had three other releases over the years (a second 2015 release, a 2016 release and a 2017 release). The 2015 releases comprise $i(t)$ and $v(t)$ sampled at 16k[S/s], V , P , and S extracted at 1[S/s], and S extracted at 1/6[S/s] at WP levels, and P extracted at 1/6[S/s] at IC and IL levels across three households, S and P extracted at 1/6[S/s] at WP and IC/IL levels respectively across two additional households. IL data comes with room occupancy information with which the authors mentioned the possibility to infer the on/off switch events of several loads. All data were acquired with 20-bit resolution ADCs. The following releases (i.e., 2015, 2016 and 2017) all comprise elongated periods of WP measurements at a specific household from 665 days (in the 2015 original release), to 890 days (in the 2015 August release), to 1264 days (in the 2016 release), to 1569 days in the final release. The low frequency data are provided in CSV and H5 formats and the high frequency measurements data are provided in Free Lossless Audio Codec (FLAC) format. The official website of the UK-DALE dataset is: <http://jack-kelly.com/data>.
- **Electricity Consumption and Occupancy (ECO) dataset [30]:** It provides electricity consumption and occupancy data from six Swiss households. The electricity consumption data comprise I , V , and ϕ extracted at 1[S/s] measured at WP and IL levels (of 19 classes) over 240 days. All file measurements are provided in CSV and MAT formats. The official website of the ECO dataset is: <https://www.vs.inf.ethz.ch/res/show.html?what=eco-data>.
- **Plug Load Appliance Identification Dataset (PLAID) [31,32]:** It provides single-phase IL electricity consumption data from 60 US households. It has a 2014 release (PLAID 1) and a 2017 release (PLAID 2). Both releases comprise two seconds to 10s windows of I and V measurements extracted at 30k[S/s], which contain the steady-state operation and the startup transient when available of 11 different appliance classes. All file measurements are provided in CSV format. The official website of PLAID is: <http://www.plaidplug.com>.
- **Residential Building Stock Assessment (RBSA) dataset [33]:** The Northwest Energy Efficiency Alliance (NEEA) has conducted the residential building stock assessment studies in 2014 (RBSA I) and in 2017 (RBSA II) aiming to shed light on the US residential energy use, its determinants, and energy saving opportunities. Both datasets provide IL electricity consumption, comprising the Pt extracted at 1/900[S/s] of 15 classes of appliances over 810 consecutive days. All file measurements are provided in TXT format. The official website of the RBSA dataset is: <https://nea.org/data/residential-building-stock-assessment>.
- **Commercial Building Energy Dataset (COMBED) [34]:** It provides electric power measurements from an Indian academic building. The dataset comprises I , P , and Pt extracted at 1/30[S/s] and measured at WP and IC levels for 30 days. All file measurements are provided in CSV and in H5 formats. The official website of COMBED is: <https://combed.github.io>.
- **High Frequency EMI Dataset (HFED) [9]:** It provides electromagnetic interference (EMI) IL consumption measurements in a controlled laboratory settings and in a residential apartment in India. While data from the residential settings are no longer available online since 2016, the dataset comprises the EMI traces from laboratory settings of 24 loads in different setups and combinations sampled at 10k[S/s] to 5M[S/s]. All file measurements are provided in CSV format. The official website of HFED is: <https://hfed.github.io>.
- **REFIT: Electrical Load Measurements [35]:** It provides electric power measurements from 20 English households. The dataset comprises P extracted at 1/8[S/s] measured at WP and IL levels (nine classes) for a continuous period of almost 712 days. All file measurements are provided in CSV format. The official website of the REFIT dataset is: <https://pureportal.strath.ac.uk/en/datasets/refit-electrical-load-measurements>.
- **Dataport dataset [12]:** It provides several climate related data and energy usage related data in real household settings from the US, and testbeds. The dataset comprises electricity consumption data such as current and P sampled at 1/60[S/s] measured at WP and IL levels in almost

1000 households. A portion of the dataset is available in H5 format. The official website of the Dataport dataset is: <https://dataport.pecanstreet.org>.

- **Dutch Residential Energy Dataset (DRED)** [36]: It provides several ambient data and Electricity monitoring data, and occupancy and general information of one household from The Netherlands. First, the ambient data comprises indoor room temperature, outdoor temperature, both sampled at 1/60[S/s], and also wind speed, humidity, and precipitation parameters sampled at 1/3600[S/s]. In addition, the occupancy information comprises room-level location information of occupants using smartphones or smartwatches sampled at 1/60[S/s]. The electricity monitoring data comprises P extracted at 1[S/s] and measured at WP and IL levels over a period of 180 days. All file measurements are provided in CSV and in H5 formats. The official website of DRED is: <http://www.st.ewi.tudelft.nl/akshay/dred>.
- **Worldwide Household and Industry Transient Energy Dataset (WHITED)** [37]: It provides single-phase measurements of numerous residential and industrial IL from Germany, Austria, Indonesia, and recently Canada. The dataset comprises 5s of $i(t)$ and $v(t)$ start-up measurements sampled at 44k[S/s] from 46 distinct classes of appliances. All file measurements are provided in FLAC format. The official website of WHITED is: <https://www.in.tum.de/i13/resources/whited>.
- **Controlled On/Off Loads Library (COOLL) dataset** [38]: It provides single-phase IL electricity consumption data taken from a French University. The dataset comprises 6s of $i(t)$ and $v(t)$ start-up measurements (i.e., containing turn-on transients) sampled at 100k[S/s] from 12 distinct classes of controllable (with on/off switches) appliances. Researchers performed measurements using an acquisition system coupled with an additional control system allowing for a precise control of the turning-on and turning-off of appliances; hence, they were able to capture 20 different variations of the energy consumption for every appliance. Each of these variations present different transients. All file measurements are provided in FLAC format. The official website of the COOLL dataset is: <https://coolldataset.github.io>.
- **Electric demand profiles (ELP)** [39,40]: It provides single-phase electric demand measurements from 23 Canadian households of different types (i.e., semi-detached, detached, and row houses). The dataset comprises \bar{P} extracted at 1/60[S/s] and monitored at WP and IL levels over a whole year. All file measurements are provided in CSV format. The official website of the ELP dataset is: <https://carleton.ca/sbes/?p=313>.
- **The Rainforest Automation Energy (RAE) dataset** [41]: It provides environmental and electrical data from two Canadian households. The environmental data consist of various measurements from different thermostats model, sampled at approximately 1/300[S/s]. The single-phase electrical data comprises $i(t)$, $v(t)$, P , Pt , Q , Qt , S , St , f , and PF sampled at 1[S/s] and monitored at WP and IC levels over a monitoring period of 72 days in the first house and 59 days in the second. All file measurements are provided in CSV format. The official website of the RAE dataset is: <https://energy.duke.edu/content/rainforest-automation-energy-dataset-rae>.
- **Laboratory-measured Industrial Appliance Characteristics (LILACD)** [42]: It provides three-phase and single-phase industrial loads consumption measurements from Germany. It comprises IL measurements of $i(t)$ and $v(t)$ sampled at 50k[S/s] and spanning from a couple of seconds to several minutes depending on the appliance type. A total number of 15 distinct classes of appliances were monitored. Acquisitions were conducted taking into account loads' states, timing patterns of loads' on/off cycles, and up to three combination of loads, producing a total of 1302 measurements: 381 of single appliances, 864 of a combination of two appliances, and 56 of a combination of three appliances. All data were acquired using a 16-bit resolution. All file measurements are provided in the Technical Data Management Streaming (TDMS) format. The official website of the LILACD is: <https://www.in.tum.de/i13/resources/lilacd>.

Table 2. Summary of all open datasets specific to electricity disaggregation; (R: residential, C: commercial, E: educational, I: industrial, NS: not specified).

