A Review of Publicly Available Energy Data Sets

Shubhankar Kapoor, Björn Sturmberg, Marnie Shaw
Battery Storage and Grid Integration Program

Research School of Electrical, Energy and Materials Engineering

The Australian National University
Canberra ACT 2601 Australia

Report submitted for Wattwatchers' My Energy Marketplace (MEM) project, supported by the Australian Renewable Energy Agency (ARENA) https://arena.gov.au/projects/wattwatchersmy-energy-marketplace/)

CRICOS Provider No. 00120C

1 Executive Summary

With increasing levels of renewable energy powering our electricity system, data that tracks energy generation, distribution and consumption has never been more important. For consumers, energy data can inform decisions around home energy management, increasing energy efficiency, participation in aggregation schemes for demand response and Virtual Power Plants (VPPs), and investing in new assets like solar, battery storage or an electric vehicle.

For network operators, energy data can improve the visibility of energy generation and consumption on the low-voltage grid and thereby improve their management of the grid. For policy makers, energy data are needed to make sound policy and regulatory decisions for shaping the energy systems of the future, including future renewable energy generation and storage.

Here, we investigate how energy data have been collected and studied to date. We review 24 of the largest, publicly-available energy data sets, to understand what information has been collected, and how the information has been used. For each dataset, we investigated the date/duration and location of the study, as well as the data variables, including resolution and number of metering points. While 19 of the data sets provide information about power, only five collected measurements of Voltage (V) and Current (I). Out of the 24 data sets, only three collected frequency measurements and two collected phase measurements.

Most of the research carried out on these data sets has made use of power and energy variables with fewer using voltage and current. Almost no research has made use of frequency measurements. Our results suggest that most research on energy data has been carried out to inform customer consumption behaviour, with noticeably less work having been done using energy data to monitor the stability of the network and to make investment decisions for the future grid. We recommend, for future studies and research, to include more information about:

- where metering devices are located in the network.
- frequency and phase values (to allow more work to be done with frequency stabilization or phasor base control).

The report includes links to each of the studied data sets.

2 Acknowledgments

The authors would like to thank Australian Renewable Energy Energy Agency (ARENA) for supporting the project.

Contents

List of Figures

List of Tables

3 Introduction

In the following study, 24 data sets have been considered, as listed in table 1.

3.1 Tabular Representation

Name	Date/ Duration	Size	Data Variables	Resolution	Location	Owner/ Institute
Adaptive Charging Network (ACN) Dataset [45]	May 1,2018 (or earlier) - Present	Over 30,000 charging sessions	Charging current, connection and disconnection time, Energy Delivered (kWh), time zone of charging station, siteID and user inputs	A session of charging the Electric Vehicle (EV)	Two workplace charging sites in California	CALTECH collaboration with PowerFlex Systems
Smart-Grid Smart-City (SGSC)	2010-2014	approx 17000 customers	Energy (kWh) for controlled, load, total household generation, peak event tariffs, PV generation, customer tariff product	30 mins	Australia	Australian Government
NextGen (NG) [71]	2016-Present	upto 5000 Households	Active and reactive power, voltage, frequency, Photo Voltaic (PV) generation battery power, energy capacity (kWh) and maximum charge and discharge rates (kW) of the batteries	5 mins	Australian Capital Territory (ACT)	ACT Government

Name	Date/ Duration	Size	Data Variables	Resolution	Location	Owner/ Institute
Low Carbon London Project (LCLP) [68]	Nov'11 - Feb'14	5,567 households	Energy (kWh), 3 Time of Use tariffs for year 2013 for limited households	30 mins	London household	UK Power Networks
SHED [64]	1 July'10-30 June'13	300 households	General Consumption of house, Controllable load consumption, Gross Generation from PV and generation capacity	30 mins	Household with rooftop PVs in Ausgrid Electricity Network	Ausgrid (NSW GOV)
TU Vienna ADRES dataset (TUVA)	One week in Sep-Dec'09 and May-Oct'10	30 households	Active and Reactive Power, voltage	1s	Upper Austria	TU Vienna
Load data for zone substations in the ActewAGL network (LZSA)	Jul 2004 - Jun 2018	All substations in Canberra	Active Power (MW)	30 mins	Canberra Substation	Commonwealth Scientific and Industrial Research Organisation (CSIRO)

Name	Date/ Duration	Size	Data Variables	Resolution	Location	Owner/ Institute
EPRI Distributed PV Monitoring	6/16/2012 07:00 UTC-6/25/2012 02:00 UTC	17 Modules	AC Power(kW), AC Rating and orientation of PVs for all modules, AC Reactive Power(kvar) for 5 modules, Irradiance(W/m^2) for 14 modules.	1s, 1/15/60 min	PV modules at Poles and PV plants	Electric Power Research Institute (EPRI)
Pecan Street Dataset [72]	2009 - Present	1,000 households	Electricity use and generation, EV charging, rooftop solar generation, energy storage and energy usage to individual circuits along with heating and cooling systems	1s/ 1min/ 15 min	Household data primarily from Texas (Pecan Street). Other states are California, New York and more	Pecan Street Inc
UMass SMART dataset [6]	3 months	3 house holds and 400+ in other one	Usage (aggregate and circuit level) and generation for 3 homes and usage for the rest	1s	USA	Laboratory for Advanced System Software
REDD [39]	3-19 days	6 households	Power Data for all and frequency and current data for 2	3 sec at appliance level, 1s and 15 kHz at aggregate	USA	Massachusetts Institute of Technology (MIT)

Name	Date/ Duration	Size	Data Variables	Resolution	Location	Owner/ Institute
AMPD [49]	April 2012-March 2014	1,051,200 readings/meter for 21 meters in one house	I, V, P, Q available at aggregate and appliance level, weather data and billing data	1 min	Household	Harvard
UK Domestic Appliance-Level Electricity (DALE) [37]	39-234 days and one for 4.3 years	5 Households	Whole home power and appliance level power demand. 3 houses have whole house reactive power. Voltage and current for whole house at 16kHz for 3 houses.	1s (for 3 homes) and 6s	UK	UK DALE
Residential Building Stock Assessment (RBSA) Metering Data	Dec’06 - Nov’10	101 Houses	Active Power (kW), Hours of usage at appliance level	15 mins	Pacific Northwest	NEEA
UCI Individual household electric power consumption dataset	18 Sep’12 - 31 Mar’13	1 household, 2075259 measurements	Active Power(kW), Q, V, I along with 3 Sub metering for energy (Wh)	1 min	Sceaux (7kms of Paris, France)	University of California Irvine (UCI)