Dataset	Year(s) of Release	Type of Setting	Acquisition Level			WP Span (Days)	Features (refer to Table 1)						Event Labels	File Format		
			WP	IC	IL		I	V	P	Q	S	Other		CSV	H5	Other
REDD (US) [7]	2011	R	✓	✓	✓	119 ^a	✓	✓	.	.	✓	DAT
UMass Smart* (US) ^b [22]	2013,17	R	✓	✓	✓	90–1095	.	✓	✓	.	✓	<i>f</i>	.	✓	.	.
BLUED (US) [11]	2012	R	✓	✓	✓	7	✓	✓	✓	✓	.	.	✓	.	.	TXT, MAT
HES (UK) [20]	2012	R	.	.	✓	30–365	.	.	✓	NS	NS	NS
Tracebase (DE) [10]	2012	R/C	.	.	✓	1883	.	.	✓	✓	✓	.
IHEPC (FR) [21]	2012	R	✓	✓	.	1460	✓	✓	✓	✓	.	<i>Pt</i>	.	.	.	TXT
AMPds (CA) [23,24]	2013,16	R	✓	✓	✓	NS	✓	✓	✓	✓	.	<i>f,φ,Pt</i>	.	✓	✓	.
iAWE (IN) [25]	2013	R	✓	✓	✓	73	✓	✓	✓	✓	.	.	✓ ^c	✓	.	.
ACS-F1&2 (CH) [26,27]	2013,14	R	.	.	✓	NS	✓	✓	✓	✓	.	<i>φ</i>	.	.	.	MAT, XML
GREEND (IT, AT) [28]	2014	R	.	.	✓	90–120	.	.	✓	✓	✓	.
UK-DALE (GB) [29]	2015,17	R	✓	✓	✓	1569	✓	✓	✓	.	✓	.	.	✓	✓	FLAC
ECO (CH) [30]	2014	R	✓	.	✓	240	✓	✓	✓	✓	.	.	.	✓	.	XML
PLAID 1&2 (US) [31,32]	2014,17	R	.	.	✓	NS	✓	✓	✓	.	.
RBSA I&II (US) [33]	2014,17	R	.	.	✓	810	<i>f,Pt</i>	.	✓	.	.
COMBED (IN) [34]	2014	E	✓	✓	.	30	✓	.	✓	.	✓	.	.	✓	✓	.
HFED (IN) [9]	2014	R/E	.	.	✓	NS	✓	.	.
REFIT (UK) [35]	2015	R	✓	.	✓	712	.	.	✓	✓	.	.
Dataport (US) [12]	2015	R	✓	✓	.	1424	✓	.	✓	✓	.
DRED (NL) [36]	2015	R	✓	.	✓	180	.	.	✓	✓	✓	.
WHITED (DE, AT, ID) [37]	2016	R	.	.	✓	NS	✓	✓	MAT
COOLL (FR) [38]	2016	E	.	.	✓	NS	✓	✓	MAT
SustDataED (PT) [43]	2016	R	✓	.	✓	10	✓	✓	✓	✓	.	.	✓	.	.	surf
ELP(CA) [40]	2011,16	R	✓	✓	✓	356	.	.	✓	✓	.	.
RAE (CA) [41]	2018	R	✓	✓	.	72	✓	✓	✓	✓	.	<i>f,φ</i>	.	.	.	TXT
BLOND-50&-250 (DE) [13]	2018	E	✓	.	✓	213–50	✓	✓	✓	.	✓	.
EMBED (US) [44]	2018	R	✓	.	✓	14, 21, 27	.	.	✓	✓	.	.	✓	✓	.	MAT
LILACD (DE) [42]	2019	I	.	.	✓	NS	✓	✓	TDMS
PRECON (PK) [45]	2019	R	✓	✓	✓	243–365	.	.	✓	✓	.	.
MORED (MA)	2020	R/I	✓	.	✓	15–81	✓	✓	✓	✓	.	.	✓	✓	.	TDMS

^a The total duration of the acquisition across six households. ^b This dataset presents various releases; two of them are EL and one is EB (refer to Section 2 for more details). ^c This dataset provides only a single day of events.

- **Pakistan Residential Electricity Consumption Dataset (PRECON) [45]:** It is the latest published EL dataset that provides single-phase and three-phase electric demand measurements from 42 Pakistani households of different demographics and socio-economic backgrounds. The dataset comprises \bar{P} extracted at 1/60[S/s] and monitored at WP and at IC and IL levels of several chosen rooms and appliances, over a duration ranging from eight months to a full year. In addition, PRECON offers meta-data containing several attributes of the monitored households, including monthly income of the household, number of residents, area of the household, etc. All file measurements are provided in CSV format. The official website of the PRECON dataset is: <http://web.lums.edu.pk/~eig/precon.html>.

2.2. Eb Datasets

As opposed to EL datasets, the EB datasets contain energy consumption measurements coupled with lists of events that describe which and when state transitions occur for each load as they are used in the premises.

Only five datasets (BLUED, SustDataED, BLOND, EMBED, and UMass Smart* Home Data Set) are EB, with EMBED being the latest and the largest covering three houses over two to three weeks [44] (see Table 2).

The following is an exhaustive description of EB datasets:

- **Building-Level fully labeled Electricity Disaggregation (BLUED) dataset [11]:** A portion of BLUED is published, providing single-phase consumption readings of $i(t)$ and $v(t)$ sampled at 12k[S/s] rates at WP, IC, and IL levels from a single household in the US over a seven days period. BLUED also contains the event list reporting loads' power state changes. All file measurements are provided in TXT and MAT formats. The official website of the dataset is: <http://portoalegre.andrew.cmu.edu:88/BLUED>. It can be downloaded upon requesting the files password by email.
- **UMass Smart* Datasets [22]:**
 - The 2013 release of Home dataset contains environmental and single-phase electrical data from three households over a period of 90 days. First, environmental data describe indoor and outdoor weather data comprising averaged temperature, humidity, wind, and rainfall metrics sampled at a 1/60[S/s] rate. Second, electrical data describe consumed energy inside households and generated energy from solar panels and wind turbines near one of the households. The electrical data comprises \bar{P} and \bar{S} extracted at a 1[S/s] rate, at WP level (from every household) and IC level, P extracted at an average rate of 1/2.5[S/s] at IL level, and $v(t)$ and f sampled at a 1[S/s] rate at WP level. In addition, event data from a single household were provided describing: (i) wall switch events (on, off, dim in percentage) at most of the wall lighting switches; (ii) thermostat events describing the homes heating and cooling systems state (on/off), changes in temperature (in Fahrenheit), and changes in their set-point (in Fahrenheit); (iii) motion events (yes/no) corresponding to an activity being detected in a previously dormant room, or an activity that has not been detected for two minutes in a previously active room, and describing occupancy in six rooms of one household; (iv) door events (open/close) corresponding to the kitchen refrigerator, its freezer compartment, and the basement freezer. The latter data of generated energy of one household comprises \bar{I} from the solar panels and turbines, and \bar{V} from the attached battery extracted at a 1/5[S/s] rate. All file measurements are provided in CSV format.
- **Public Datasets for Sustainability and Electric Energy Research: Energy Disaggregation Research (SustData-ED) [43]:** SustData-ED, an extension of the SustData dataset is published, providing single-phase electricity consumption data, and room occupancy data of a Portuguese single-family household over 10 days. SustData-ED contains $i(t)$ and $v(t)$ sampled at 12.8k[S/s] rates and P and Q extracted at 50[S/s] rates at WP levels, and I and V extracted at 50[S/s] rates at

IL levels (17 individual loads), and supplemented with event lists describing the state transitions of those loads. All file measurements are provided as bespoke SURF files. The official website of SustData-ED is: <http://aveiro.m-iti.org/data>. Data can be downloaded upon requesting the logging credentials by email.

- Building-Level Office eNvironment Dataset (BLOND) [13]:** It provides three-phase power grid electricity consumption measurements from a German academic building. It was released under two versions: BLOND-50 and BLOND-250. Both versions comprise WP and IL level readings of $i(t)$ and $v(t)$, with the former spanning 213 days of measurements sampled at 50k[S/s] for WP data and 6.4k[S/s] for IL data, and the latter spanning 50 days at 250k[S/s] for WP data and 50k[S/s] for IL data. Additionally, the ground truth data (i.e., IL) are coupled with relative events. All file measurements are provided in H5 format. The official website of the BLOND dataset is: <https://mediatum.ub.tum.de/1375836>.
- Energy Monitoring through Building Electricity Disaggregation (EMBED) dataset [44]:** EMBED is one of latest open EB datasets, providing single-phase electricity consumption from three apartments in the US over periods spanning 14, 21, and 27 days. EMBED contains WP data of P and Q extracted at 60[S/s], and IL data of P extracted at 1 to 2[S/s] rates, in addition to fully labeled events. All file measurements are provided as MAT files, except IL measurements that are provided as CSV files. The official website of EMBED is: <http://embed-dataset.org>.

Typically, datasets such as REDD and UK-DALE are employed frequently for various residential disaggregation studies in the literature. This can be attributed to many aspects, namely, being one of the first open datasets and having rich data acquired for long durations of time. Nevertheless, newer datasets such as EMBED, PRECON, and MORED can provide more consumption data and help alleviate the lack of diversity in the data when developing solutions by researchers or distribution companies alike.

As can be noticed from the summary above that a taxonomy based on the existence of event lists in the datasets is sufficient to determine which one to use for an event-based process. For instance, Figure 2 presents a taxonomy based on the provided types of data, while Figure 3 presents a taxonomy based on the provided consumption data and their sampling rates. These taxonomies can help discriminate between datasets, showcase their intrinsic features, and help researchers pinpoint the right dataset or combinations of datasets to use. In addition, such taxonomies can also help identify specific data scarcity or potential problems in the available open datasets.

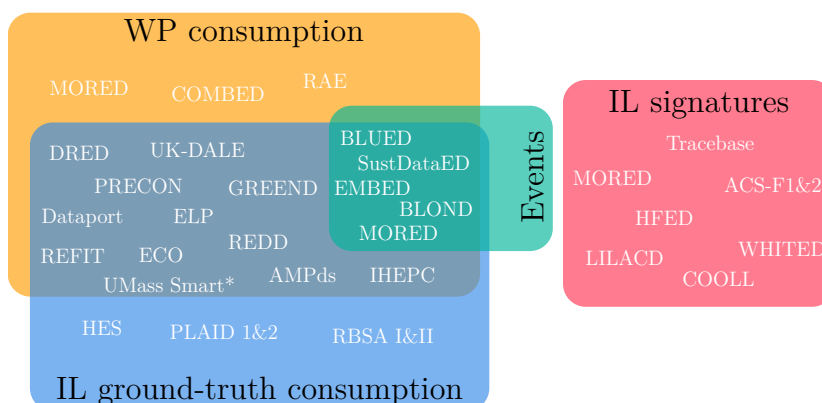


Figure 2. Venn diagram of a taxonomy based solely on the provided types of data.

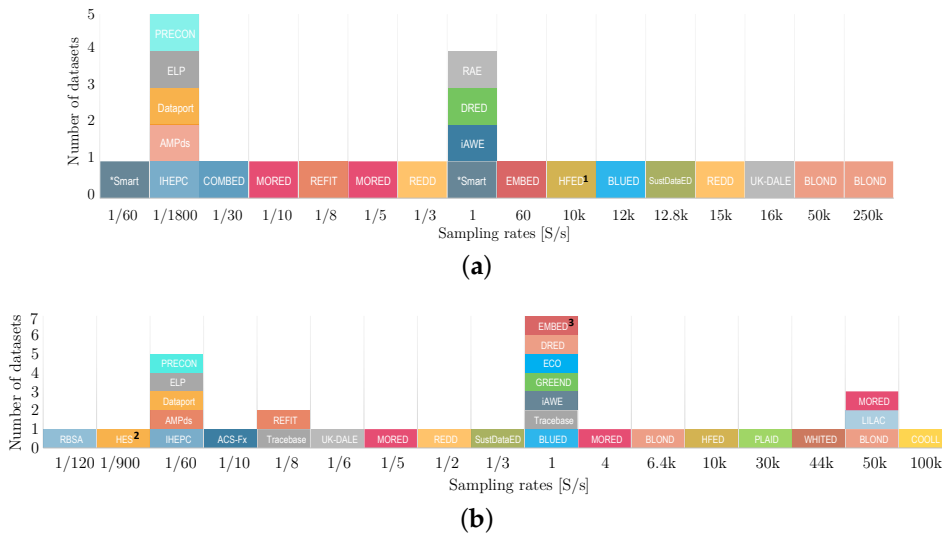


Figure 3. Histograms of a taxonomy based on the provided electricity consumption data and sampling rates. (a) Datasets providing WP and/or IC data; (b) Datasets providing IL data. ¹: Sampling rates of HFED are between 10k(S/s) and 5M(S/s); ²: Sampling rate of HES is 1/120[S/s] or 1/300[S/s]; ³: Sampling rate of EMBED is 1[S/s] or 2[S/s].

3. Data Acquisition Campaign

During spring and summer of 2019 and 2020, a data acquisition campaign was conducted to collect data reflecting the electricity consumption of different urban premises in different Moroccan cities. We note that the consent of all participating residents was gained prior to any data acquisition. In addition, anonymity was conserved in all gathered data in order to maintain their right for privacy.

3.1. Demographic Sample

Most Moroccan households are fed with a 220 V 50 Hz two-wire single-phase power circuit, with the exception of large houses in affluent neighborhoods that can be fed with a three-phase voltage supply. Morocco is a developing country with a large proportion of its population having low purchasing power; hence, availability and usage of power-hungry loads such as air-conditioners and microwaves are generally limited to affluent neighborhoods. In addition, given that the electricity consumption patterns and types of loads may differ based on the socio-economic status of the residents, we have taken into account the size, type, and location of premises, with the view of obtaining a representative sample of the country’s electricity consumption profiles. As can be seen in Figure 4, MORED acquisition campaign targeted residential premises such as apartments and semi-detached houses from affluent neighborhoods, and apartments from disadvantaged neighborhoods of Moroccan urban areas, along with laboratories at the International University of Rabat. In this section, we describe the mechanisms of the data acquisition campaign.

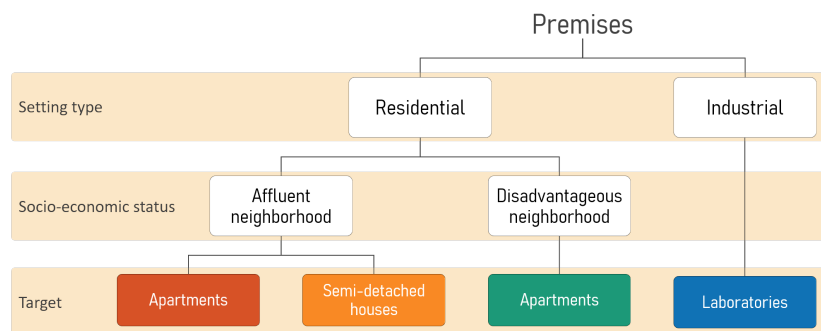


Figure 4. Premises targeted for MORED data acquisition.

3.2. Acquisition System

Power acquisitions were made possible using two different systems: emonPi and IMPEC [46]. The former was used for WP and IL ground-truth consumption acquisition at low sampling rates. IMPEC was used for IL signature consumption acquisition at high sampling rates.

IMPEC—an integrated system for monitoring and processing electricity consumption in buildings (see Figure 5a)—is a system that we have developed and presented in [46]. IMPEC houses a FPGA, a real-time processor, and two IO modules of 24-bit and 16-bit resolution ADCs for current and voltage measurements respectively. In fact, voltage is measured directly using the IO module, while the current is measured using split-core current transformers (SCT), with a single SCT for a single-phase power supply and three SCTs for a three-phase supply. Compared with commercial acquisition systems, IMPEC is tailored for NILM purposes and offers several advantages, such as acquiring electricity consumption at a high sampling frequency (50k[S/s]), thereby allowing one to extract electrical features at a frequency of 4[S/s], and logging data in TDMS files. It also provides control over key acquisition aspects via its user graphical interface (i.e., acquisition rate, extraction rate, length of snapshots of raw waveforms), and allows metadata information logging (e.g., information about premises and appliances).

EmonPi is an open-source energy monitoring system based on raspberry Pi (see Figure 5b) that houses an Arduino compatible micro-controller (i.e., The Atmel ATmega328) with 10-bit resolution ADCs for current and voltage measurements. Similarly to IMPEC, the current measured with emonPi is based on SCTs, while the voltage is measured using a plug-in AC-AC adapter. Multiple emonPis were used to extract features describing WP consumption and IL ground-truth consumption of several appliances of a household at a frequency of 1/5[S/s] or 1/10[S/s] and log data in CSV files. In another hand, a Web app and a native-android app called EmonCMS corresponding to each emonPi were used by household residents in order to monitor their real-time and daily power consumption.