Name	Date/ Duration	Size	Data Variables	Resolution	Location	Owner/ Institute
Building-Level Office Environment Dataset of Typical Electrical Appliances (BLOND-50) [40]	213 days	1 Office Building: 53 appliances (17 classes) in a 3-phase power grid	Voltage and current readings for aggregated circuits and matching fully-labeled ground truth data (Individual appliance measurements)	50kS/s (aggregate) and 6.4kS/s (individual appliances)	Germany	TU Munich (TUM)
Building-Level Fully labeled Electricity Disaggregation dataset (BLUED) [23]	8 days	1 house	Current and voltage at circuit/appliance level for 2 phases along with ground truth for events	12kHZ or 0000833 s	USA	Carnegie Mellon University (CMU)
Controlled On/Off Loads Library (COOLL) [63]	June'16	N/A	V and I at appliance level, for 42 controllable appliances of 12 types	100 kHz	PRISME laboratory of the University of Orléans, France	University of Orléans

Name	Date/ Duration	Size	Data Variables	Resolution	Location	Owner/ Institute
Dutch Residential Energy Dataset (DRED) [75]	5th July to 5th Dec 2015	1 Household	Aggregated energy consumption and appliance level consumption, occupancy information, room/outdoor temperature and other weather info	1Hz or 1s for Power/ Energy	Netherlands	TU Delft
Electricity Consumption & Occupancy (ECO) [9]	01.06.12 to 31.01.13	6 Households	Current, voltage, and phase shift for each of the three phases in the household and plug level data for some appliances. Occupancy information.	1Hz or 1s	Switzerland	ETH Zurich
GREEND Dataset [54]	Over a year 2014-02-12 to 2014-10-30	9 Households	Active power measurements of 9 individual appliances and the household aggregate power demand	1Hz or 1s	Austria and Italy	Alpen-Adria-Universität Klagenfurt
IAWE [7]	May-August 2013: 73 days	1 household	P, V, I, f, phase at meter level and for major appliances. Water and ambient monitoring	1/6s	New Delhi, India	Indraprastha Institute of Information Technology

Name	Date/ Duration	Size	Data Variables	Resolution	Location	Owner/ Institute
REFIT: Electrical Load Measurements [57]	Oct 2013 - Jun 2015	20 Households	Power(W) at aggregate and appliance level	8 seconds	UK	University of Strathclyde and others
Tracebase data set (TBS) [65]	2012 and 2013	1883 days	Power for 158 appliance instances, of 43 different appliance types	1s	Darmstadt and some in Sydney	Andreas Reinhardt

Table 1: List of Data Sets

3.2 Access to Data Sets

While most of the data sets mentioned in the above table can be downloaded with the links mentioned in this study in Section 5, access to the following must be requested:

1. Pecan Street Dataset
2. NextGen Dataset
3. TU Vienna ADRES dataset
4. UMass Smart Dataset
5. REDD

NOTE: COOLL and IAWE are available through google drive on their respective web pages.

3.3 Overview of the Data Sets

Energy data sets typically contain measurements for a range of variables, including voltage, current, power, reactive power, frequency and phase information. Fig. 1 provides an overview of the number of data sets that have recorded each of these variables. The number of nodes in each of the data sets is shown in Fig. 2. Here, a node refers to an independent measurement point, e.g. a house or a connection point. The bar graph in Fig. 3 shows the number of data sets with PV generation. Finally, the bar graph in Fig. 4 shows the resolution of the data recorded for each of the data sets.