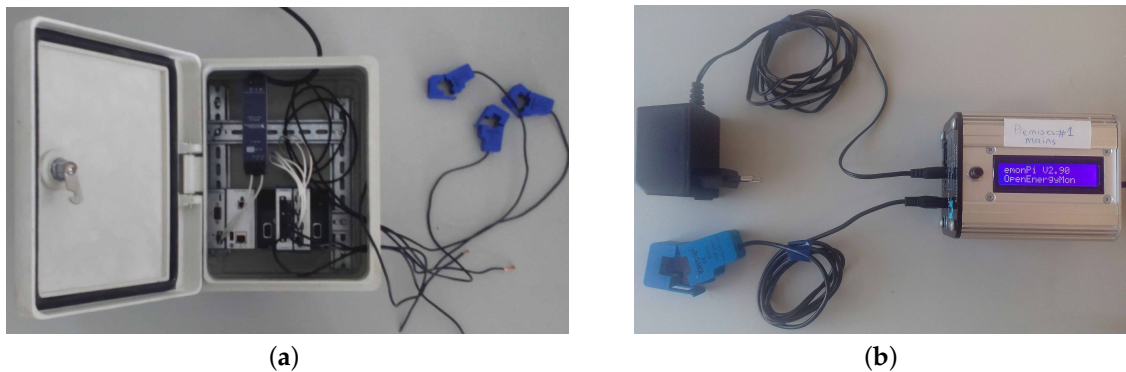


Figure 5. Acquisition systems used in the acquisition campaign of MORED. (a) Three-phase IMPEC system; (b) emonPi system.

3.3. Acquisition Method

As showcased in Figure 6, different types of data were collected, reflecting electricity consumption in Moroccan premises during the acquisition campaign of MORED: WP consumption data, IL ground-truth consumption data, IL signature consumption data, and event data.

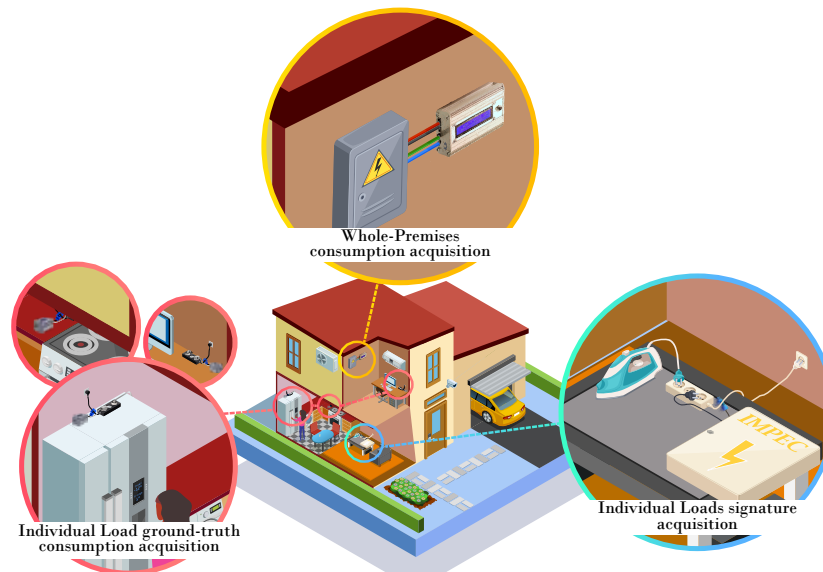


Figure 6. The three different types of the electricity consumption data contained in MORED.

3.3.1. WP Electricity Consumption Data

In order to acquire WP electricity consumption, emonPi systems were installed at the mains of Moroccan households, with one SCT hooked at one phase, for an extended duration of time (see Section 4 for more details). Figure 7 presents the WP electricity consumption acquired from household 2.

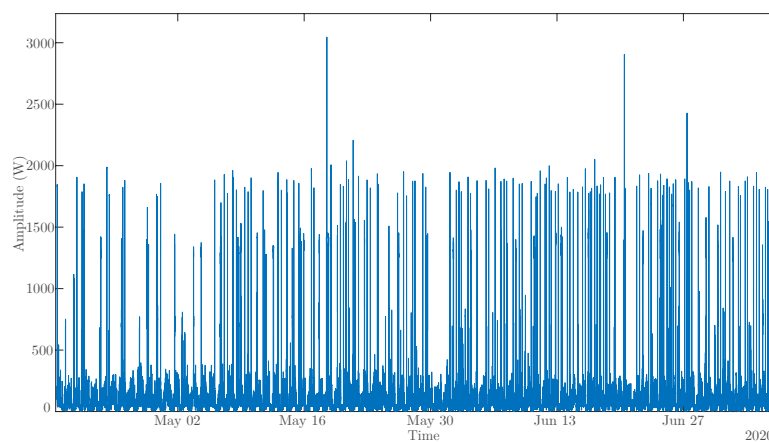


Figure 7. WP electricity consumption of household 2.

3.3.2. IL Electricity Consumption Data

- *Preliminary Notions*

Differentiation between electrical loads is dependent mostly on their respective signatures, which are in turn mostly defined by loads consumption type [47]. In order to acquire IL electricity consumption, the IMPEC system was configured to operate following different processes depending on the load's consumption type. In fact, loads can be classified, based on their consumption patterns, into three main categories: on/off, finite-state, and continuously variable devices (see Figure 8) [48].

1. On/off devices: These are loads with two different states where each one corresponds to a constant energy consumption level (see Figure 8a). This type covers loads including light bulbs, toaster, etc.
2. Finite-state machine (FSM) devices: Also called multi-state devices, these are loads with finite numbers of switching states (see Figure 8b). This type of load covers washing machines, fruit-mixers, etc.

- Continuously variable devices (CVD): These are loads with no apparent finite numbers of switching states (see Figure 8c). This type covers loads such as computers, refrigerators, etc.

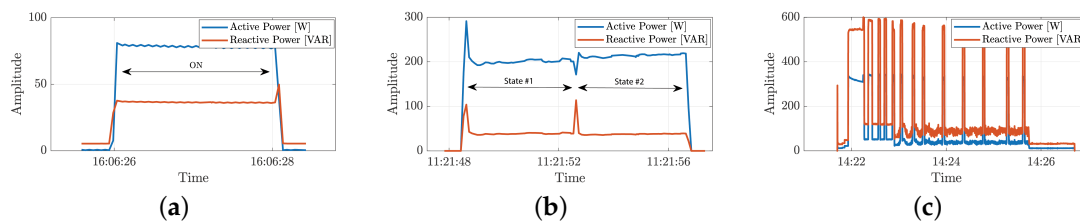


Figure 8. Electricity consumption of three types of devices. (a) On/Off device: hair curler; (b) FSM device: blender; (c) CVD: 3D printer.

• *IL Signature Acquisition Method*

The process to capture signatures is illustrated in the diagram shown in Figure 9. We first proceed by visually analyzing the electrical consumption of one load to define how to proceed with its signature acquisition. Afterwards, the acquisition is performed following a process of three main steps: pre-activation, signature acquisition, and post-activation. The pre- (post-)activation is a period spanning approximately two seconds that precedes (follows) the signature acquisition to enable a clear capture of the turn-on (-off) transients, essential to distinguishing between loads with similar consumption [47,49].

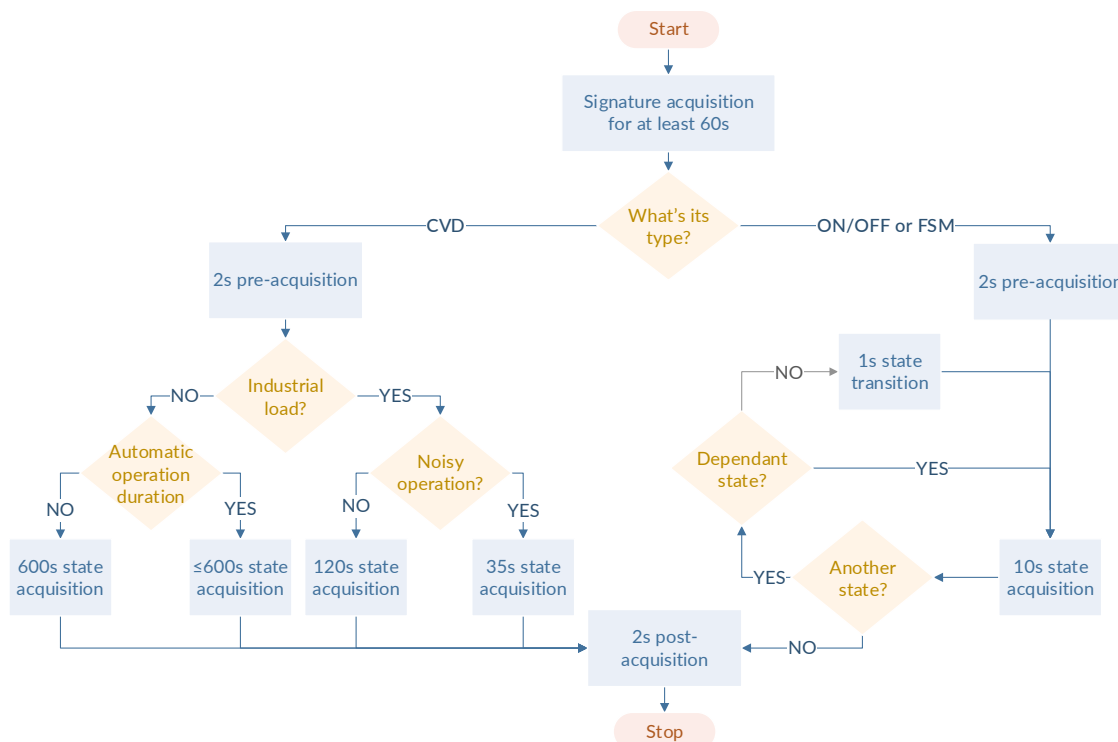


Figure 9. Flowchart underlying the process of IL signature acquisition.

The signature acquisition of a load is performed in two ways:

- In the case of an on/off or FSM load, each state is recorded for 10 s. The transition between states may be direct, or occurring during a one second interval in case of loads with buttons that can halt the execution of states upon their release.
- In the case of a CVD, longer durations of acquisitions are helpful to capturing realistic images of its usual consumption, yet some loads might have some characteristics that can make acquisitions

hard to conduct for long periods; these are mainly devices having automatic or variable operation times that depend on their initial physical states (e.g., initial temperature of an electric kettle), or equipment for industrial applications which experience high noise emissions or is difficult to operate.

As demonstrated previously in the literature [38,50], several aspects can define transients: type of the load, initial physical state of the load, task to be performed, and the time instant corresponding to the turning-on or turning-off of the device with respect to the voltage waveform time-cycle (i.e., the precise value of the voltage phase when the appliance was turned on or off). Therefore, we can observe in the measurements of an appliance different variations of the transient depending on those aspects. Thus, we recorded 10 measurement instances for each load in order to provide a richer representation of the load's consumption.

- *IL Ground-truth Acquisition Method*

Ground-truth data acquisition is achieved in a household by connecting to power extension cords one load and one emonPi. These extension cords would be connected in plugs that are solely used for supplying power to their corresponding devices (i.e., load and emonPi) for the duration of the acquisition. Synchronization between all emonPis available in a household is guaranteed once they are connected to the Internet by updating their internal clock with the country's clock. Loads in a household are typically chosen to be monitored based on the answer of two questions: whether it is possible to access the load's cables and whether its consumption is higher than the monitoring system's consumption. Therefore, unmonitored loads generally include routers, phone chargers, and built-in appliances that can be usually found in the kitchen. Figure 10 showcases the WP electricity consumption of household 5 (see Table 3). These data were annotated using events that were identified from processing the corresponding IL ground-truth consumption data. The portions of the WP consumption that are not annotated in this figure are due to the limited number of monitored loads in the premises.

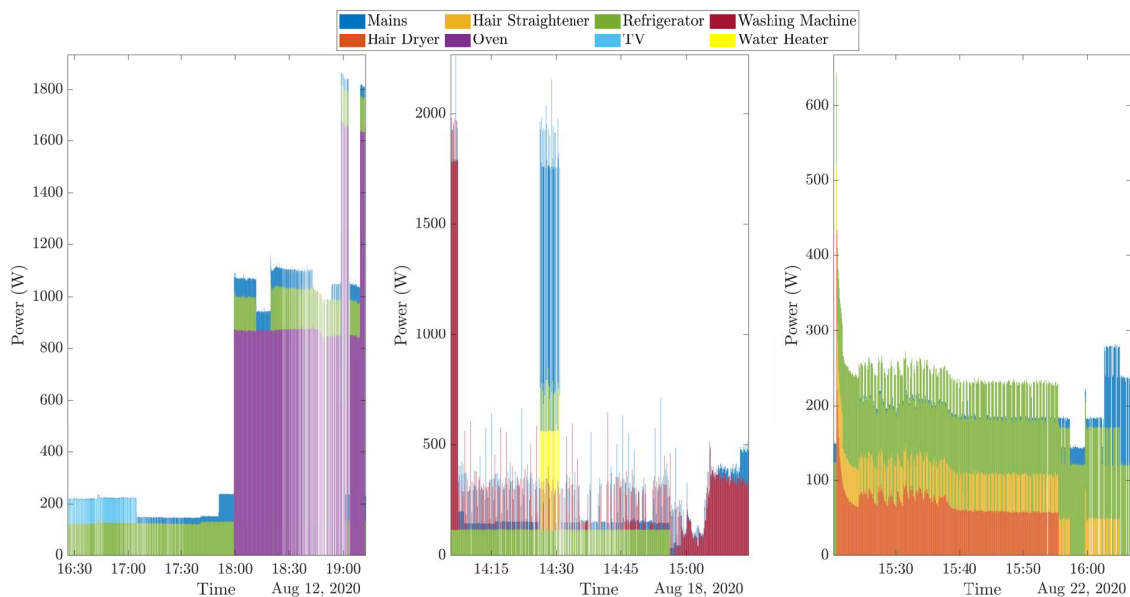


Figure 10. WP electricity consumption of household 5 annotated using IL ground-truth data.

Table 3. Summary of households monitored for WP and IL ground-truth consumption data in MORED. Aff/Apt: affluent neighborhood apartment; Aff/SDH: affluent neighborhood semi-detached house; Dis/Apt: disadvantaged neighborhood apartment; R: rented; B: bought.

Properties		Premises					
		1	2	3	4	5	6
Premises	Aff/Apt	.	.	.	✓	✓	✓
	Aff/SDH	.	✓
	Dis/Apt	✓	.	✓	.	.	.
	City	Sale	Sale	Sale	Tetúan	Rabat	Tetúan
	Number of occupants	2	7	5	3	2	4
	Number of elderly (>65)	0	0	0	0	0	0
	Number of adults (<65 and >18)	2	4	4	3	2	3
	Number of rooms	5	5	5	5	5	6
	Number of floors	1	3	1	1	1	1
	Area (m ²)	50	100	70	103	100	103
	Ownership	R	B	R	B	R	B
WP measurements	Sampling rate (S/s)	1/5	1/5	1/5	1/5	1/5	1/5
	Total acquisition duration (days)	24	14	80	60	44	40
	Mean daily consumption (kW)	110	115	220	281	280	295
IL ground-truth measurements	Sampling rate (S/s)	1/5	1/5	1/5	1/5	1/5	1/5
	Total number of monitored Loads	8	3	3	4	7	3

3.3.3. Event Data

In order to complement the WP data, events reflecting state changes of monitored loads were added to MORED. For IL signatures, state events were identified directly from the load's acquisition process. For IL ground-truth measurements, on-off events were identified manually post-acquisition using a thresholding technique. A load is considered turned ON only when its corresponding active power, P , exceeds a specific threshold. These events are exported in a CSV file containing their corresponding timestamps and the corresponding equipment name. Figure 11 presents the concurrence plot of processed events identified from IL ground-truth consumption of 13 distinct loads acquired from six monitored premises. In this plot, multiple loads can be seen to operate simultaneously (e.g., refrigerator and freezer, monitor and laptop).