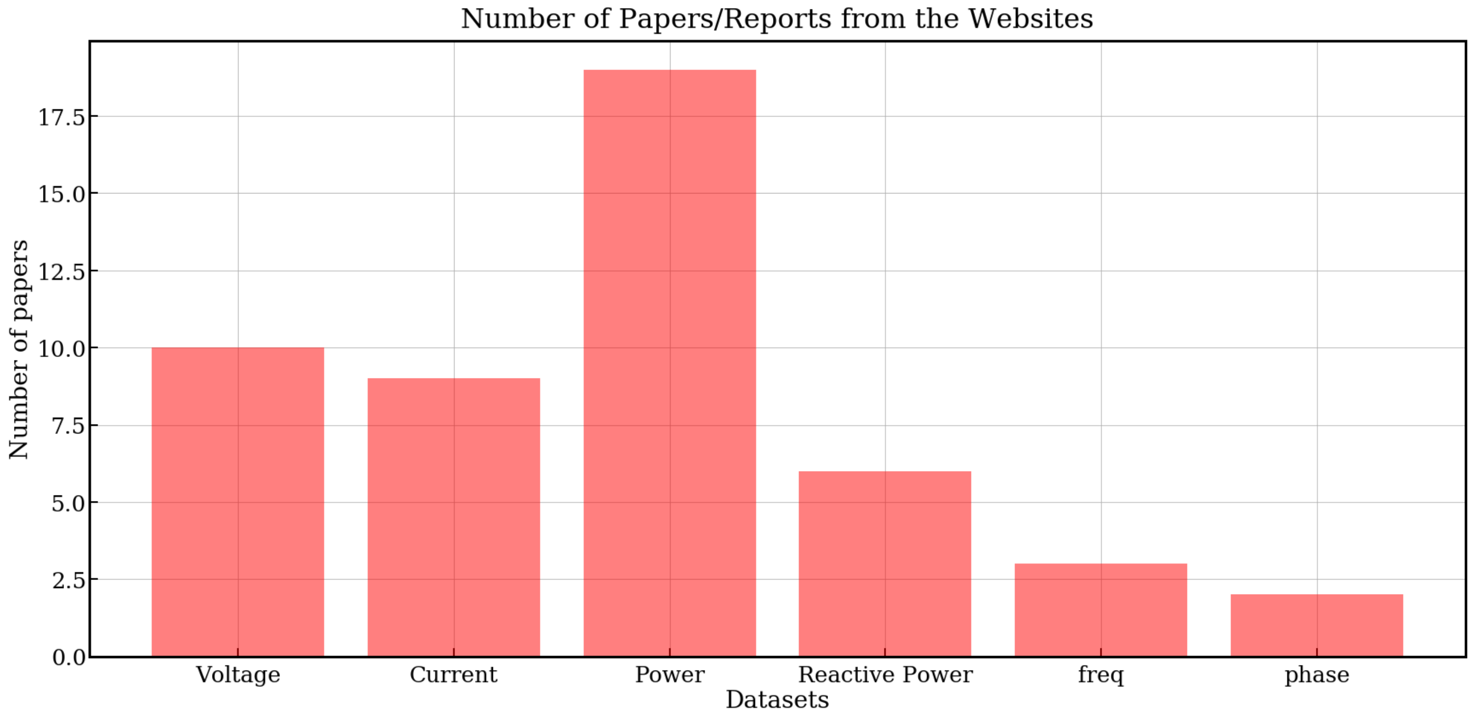

Figure 1: Variables reported in each of the data sets

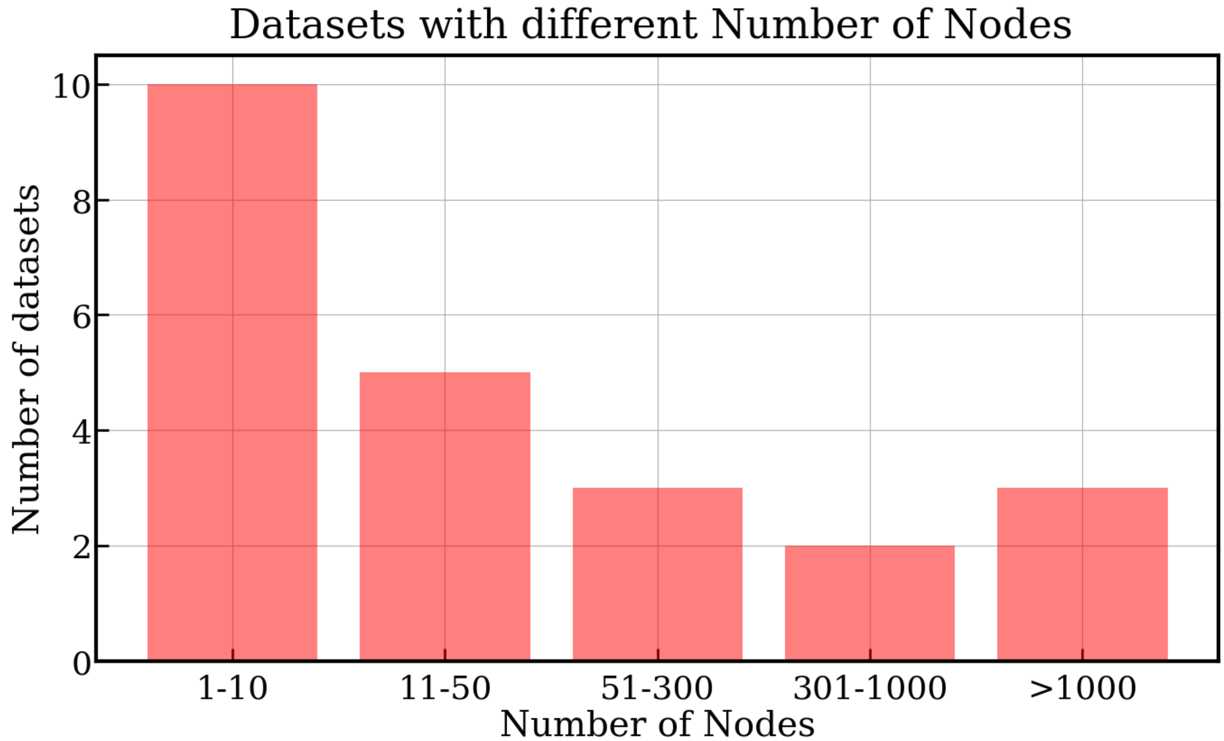

Figure 2: Number of nodes in each of the data sets

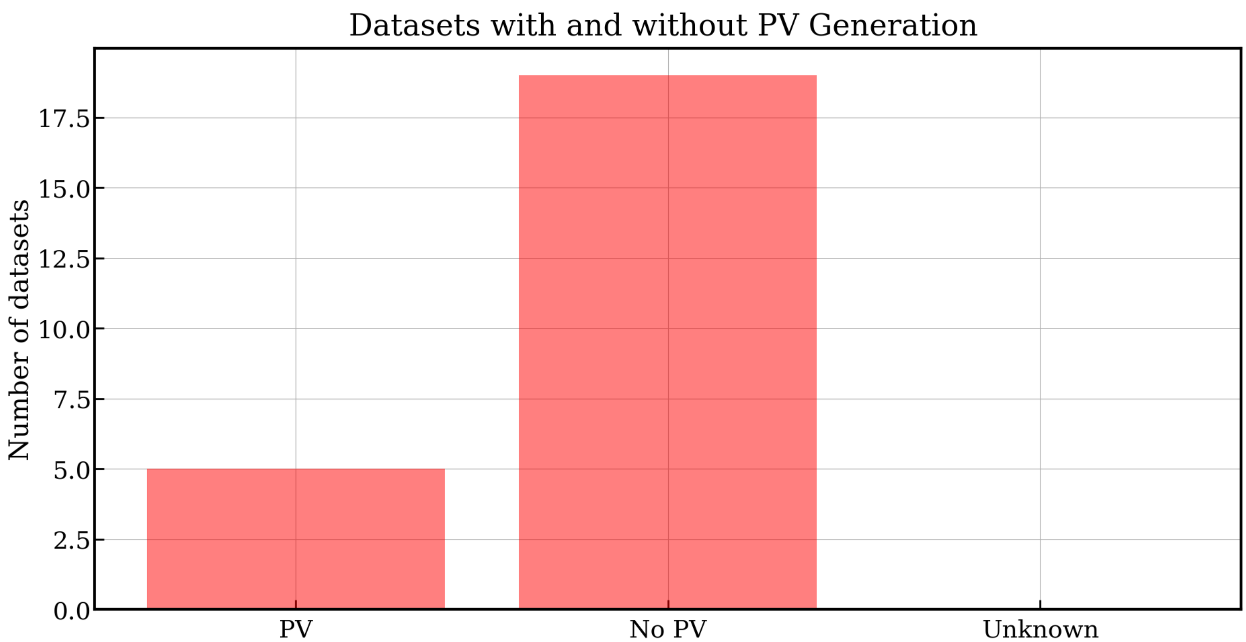

Figure 3: Data Sets with and without PV Generation

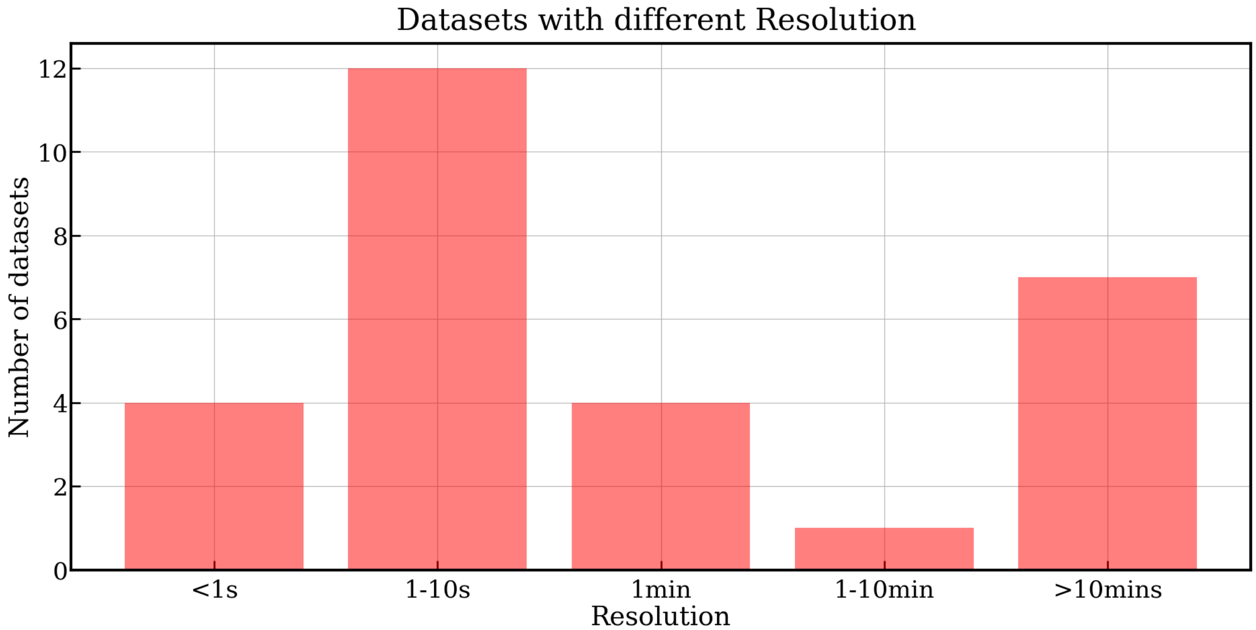

Figure 4: Number of Data Sets with different resolution

3.4 Gaps in the data sets

Figure 1 shows that 19 out of 23 data sets (ACN not included) report power/energy and the four data sets that do not report power and energy, instead report voltage (V) and current (I). Note in Fig. 1, that only two of the data sets report phase values and three report frequency values. Even the most comprehensive and highly-cited dataset, the Pecan street dataset, does not report frequency. Additionally, no data set provided information regarding the location of nodes on the network.

4 Purpose & papers

4.1 Publication(s) based on the data sets