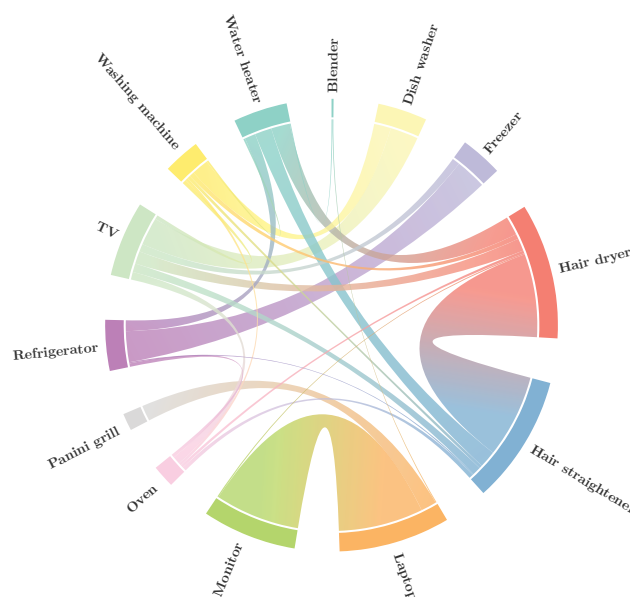


Figure 11. Concurrence plot using events identified from the IL ground-truth consumption data from six households.

4. Dataset Summary

MORED comprises three kinds of electricity consumption data: labeled WP and ground-truth IL consumption, WP consumption, and labeled IL signature consumption. This dataset, available for download at <https://moredataset.github.io/MORED>, is intended to provide more data generally and help alleviate the scarcity of EB electricity consumption datasets in the field specifically. Summaries of targeted premises and loads for WP and IL ground-truth data, WP data, and IL signature data are presented in Tables 3–5, respectively.

The main features of the MORED dataset are:

- WP electricity consumption: WP consumption data were acquired from a total of 13 households of different socio-economic statuses for extended periods of time. They comprise V and P extracted at $1/5[S/s]$ for affluent neighborhood premises and at $1/10[S/s]$ for disadvantaged neighborhood premises.
- Labeled IL ground-truth electricity consumption: IL ground-truth data were acquired for a combined duration of over 300 days from six residential premises comprising V and P extracted at $1/5[S/s]$. Events corresponding to IL ground-truth data only describe on/off operational states of the monitored loads and are all reported in a CSV file.
- Fully labeled IL signature electricity consumption: IL signature data were acquired from 37 classes of loads (a total of 52 loads) comprising i and v sampled at $50k[S/s]$ (which is a higher rate than those of most datasets), and I , V , P , and Q extracted at $4[S/s]$. Events corresponding to IL signature data describe all operational states of the monitored loads and are reported in a CSV file, following the process described in the previous section.
- Metadata: These offer detailed information for all three sets of data (an appliance type, prior knowledge about the typical time of its use per day, its room-location in the premises, etc.) following the NILM Metadata schema [18]. They are reported in YAML text files.
- Measurements data are all provided in CSV files for easy and free access. Yet, IL signature data are also provided in TDMS files since they were originally logged in this format.

Figure 12 illustrates the organization of the dataset directory. Data are divided into three main subdirectories, each corresponding to a specific type of data (i.e., WP and IL ground-truth (WPILGT), WP, and IL signature (ILS)).

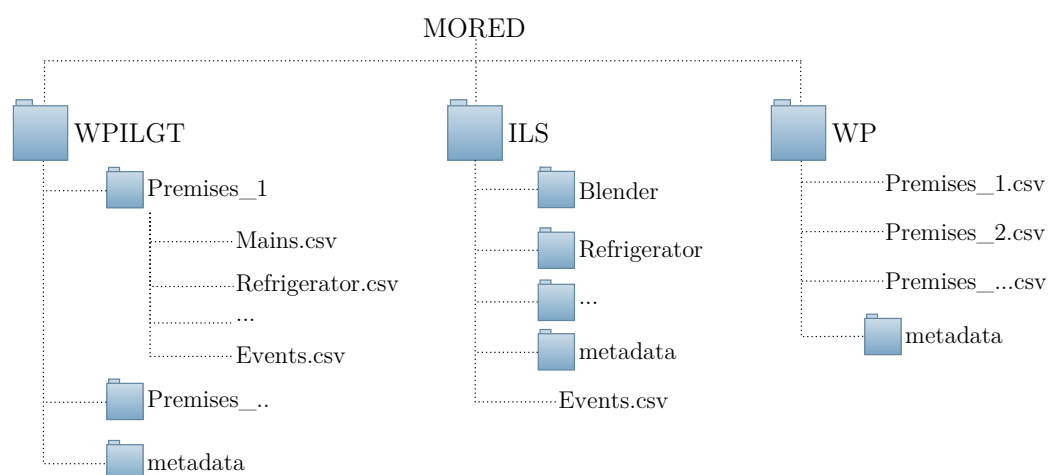


Figure 12. Tree describing the content of the MORED directory.

Table 4. Summary of households monitored for WP electricity consumption data only in MORED. Aff/Apt: affluent neighborhood apartment; Aff/SDH: affluent neighborhood semi-detached house; Dis/Apt: disadvantaged neighborhood apartment; R: rented; B: bought.

Properties	Premises						
	1	2	3	4	5	6	7
Aff/Apt	✓
Aff/SDH	✓	.
Dis/Apt	✓	✓	✓	✓	✓	.	.
City of residence	Sale	Sale	Sale	Sale	Sale	Sale	Tetúan
Number of occupants	2	3	8	4	2	2	3
Number of elderly (>65)	0	0	0	0	2	0	0
Number of adults (<65 and >18)	2	3	5	4	0	2	3
Number of rooms	5	5	5	5	5	5	3
Number of floors	1	1	1	1	1	2	1
Area (m ²)	50	50	50	50	50	100	103
Ownership	R	R	R	R	R	B	B
Sampling rate (S/s)	1/10	1/10	1/10	1/10	1/10	1/5	1/5
Total duration (days)	45	80	36	33	28	82	15
Mean daily consumption (kW)	110	115	220	281	28	187	227

Table 5. Summary of monitored loads for IL signature data in MORED.

Load	Residential				Lab.	#Brand	Consumption Type	Power Supply	#Takes
	1	2	3	4					
Blender	✓	✓	✓	✓	.	4	FSM	1UI	40
Hair dryer	✓	✓	.	✓	.	3	FSM	1UI	30
Laptop	✓	2	CVD	1UI	21
Kettle	.	.	.	✓	.	1	ON/OFF	1UI	10
Hand mixer	.	.	.	✓	.	1	FSM	1UI	10
Stand mixer	.	.	.	✓	.	1	FSM	1UI	10
Toaster	.	.	.	✓	.	1	ON/OFF	1UI	10
Monitor	✓	1	CVD	1UI	10
Panini grill	✓	✓	.	.	.	2	ON/OFF	1UI	20
Refrigerator	✓	✓	✓	.	.	3	CVD	1UI	31
Space heater	✓	1	ON/OFF	1UI	10
Washing machine	✓	.	✓	.	.	2	CVD	1UI	21
WIFI fiber adapter	✓	1	ON/OFF	1UI	12
WIFI modem	✓	1	ON/OFF	1UI	10
Curling brush	.	✓	.	.	.	1	ON/OFF	1UI	10
Hair curler	.	✓	.	.	.	1	ON/OFF	1UI	10
Hair straightener	.	✓	.	.	.	1	ON/OFF	1UI	10
Ironing machine	✓	✓	.	.	.	2	ON/OFF	1UI	20
Oven	.	✓	✓	.	.	2	FSM	1UI	20
Vacuum cleaner	.	✓	.	.	.	1	ON/OFF	1UI	10
Water heater	.	✓	✓	.	.	2	ON/OFF	1UI	20
PS3	.	.	✓	.	.	1	CVD	1UI	10
TV	.	.	✓	.	.	1	CVD	1UI	10
3D printer	✓	1	CVD	1UI	10
Band saw	✓	1	CVD	1UI	10
Desktop computers	✓	1	CVD	1UI	21
Drill press	✓	1	CVD	1UI	10
Drilling machine	✓	1	CVD	1UI	10
Fan heater	✓	1	CVD	1UI	10
Heat gun	✓	1	CVD	1UI	10
Immersion heater	✓	1	CVD	1UI	10
Impact drill	✓	2	CVD	1UI	21
Jigsaw	✓	1	CVD	1UI	10
Metal cutting saw	✓	1	CVD	1UI	10
Paint curing lamp	✓	1	CVD	1UI	10
Sliding miter saw	✓	1	CVD	1UI	10
Combined grinder & polisher	✓	1	CVD	3UI	10