The bar graph in Fig. 5 represents the number of papers/reports that have been published for each of the data sets. The number is taken from the website for each data set. Fig. 6 shows the number of citations for the original paper outlining each dataset (note that UMass SMART Dataset can be found in both bar graphs).

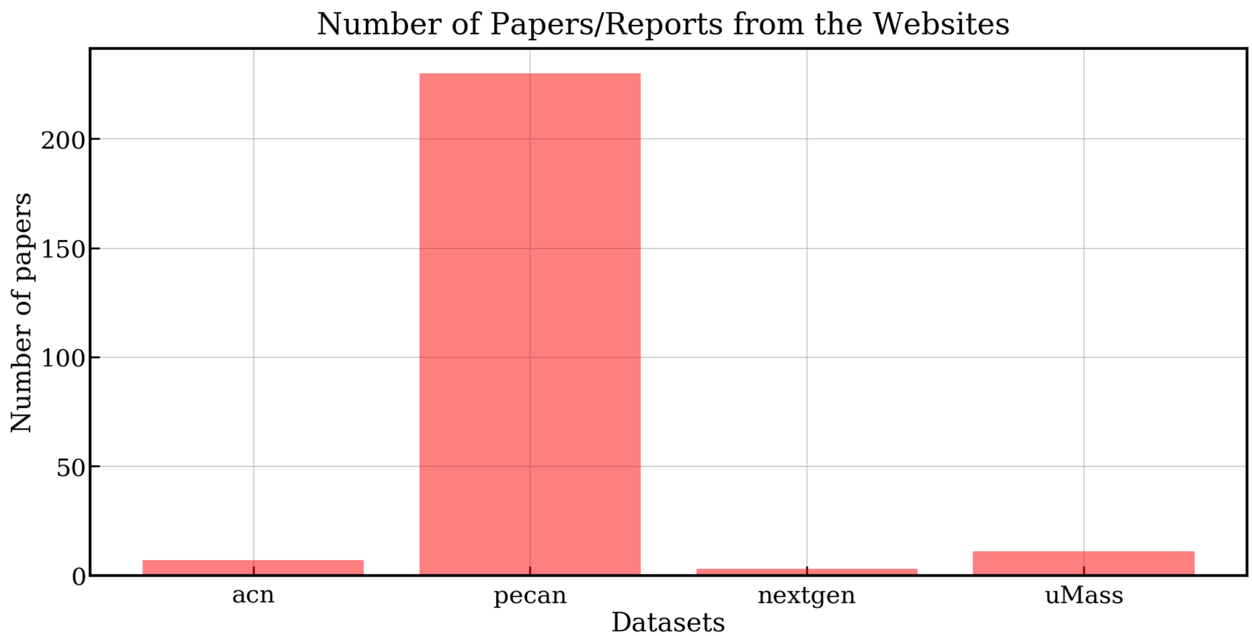

Figure 5: Number of Papers & Reports for the Data Sets from their Websites

For those data sets in this report without a website to collect reports published, we searched for these citations manually. These included:

1. Smart-Grid Smart-City
2. ACTEWAGL Substation Data
3. Low Carbon London Project
4. Solar Home Electricity Data
5. TU Vienna ADRES dataset
6. EPRI Distributed PV Monitoring
7. RBSA Metering data
8. UCI Individual household electric power consumption dataset

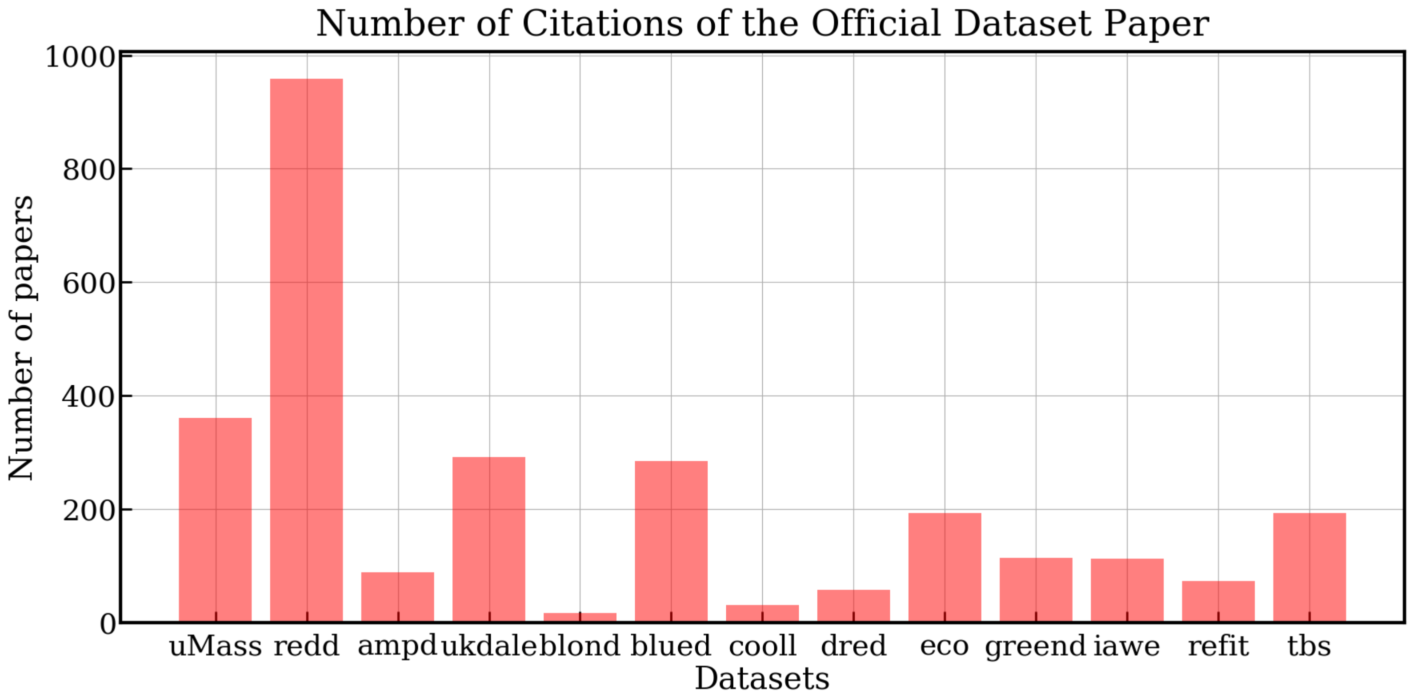

Figure 6: Number of Citations for the original paper outlining each of the data sets

4.2 Key Results produced using the Data Sets

The following table highlights the different streams of work done using different data sets. The categories are as follows:

1. Control, Demand Response, Optimization: includes control of appliances/EVs using different strategies. Optimizing behind the meter storage, PV orientation.
2. Monitoring Capabilities: visualization techniques for data monitoring. Drawing comparisons with simulated data, evaluation of different methods and techniques.
3. Detection Capabilities: detecting various events like charging of EV, switching events, occupancy detection.
4. State Estimation
5. Prediction and Forecasting: predicting usage behaviour, forecasting load profiles, estimating PV generation and optimal size, segmenting and classifying residential customers, appliances.
6. Analysis: analysing user behaviour, EV behaviour, effect of PVs. Analysing different methods for existing techniques.

7. Pricing: defining different pricing under different scenarios.

8. Customer Behaviour: understanding customer behaviour, customer segmentation and the effects of different polices and technologies on customers.

9. Privacy: privacy of data.

10. Policy: examining government promises and the effect of technology on policy making and politics.

11. System Architecture: defining system architecture.

We chose to review the following five datasets for the following reasons:

1. ACN: provides comprehensive information for EV charging.

2. SGSC: one of the largest commercial-scale trial deployment in the world with enormous number of households which aimed at analyzing cost and benefits of the deployment of smart grid technologies, building awareness of economic and environmental benefits, gathering information and data for decision making.

3. PECAN STREET DATASET: provides data for EV charging, rooftop solar generation, energy storage at circuit level with low sampling time interval.

4. REDD: one of the most cited data sets used for Non-Intrusive Appliance Load Monitoring (NIALM)

5. ECO: another NIALM data set along with the occupancy information for the household.

DATASET	**RESEARCH QUESTIONS AND GOALS**	REF.
ACN Dataset		
Control, Demand Response, Optimization	Proposes online scheduling for EV charging using MPC and convex optimization.	[43]
	Scheduling the charging of EVs to minimize peak demand.	[45]
	Optimizing on site solar sizing to meet EV demand.	[45]
	Proposes a decentralized algorithm for scheduling EV charging.	[24]
	Proposed an online linear programming (OLP) for EV charging.	[27]

DATASET	RESEARCH QUESTIONS AND GOALS	REF.
Monitoring Capabilities	Build an ACN simulator to evaluate algorithms and test assumptions.	[44]
System Architecture	Describes the ACN architecture with its software/hardware components.	[43]
	Presents ACN system design including power distribution system, control and communication system with formulation of charging problem.	[42]
Detection Capabilities	-	
State Estimation	N/A	
Prediction and Forecasting	Learning and predicting user behaviours with Guassian mixture models (GMMS)	[45]
Analysis	Analysing number of sessions and usage with different pricing, on different days.	[45]
Pricing	Proposes an algorithm that decides on current charging rates to meet all EVs' energy demands.	[58]
Customer Behaviour	Observes drivers' laxity on different days.	[45]
Privacy	-	
Policy	-	
Smart-Grid Smart-City		
Control, Demand Response, Optimization	Proposes a smart home energy management system with financial benefits to households.	[35]
	Proposes an optimization method for PV orientation and sizing to maximize PV returns.	[19]

DATASET	RESEARCH QUESTIONS AND GOALS	REF.
Monitoring Capabilities	Monitors the impacts on household load due to demand response using Principal Component Analysis and Self-Organizing map for different tariff plans.	
System Architecture	Illustrates different layers of SGSC	[59]
Detection Capabilities	Highlights the importance of fault detection and its economic benefits.	[59]
State Estimation	N/A	
Prediction and Forecasting	Develops an electricity demand model.	[21]
Analysis	Analyzed and discussed the future of smart electricity networks/grid in Australia.	[28]
	Investigates the ongoing research program in different countries for smart grid, draws a comparison, highlights its benefits and issues.	[69]
	Analysed the behaviour of feedback devices on reducing electricity consumption for different demographics.	[20]
Pricing	Highlights the different factors for improving customer pricing outcomes.	[59]
Customer Behaviour	Surveyed the household participants to get an understanding of the user behaviour.	[41]
	Highlights the change in cutomer behaviour to different pricing.	[55]
Privacy	Proposes a lightweight security and privacy-preserving scheme for electricity consumption aggregation.	[1]

DATASET	RESEARCH QUESTIONS AND GOALS	REF.
Policy	Examines the social and environmental promises made by the government on behalf of SGSC program.	[48]
	Looks at the impact of smart electricity network on politics of urbanism	[14]
	Expected benfits of $28 bn till 2034 from smart grid technology.	[59]
PECAN STREET DATASET		
Control, Demand Response, Optimization	Optimizing behind the meter storage for customers to reduce their electricity bills under different utility tariffs.	[78]
	Demand-side control of domestic assets in islanded residential microgrids.	[32]
	Investigates the potential of HVAC systems on demand side control.	[76]
	An Artificial Neural network (ANN) has been proposed for voltage regularization in distribution systems.	[3]
	An incentive based demand response system proposed.	[83]
	More resources for these can be observed here.	[53] [81] [84] [74] [5]
Monitoring Capabilities	Compares the simulated data with real data at circuit level.	[25]
	Evaluates the issues with Advanced Metering Infrastructure (AMI) and makes recommendations to improve AMI.	[11]
System Architecture	-	
Detection Capabilities	Proposes to detect EV charging using smart meter data.	[31]
State Estimation	Bayesian State Estimation using deep learning.	[52]

DATASET	RESEARCH QUESTIONS AND GOALS	REF.
Prediction and Forecasting	A probabilistic model to predict usage appliance usage behaviour.	[82]
	Discovering three fundamental load shape profiles and clustering distributions for different buildings.	[60]
	Segmenting residential consumers based on their appliance level power consumption which can be helpful for demand response.	[13]
	Simulating EV electricity demand.	[30]
	A neural network approach to estimate residential PV size, tilt and azimuth.	[50]
	Other resources can be found here	[33] [79]
Analysis	Analyzing EV behaviour and its effect on household power demands.	[18]
	Analyzing the Vehicle to Grid (V2G) behaviour and its impact on the grid.	[29]
	Analysed the effect of azimuth and lilt of solar PV on the generation and its environmental and economic impacts.	[66]
Pricing	studied the effects of real-time retail pricing on the stability and volatility of power systems.	[67]
Customer Behaviour	Approach for targeting customers for energy efficient programs.	[47]
Privacy	Proposed a privacy preserving mechanism for smart meter data release	[70]
Policy	Highights the importance of simulation tools for policy and investment decisions.	[26]
REDD		