5. Conclusions and Future Work

In the first part of this article, we presented a summary of the open electricity consumption datasets, in order to provide an exhaustive overview of the current state of the art. In the second part, we presented MORED: the first public dataset of Moroccan buildings' electricity consumption. Unlike its predecessors, MORED contains three types of electricity consumption data: labeled WP and IL ground-truth consumption data acquired at 1/5[S/s] rates, fully labeled IL signature acquired at 50k[S/s] and 4[S/s], and WP consumption data acquired at 1/5[S/s] or 1/10[S/s] rates. The aim for building such a dataset is to make a contribution to the fields of energy disaggregation and smart energy management systems. Nevertheless, MORED is considered a work in progress as the authors strive to continuously add new measurements from residential and industrial buildings following the same approaches described in this paper.

Author Contributions: M.A.A.: methodology, software, writing—original draft, data curation. D.B.L.: methodology, supervision, writing—review and editing. C.E.: supervision, writing—review and editing, data curation. M.G.: funding acquisition, methodology, project administration, supervision, writing—review and editing. A.K.: supervision, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the United States Agency for International Development (USAID) (grant number 2000007744) within the PVBUILD project; and by IRESEN, Morocco within the MORESOLAR project.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Lin, Y.H.; Hu, Y.C. Residential Consumer-Centric Demand-Side Management Based on Energy Disaggregation-Piloting Constrained Swarm Intelligence: Towards Edge Computing. *Sensors* **2018**, *18*, 1365. [[CrossRef](#)]
2. Huang, Z.; Zhu, T.; Gu, Y.; Li, Y. Shepherd: Sharing Energy for Privacy Preserving in Hybrid AC-DC Microgrids. In Proceedings of the Seventh International Conference on Future Energy Systems, Association for Computing Machinery, e-Energy '16, New York, NY, USA, 21–24 June 2016; pp. 1–10. [[CrossRef](#)]
3. Shoji, T.; Hirohashi, W.; Fujimoto, Y.; Amano, Y.; Ichi Tanabe, S.; Hayashi, Y. Personalized Energy Management Systems for Home Appliances Based on Bayesian Networks. *J. Int. Counc. Electr. Eng.* **2015**, *5*, 64–69. [[CrossRef](#)]
4. Carrie Armel, K.; Gupta, A.; Shrimali, G.; Albert, A. Is disaggregation the holy grail of energy efficiency? The case of electricity; Special Section: Transition Pathways to a Low Carbon Economy. *Energy Policy* **2013**, *52*, 213–234. [[CrossRef](#)]
5. Figueiredo, M. Contributions to Electrical Energy Disaggregation in a Smart Home. Ph.D. Thesis, Faculty of Sciences and Technology of University of Coimbra, Coimbra, Portugal, 2014.
6. Ahajjam, M.A.; Licea, D.B.; Ghogho, M.; Kobbane, A. *Electric Power Quality Disturbances Classification based on Temporal-Spectral Images and Deep Convolutional Neural Networks*; 2020 International Wireless Communications and Mobile Computing (IWCMC); IEEE: Piscataway, NJ, USA, 2020; pp. 1701–1706. [[CrossRef](#)]
7. Kolter, J.Z.; Johnson, M.J. REDD: A Public Data Set for Energy Disaggregation Research. In Proceedings of the SustKDD workshop on Data Mining Applications in Sustainability, San Diego, CA, USA, 21 August 2011; pp. 59–62.
8. Liang, J.; Ng, S.K.; Kendall, G.; Cheng, J.W. Load signature study—Part I: Basic concept, structure, and methodology. *IEEE Trans. Power Deliv.* **2009**, *25*, 551–560. [[CrossRef](#)]
9. Gulati, M.; Ram, S.S.; Singh, A. An in Depth Study into Using EMI Signatures for Appliance Identification. In Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings; Association for Computing Machinery, BuildSys '14, New York, NY, USA, 4–6 November 2014; pp. 70–79. [[CrossRef](#)]
10. Reinhardt, A.; Baumann, P.; Burgstahler, D.; Hollick, M.; Chonov, H.; Werner, M.; Steinmetz, R. On the accuracy of appliance identification based on distributed load metering data. In *2012 Sustainable Internet and ICT for Sustainability (SustainIT)*; IEEE: Piscataway, NJ, USA, 2012; pp. 1–9.

11. Anderson, K.; Ocneanu, A.; Benitez, D.; Carlson, D.; Rowe, A.; Berges, M. BLUED: A Fully Labeled Public Dataset for Event-Based Non-Intrusive Load Monitoring Research. In Proceedings of the 2nd KDD Workshop on Data Mining Applications in Sustainability (SustKDD), Beijing, China, 12 August 2012; pp. 1–5.
12. Parson, O.; Fisher, G.; Hersey, A.; Batra, N.; Kelly, J.; Singh, A.; Knottenbelt, W.; Rogers, A. Dataport and NILMTK: A building data set designed for non-intrusive load monitoring. In Proceedings of the 2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Orlando, FL, USA, 14–16 December 2015; pp. 210–214. [[CrossRef](#)]
13. Kriechbaumer, T.; Jacobsen, H.A. BLOND, a building-level office environment dataset of typical electrical appliances. *Sci. Data* **2018**, *5*, 180048. [[CrossRef](#)] [[PubMed](#)]
14. Pereira, L.; Nunes, N. Performance evaluation in non-intrusive load monitoring: Datasets, metrics, and tools—A review. *WIREs Data Min. Knowl. Discov.* **2018**, *8*, e1265. [[CrossRef](#)]
15. Mayhorn, E.T.; Sullivan, G.P.; Butner, R.S.; Hao, H.; Baechler, M.C. *Characteristics and Performance of Existing load Disaggregation Technologies*; Technical Report; Pacific Northwest National Lab. (PNNL): Richland, WA, USA, 2015.
16. Batra, N.; Kelly, J.; Parson, O.; Dutta, H.; Knottenbelt, W.; Rogers, A.; Singh, A.; Srivastava, M. NILMTK: An Open Source Toolkit for Non-Intrusive Load Monitoring. In Proceedings of the 5th International Conference on Future Energy Systems, e-Energy '14, Cambridge, UK, 11–13 June 2014; pp. 265–276. [[CrossRef](#)]
17. 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: Demo Abstract. In Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings, BuildSys '14, Memphis, TN, USA, 4–6 November 2014; pp. 182–183. [[CrossRef](#)]
18. Kelly, J.; Knottenbelt, W. Metadata for Energy Disaggregation. In Proceedings of the 2014 IEEE 38th International Computer Software and Applications Conference Workshops, Vasteras, Sweden, 21–25 June 2014; pp. 578–583.
19. Pereira, L. EMD-DF: A Data Model and File Format for Energy Disaggregation Datasets. In Proceedings of the 4th ACM International Conference on Systems for Energy-Efficient Built Environments, Association for Computing Machinery, BuildSys '17, Delft, The Netherlands, 8–9 November 2017; pp. 1–2. [[CrossRef](#)]
20. Zimmermann, J.P.; Evans, M.; Griggs, J.; King, N.; Harding, L.; Roberts, P.; Evans, C. *Household Electricity Survey: A Study of Domestic Electrical Product Usage*; Intertek Testing & Certification Ltd.: Syracuse, NY, USA, 2012.
21. Dua, D.; Graff, C. *Individual Household Electric Power Consumption Data Set-UCI*; Machine Learning Repository; UCI: Irvine, CA, USA, 2012.
22. Barker, S.; Mishra, A.; Irwin, D.; Cecchet, E.; Shenoy, P.; Albrecht, J. Smart*: An open data set and tools for enabling research in sustainable homes. In Proceedings of the 2012 Workshop on Data Mining Applications in Sustainability, Beijing, China, 12 August 2012; Volume 111, p. 108.
23. Makonin, S.; Popowich, F.; Bartram, L.; Gill, B.; Bajić, I.V. AMPds: A public dataset for load disaggregation and eco-feedback research. *IEEE Electr. Power Energy Conf.* **2013**, 1–6. [[CrossRef](#)]
24. Makonin, S.; Ellert, B.; Bajic, I.V.; Popowich, F. Electricity, water, and natural gas consumption of a residential house in Canada from 2012 to 2014. *Sci. Data* **2016**, *3*, 1–12. [[CrossRef](#)] [[PubMed](#)]
25. Batra, N.; Gulati, M.; Singh, A.; Srivastava, M.B. It's Different: Insights into Home Energy Consumption in India. In Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings, BuildSys'13, Roma, Italy, 13–14 November 2013; pp. 1–8. [[CrossRef](#)]
26. Gisler, C.; Ridi, A.; Zufferey, D.; Khaled, O.A.; Hennebert, J. Appliance consumption signature database and recognition test protocols. In Proceedings of the 2013 8th International Workshop on Systems, Signal Processing and their Applications (WoSSPA), Algiers, Algeria, 12–15 May 2013; pp. 336–341. [[CrossRef](#)]
27. Ridi, A.; Gisler, C.; Hennebert, J. ACS-F2-A new database of appliance consumption signatures. In Proceedings of the 2014 6th International Conference of Soft Computing and Pattern Recognition (SoCPaR), Tunis, Tunisia, 11–14 August 2014; pp. 145–150. [[CrossRef](#)]
28. Monacchi, A.; Egarter, D.; Elmenreich, W.; D'Alessandro, S.; Tonello, A.M. GREEND: An energy consumption dataset of households in Italy and Austria. In Proceedings of the 2014 IEEE International Conference on Smart Grid Communications (SmartGridComm), Venice, Italy, 3–6 November 2014; pp. 511–516. [[CrossRef](#)]
29. 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*. [[CrossRef](#)] [[PubMed](#)]