DATASET	RESEARCH QUESTIONS AND GOALS	REF.
Control, Demand Response, Optimization	Proposes a Model Predictive Control (MPC) to control heating/cooling loads, Electric Water Heater (EWH), PV and battery storage in a residential buildings with time-varying pricing.	[80]
Monitoring Capabilities	Designed an interactive visualisation to learn from electricity consumption data.	[16]
	Non-Intrusive Appliance Load Monitoring (NIALM) without having sub-metering appliance readings.	[61]
	More resources can be found here	[8] [36]
System Architecture	Describes the REDD hardware and software systems.	[39]
Detection Capabilities	Proposes a load disaggregation approach and detects power states as on/off for appliances.	[17]
State Estimation	-	
Prediction and Forecasting	Propses multi-label classification algorithms for disaggregating appliances in a power signal.	[73]
	Proposes labelling of appliances. (Also uses AMPD datastet)	[62]
	Bench marking methods for forecasting electricity demand on the household level.	[77]
Analysis	Provides comprehensive overview of NIALM techniques and methods.	[86]
Pricing	Proposed a pricing optimization model proposed for demand response management to maximize retailer's profit	[51]
Customer Behaviour	Proposed an algorithm to learn customer behaviour from appliance level perspective.	[51]
Privacy	Proposes a Battery-based Load Hiding (BLH) algorithm to achieve differential privacy.	[85]
Policy		

DATASET	RESEARCH QUESTIONS AND GOALS	REF.
ECO		
Control, Demand Response, Optimization	-	
Monitoring Capabilities	Proposed an algorithm to derive appliance level data. (Uses other NIALM data sets as well)	[75]
System Architecture	Discusses the architecture for smart home energy system	[46]
Detection Capabilities	Proposed an unsupervised method for occupancy detection. Proposed an algorithm to extract occupancy information.	[10] [38]
State Estimation	-	
Prediction and Forecasting	Proposed a method for modelling appliance power consumption. (Also used TBS for their study) Propses Spatiotemporal Pattern Network (STPN) to predict energy/power consumption for demand side and supply side.	[15] [34]
Analysis	Proposes a data analytical engine. Presented a review of unsupervised NIALM algorithms along with brief overview of NIALM data sets. A survey of NIALM systems and the associated methods with it. (Uses other NIALM data sets as well) Similar resources can be found here.	[4] [12] [22] [56]
Pricing	-	
Customer Behaviour	-	

DATASET	RESEARCH QUESTIONS AND GOALS	REF.
Privacy	Proposed a method for the user to decide the extent of the information being shared at appliance level to achieve certain level of privacy.	[2]
Policy	-	

Table 2: Work done using different Data Sets

4.3 Summary of research carried out on energy data sets

- Algorithms/optimization techniques for controlling, scheduling and optimizing Distributed Energy Resources (DERs) to maximize financial benefits, minimize peak demands and regularize voltage. Different algorithms for dynamic pricing have been proposed to meet net load demand and reduce peak demand.
- Different methods and techniques have been proposed for Non-Intrusive Appliance Load Monitoring (NIALM).
- Different methods for occupancy detection have been proposed using the NIALM data sets.
- Capabilities detecting events like switching events and charging of EVs have been proposed.
- A lot of work has been done forecasting residential usage at aggregate and appliance level; methods for forecasting EV usage have been proposed. Additionally, work has been proposed to label appliances by using their load profile data.
- There have been some reports that have surveyed the household participants to understand their behaviour towards energy conservation programs. Some work highlights the impact(s) of energy conservation programs on policy making and governing bodies.

4.4 Recommendations

There is a direct relationship between the studies summarized in the last section and the data sets that are available. Most work has been done using power and energy measurements, including forecasting, monitoring, control techniques. Some of the work that has been done has been based on voltage (V) and current (I) measurements. However, far less work has been done using frequency and phase values. Therefore we recommend, for future work:

1. it would be ideal to include frequency and phase measurements in the dataset(s). This will assist in working with frequency stabilization or phasor base control.

2. it is recommended to have data sets with network information and behind the meter data measured simultaneously. For example, it would be ideal for future studies to know where the metering devices are located on the network.

5 Links to the datasets

1. **ACN Dataset**: https://ev.caltech.edu/dataset

2. **Smart-Grid Smart-City**: https://data.gov.au/dataset/ds-dga-4e21dea3-9b87-4610-94c7-15a8a77907ef/details

3. **Nextgen Dataset**: https://www.energydata.act.gov.au/pubhome

4. **Low Carbon London Project**: https://data.london.gov.uk/dataset/smartmeter-energy-use-data-in-london-households

5. **Solar Home Electricity Data**: https://www.ausgrid.com.au/Industry/Our-Research/Data-to-share/Solar-home-electricity-data

6. **TU Vienna ADRES dataset**: https://www.ea.tuwien.ac.at/projects/adres_concept/EN/

7. **Load data for zone substations in the ActewAGL network**: https://near.csiro.au/assets/1a650730-c398-4800-8f31-cedec0dd4e0d

8. **EPRI Distributed PV Monitoring**: https://dpv.epri.com/measurement_data.html

9. **Pecan Street Dataset**: https://dataport.pecanstreet.org/ (Available on Request: Academic level data)

10. **UMass SMART dataset**: http://traces.cs.umass.edu/index.php/Smart/Smart

11. **REDD**: http://redd.csail.mit.edu/

12. **AMPD**: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/FIE0S4

13. **UK DALE**: https://jack-kelly.com/data/

14. **RBSA Metering Data**: https://neea.org/data/residential-building-stock-assessment
15. **UCI Individual household electric power consumption dataset**: https://archive.ics.uci.edu/ml/data sets/individual+household+electric+power+consumption
16. **BLOND-50**: https://mediatum.ub.tum.de/1375836
17. **BLUED**: http://portoalegre.andrew.cmu.edu:88/BLUED/
18. **COOLL**: https://coolldataset.github.io/
19. **DRED**: http://www.st.ewi.tudelft.nl/akshay/dred/
20. **ECO**: https://www.vs.inf.ethz.ch/res/show.html?what=eco-data
21. **GREEND**: https://sourceforge.net/projects/greend/
22. **IAWE**: http://iawe.github.io/
23. **REFIT**: https://pureportal.strath.ac.uk/en/data sets/refit-electrical-load-measurements
24. **TBS**: https://github.com/areinhardt/tracebase

References

[1] Asmaa Abdallah and Xuemin Sherman Shen. A lightweight lattice-based homomorphic privacy-preserving data aggregation scheme for smart grid. *IEEE Transactions on Smart Grid*, 9(1):396–405, 2016.

[2] Waqar Asif, Muttukrishnan Rajarajan, and Marios Lestas. Increasing user controllability on device specific privacy in the internet of things. *Computer Communications*, 116:200–211, 2018.

[3] Krishna Sandeep Ayyagari, Reynaldo Gonzalez, Yufang Jin, Miltiadis Alamaniotis, Sara Ahmed, and Nikolaos Gatsis. Artificial neural network-based adaptive voltage regulation in distribution systems using data-driven stochastic optimization. In *2019 IEEE Energy Conversion Congress and Exposition (ECCE)*, pages 5840–5847. IEEE.

[4] Peter Bailis, Edward Gan, Samuel Madden, Deepak Narayanan, Kexin Rong, and Sahaana Suri. Macrobase: Prioritizing attention in fast data. In *Proceedings of the 2017 ACM International Conference on Management of Data*, pages 541–556, 2017.

[5] Hale Bakir, Umair Younas, Derya Betul Unsal Celimli, and Ahmet Afsin Kulaksiz. Konut tipi şebeke bağlantılı fotovoltaik sistemde adaptif nöro bulanık arayüz sistemi tabanlı enerji yönetimi uygulaması.

[6] Sean Barker, Aditya Mishra, David Irwin, Emmanuel Cecchet, Prashant Shenoy, Jeannie Albrecht, et al. Smart*: An open data set and tools for enabling research in sustainable homes. *SustKDD, August*, 111(112):108, 2012.

[7] Nipun Batra, Manoj Gulati, Amarjeet Singh, and Mani B Srivastava. It's different: Insights into home energy consumption in india. In *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, pages 1–8, 2013.

[8] Nipun Batra, Jack Kelly, Oliver Parson, Haimonti Dutta, William Knottenbelt, Alex Rogers, Amarjeet Singh, and Mani Srivastava. Nilmtk: an open source toolkit for non-intrusive load monitoring. In *Proceedings of the 5th international conference on Future energy systems*, pages 265–276, 2014.

[9] Christian Beckel, Wilhelm Kleiminger, Romano Cicchetti, Thorsten Staake, and Silvia Santini. The eco data set and the performance of non-intrusive load monitoring algorithms. In *Proceedings of the 1st ACM conference on embedded systems for energy-efficient buildings*, pages 80–89, 2014.

[10] Vincent Becker and Wilhelm Kleiminger. Exploring zero-training algorithms for occupancy detection based on smart meter measurements. *Computer Science-Research and Development*, 33(1-2):25–36, 2018.

[11] Logan Blakely, Matthew J Reno, and Kavya Ashok. Ami data quality and collection method considerations for improving the accuracy of distribution models. In *IEEE Photovolt. Spec. Conf.(PVSC)*, 2019.

[12] Roberto Bonfigli, Stefano Squartini, Marco Fagiani, and Francesco Piazza. Unsupervised algorithms for non-intrusive load monitoring: An up-to-date overview. In *2015 IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC)*, pages 1175–1180. IEEE, 2015.

[13] Lily Buechler and Zonghe Chua. Appliance-level residential consumer segmentation from smart meter data.

[14] Harriet Bulkeley, Pauline M McGuirk, and Robyn Dowling. Making a smart city for the smart grid? the urban material politics of actualising smart electricity networks. *Environment and Planning A: Economy and Space*, 48(9):1709–1726, 2016.

[15] Nadezda Buneeva and Andreas Reinhardt. Ambal: Realistic load signature generation for load disaggregation performance evaluation. In *2017 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 443–448. IEEE, 2017.

[16] Enrico Costanza, Sarvapali D Ramchurn, and Nicholas R Jennings. Understanding domestic energy consumption through interactive visualisation: a field study. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 216–225, 2012.

[17] Dominik Egarter and Wilfried Elmenreich. Autonomous load disaggregation approach based on active power measurements. In *2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, pages 293–298. IEEE, 2015.

[18] Hunter Estes, Surya Santoso, and Grant Fisher. Analysis of high-resolution electric vehicle charging on time-of-use grid demands. In *2015 IEEE Power & Energy Society General Meeting*, pages 1–5. IEEE, 2015.

[19] Jeremy Every, Li Li, Youguang G Guo, and David G Dorrell. Maximizing investment value of small-scale pv in a smart grid environment. In *2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA)*, pages 385–390. IEEE, 2016.

[20] H Fan, IF MacGill, and AB Sproul. Impact of feedback interventions on residential electricity demand in australia's first large scale smart grid project. In *2015 IEEE Innovative Smart Grid Technologies-Asia (ISGT ASIA)*, pages 1–6. IEEE, 2015.

[21] H Fan, IF MacGill, and AB Sproul. Statistical analysis of driving factors of residential energy demand in the greater sydney region, australia. *Energy and Buildings*, 105:9–25, 2015.

[22] Anthony Faustine, Nerey Henry Mvungi, Shubi Kaijage, and Kisangiri Michael. A survey on non-intrusive load monitoring methodies and techniques for energy disaggregation problem. *arXiv preprint arXiv:1703.00785*, 2017.

[23] Adrian Filip. Blued: A fully labeled public dataset for event-based nonintrusive load monitoring research. In *2nd Workshop on Data Mining Applications in Sustainability (SustKDD)*, page 2012, 2011.

[24] Lingwen Gan, Ufuk Topcu, and Steven H Low. Optimal decentralized protocol for electric vehicle charging. *IEEE Transactions on Power Systems*, 28(2):940–951, 2012.

[25] Brock Glasgo, Chris Hendrickson, and Inês Lima Azevedo. Assessing the value of information in residential building simulation: Comparing simulated and actual building loads at the circuit level. *Applied Energy*, 203:348–363, 2017.

[26] Brock Glasgo, Nyla Khan, and Inês Lima Azevedo. Simulating a residential building stock to support regional efficiency policy. *Applied Energy*, 261:114223, 2020.

[27] Linqi Guo, Karl F Erliksson, and Steven H Low. Optimal online adaptive electric vehicle charging. In *2017 IEEE Power & Energy Society General Meeting*, pages 1–5. IEEE, 2017.

[28] Ahmed MA Haidar, Kashem Muttaqi, and Danny Sutanto. Smart grid and its future perspectives in australia. *Renewable and Sustainable Energy Reviews*, 51:1375–1389, 2015.

[29] Chioke B Harris and Michael E Webber. A temporal assessment of vehicle use patterns and their impact on the provision of vehicle-to-grid services. *Environmental Research Letters*, 7(3):034033, 2012.

[30] Chioke B Harris and Michael E Webber. An empirically-validated methodology to simulate electricity demand for electric vehicle charging. *Applied energy*, 126:172–181, 2014.

[31] Volker Hoffmann, Bjørn Ingeberg Fesche, Karoline Ingebrigtsen, Ingrid Nytun Christie, and Morten Punnerud Engelstad. Automated detection of electric vehicles in hourly smart meter data. 2019.

[32] Gerwin Hoogsteen, Thijs van der Klauw, Albert Molderink, Johann L Hurink, Gerard JM Smit, Xianyong Feng, and Robert E Hebner. Balancing islanded residential microgrids using demand side management. In *2016 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–5. IEEE, 2016.