30. 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, BuildSys '14, Memphis, TN, USA, 4–6 November 2014; pp. 80–89. [[CrossRef](#)]
31. Gao, J.; Giri, S.; Kara, E.C.; Bergés, M. PLAID: A Public Dataset of High-Resolution Electrical Appliance Measurements for Load Identification Research: Demo Abstract. In Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings, BuildSys '14, Memphis, TN, USA, 4–6 November 2014; pp. 198–199. [[CrossRef](#)]
32. De Baets, L.; Develder, C.; Dhaene, T.; Deschrijver, D.; Gao, J.; Berges, M. Handling imbalance in an extended PLAID. In Proceedings of the Sustainable Internet and ICT for Sustainability (SustainIT), Funchal, Portugal, 6–7 December 2017; pp. 1–5. [[CrossRef](#)]
33. Inc, E. *Residential Building Stock Assessment: Metering Study*; Technical Report; Northwest Energy Efficiency Alliance: Portland, OR, USA, 2014.
34. Batra, N.; Parson, O.; Berges, M.; Singh, A.; Rogers, A. A comparison of non-intrusive load monitoring methods for commercial and residential buildings. *arXiv* **2014**, arXiv:cs.SY/1408.6595.
35. Murray, D.; Liao, J.; Stankovic, L.; Stankovic, V.; Hauxwell-Baldwin, R.; Wilson, C.; Coleman, M.; Kane, T.; Firth, S. A data management platform for personalised real-time energy feedback. In Proceedings of the 8th International Conference on Energy Efficiency in Domestic Appliances and Lighting, Lucerne, Switzerland, 26–28 August 2015; Volume 4, pp. 1–15. [[CrossRef](#)]
36. Uttama Nambi, A.S.; Reyes Lua, A.; Prasad, V.R. LocED: Location-Aware Energy Disaggregation Framework. In Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments, BuildSys '15, Seoul, Korea, 4–5 November 2015; pp. 45–54. [[CrossRef](#)]
37. Kahl, M.; Haq, A.; Kriechbaumer, T.; Jacobsen, H.A. WHITED-A Worldwide Household and Industry Transient Energy Data Set. In Proceedings of the 2016 3rd International Workshop on Non-intrusive Load Monitoring (NILM), Vancouver, BC, Canada, 14–15 May 2016; pp. 1–5.
38. Meziane, M.N.; Picon, T.; Ravier, P.; Lamarque, G.; Bunetel, J.L.; Raingeaud, Y. A measurement system for creating datasets of on/off-controlled electrical loads. In Proceedings of the IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC), Florence, Italy, 7–10 June 2016; pp. 1–5. [[CrossRef](#)]
39. Saldanha, N.; Beausoleil-Morrison, I. Measured end-use electric load profiles for 12 Canadian houses at high temporal resolution. *Energy Build.* **2012**, *49*, 519–530. [[CrossRef](#)]
40. Johnson, G.; Beausoleil-Morrison, I. Electrical-end-use data from 23 houses sampled each minute for simulating micro-generation systems. *Appl. Therm. Eng.* **2017**, *114*, 1449–1456. [[CrossRef](#)]
41. Makonin, S.; Wang, Z.; Tumpach, C. RAE: The Rainforest Automation Energy Dataset for Smart Grid Meter Data Analysis. *Data* **2018**, *3*, 8. [[CrossRef](#)]
42. Kahl, M.; Krause, V.; Hackenberg, R.; Haq, A.U.; Horn, A.; Jacobsen, H.A.; Kriechbaumer, T.; Petzenhauser, M.; Shamonin, M.; Udalzew, A. Measurement system and dataset for in-depth analysis of appliance energy consumption in industrial environment. *tm-Tech. Mess.* **2019**, *86*, 1–13. [[CrossRef](#)]
43. Ribeiro, M.; Pereira, L.; Quintal, F.; Nunes, N. SustDataED: A Public Dataset for Electric Energy Disaggregation Research. In Proceedings of the ICT for Sustainability, Amsterdam, The Netherlands, 29 August–1 September 2016; pp. 244–245. [[CrossRef](#)]
44. Jazizadeh, F.; Afzalan, M.; Becerik-Gerber, B.; Soibelman, L. EMBED: A Dataset for Energy Monitoring through Building Electricity Disaggregation. In Proceedings of the Ninth International Conference on Future Energy Systems, e-Energy '18, Karlsruhe, Germany, 12–15 June 2018; pp. 230–235. [[CrossRef](#)]
45. Nadeem, A.; Arshad, N. PRECON: Pakistan Residential Electricity Consumption Dataset. In Proceedings of the Tenth ACM International Conference on Future Energy Systems, e-Energy '19, Phoenix, AZ, USA, 25–28 June 2019; pp. 52–57. [[CrossRef](#)]
46. Ahajjam, M.A.; Bonilla Licea, D.; Ghogho, M.; Kobbane, A. IMPEC: An Integrated System for Monitoring and Processing Electricity Consumption in Buildings. *Sensors* **2020**, *20*, 1048. [[CrossRef](#)] [[PubMed](#)]
47. Zoha, A.; Gluhak, A.; Imran, M.; Rajasegarar, S. Non-Intrusive Load Monitoring Approaches for Disaggregated Energy Sensing: A Survey. *Sensors* **2012**, *12*, 16838–16866. [[CrossRef](#)] [[PubMed](#)]
48. Hart, G.W. Nonintrusive appliance load monitoring. *Proc. IEEE* **1992**, *80*, 1870–1891. [[CrossRef](#)]

49. Zeifman, M.; Roth, K. Nonintrusive appliance load monitoring: Review and outlook. *IEEE Trans. Consum. Electron.* **2011**, *57*, 76–84. [[CrossRef](#)]
50. Leeb, S.B.; Shaw, S.R.; Kirtley, J.L. Transient event detection in spectral envelope estimates for nonintrusive load monitoring. *IEEE Trans. Power Deliv.* **1995**, *10*, 1200–1210. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).