[33] Yunyou Huang, Jianfeng Zhan, Chunjie Luo, Lei Wang, Nana Wang, Daoyi Zheng, Fanda Fan, and Rui Ren. An electricity consumption model for synthesizing scalable electricity load curves. *Energy*, 169:674–683, 2019.

[34] Zhanhong Jiang, Chao Liu, Adedotun Akintayo, Gregor P Henze, and Soumik Sarkar. Energy prediction using spatiotemporal pattern networks. *Applied Energy*, 206:1022–1039, 2017.

[35] Chanaka Keerthisinghe, Gregor Verbič, and Archie C Chapman. A fast technique for smart home management: Adp with temporal difference learning. *IEEE Transactions on smart grid*, 9(4):3291–3303, 2016.

[36] Jack Kelly and William Knottenbelt. Neural nilm: Deep neural networks applied to energy disaggregation. In *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*, pages 55–64, 2015.

[37] Jack Kelly and William Knottenbelt. The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. *Scientific Data*, 2(150007), 2015.

[38] Wilhelm Kleiminger, Christian Beckel, Thorsten Staake, and Silvia Santini. Occupancy detection from electricity consumption data. In *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, pages 1–8, 2013.

[39] J Zico Kolter and Matthew J Johnson. Redd: A public data set for energy disaggregation research. In *Workshop on data mining applications in sustainability (SIGKDD), San Diego, CA*, volume 25, pages 59–62, 2011.

[40] Thomas Kriechbaumer and Hans-Arno Jacobsen. Blond: Building-level office environment dataset, 2017.

[41] E Langham, J Downes, T Brennan, J Fyfe, SH Mohr, P Rickwood, and S White. Smart grid, smart city, customer research report. 2014.

[42] George Lee, Ted Lee, Zhi Low, Steven H Low, and Christine Ortega. Adaptive charging network for electric vehicles. In *2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pages 891–895. IEEE, 2016.

[43] Zachary J Lee, Daniel Chang, Cheng Jin, George S Lee, Rand Lee, Ted Lee, and Steven H Low. Large-scale adaptive electric vehicle charging. In *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pages 1–7. IEEE, 2018.

[44] Zachary J Lee, Daniel Johansson, and Steven H Low. Acn-sim: An open-source simulator for data-driven electric vehicle charging research. In *2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pages 1–6. IEEE, 2019.

[45] Zachary J. Lee, Tongxin Li, and Steven H. Low. ACN-Data: Analysis and Applications of an Open EV Charging Dataset. In *Proceedings of the Tenth International Conference on Future Energy Systems*, e-Energy '19, June 2019.

[46] Wenjin Jason Li, Xiaoqi Tan, and Danny HK Tsang. Smart home energy management systems based on non-intrusive load monitoring. In *2015 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 885–890. IEEE, 2015.

[47] Huishi Liang, Jin Ma, Rongfu Sun, and Yanling Du. A data-driven approach for targeting residential customers for energy efficiency programs. *IEEE Transactions on Smart Grid*, 2019.

[48] Heather Lovell. The promise of smart grids. *Local Environment*, 24(7):580–594, 2019.

[49] Stephen Makonin, Bradley Ellert, Ivan V Bajić, and Fred Popowich. Electricity, water, and natural gas consumption of a residential house in canada from 2012 to 2014. *Scientific data*, 3:160037, 2016.

[50] Karl Mason, Matthew J Reno, Logan Blakely, Sadegh Vejdan, and Santiago Grijalva. A deep neural network approach for behind-the-meter residential pv size, tilt and azimuth estimation. *Solar Energy*, 196:260–269, 2020.

[51] Fan-Lin Meng and Xiao-Jun Zeng. A profit maximization approach to demand response management with customers behavior learning in smart grid. *IEEE Transactions on Smart Grid*, 7(3):1516–1529, 2015.

[52] Kursat Rasim Mestav, Jaime Luengo-Rozas, and Lang Tong. Bayesian state estimation for unobservable distribution systems via deep learning. *IEEE Transactions on Power Systems*, 34(6):4910–4920, 2019.

[53] Amin Mirakhorli and Bing Dong. An open source smart building energy management platform through volttron. In *Proceedings of building simulation*, 2017.

[54] Andrea Monacchi, Dominik Egarter, Wilfried Elmenreich, Salvatore D'Alessandro, and Andrea M Tonello. Greend: An energy consumption dataset of households in italy and austria. In *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 511–516. IEEE, 2014.

[55] Omid Motlagh, Greg Foliente, and George Grozev. Knowledge-mining the australian smart grid smart city data: A statistical-neural approach to demand-response analysis. In *Planning Support Systems and Smart Cities*, pages 189–207. Springer, 2015.

[56] Jacob A Mueller and Jonathan W Kimball. Accurate energy use estimation for nonintrusive load monitoring in systems of known devices. *IEEE Transactions on Smart Grid*, 9(4):2797–2808, 2016.

[57] David Murray, Lina Stankovic, and Vladimir Stankovic. An electrical load measurements dataset of united kingdom households from a two-year longitudinal study. *Scientific data*, 4(1):1–12, 2017.

[58] Yorie Nakahira, Niangjun Chen, Lijun Chen, and Steven H Low. Smoothed least-laxity-first algorithm for ev charging. In *Proceedings of the Eighth International Conference on Future Energy Systems*, pages 242–251, 2017.

[59] Michelle Norris, Richard Sharp, Ranelle Cliff, Samuel Koci, Hugh Gardner, Melanie Koerner, Ezra Beeman, Geoff Erder, Andrew Harpham, James Allen, Edward Langham, and Jenni Downes. Smart grid smart city: shaping australia's energy future, executive report. 2014.

[60] June Young Park, Xiya Yang, Clayton Miller, Pandarasamy Arjunan, and Zoltan Nagy. Apples or oranges? identification of fundamental load shape profiles for benchmarking buildings using a large and diverse dataset. *Applied energy*, 236:1280–1295, 2019.

[61] Oliver Parson, Siddhartha Ghosh, Mark Weal, and Alex Rogers. Non-intrusive load monitoring using prior models of general appliance types. In *Twenty-Sixth AAAI Conference on Artificial Intelligence*, 2012.

[62] Lucas Pereira and Nuno J Nunes. Semi-automatic labeling for public non-intrusive load monitoring datasets. In *2015 Sustainable Internet and ICT for Sustainability (SustainIT)*, pages 1–4. IEEE, 2015.

[63] Thomas Picon, Mohamed Nait Meziane, Philippe Ravier, Guy Lamarque, Clarisse Novello, Jean-Charles Le Bunetel, and Yves Raingeaud. COOLL: Controlled on/off loads library, a public dataset of high-sampled electrical signals for appliance identification. *arXiv preprint arXiv:1611.05803 [cs.OH]*, 2016.

[64] Elizabeth L Ratnam, Steven R Weller, Christopher M Kellett, and Alan T Murray. Residential load and rooftop pv generation: an australian distribution network dataset. *International Journal of Sustainable Energy*, 36(8):787–806, 2017.

[65] Andreas Reinhardt, Paul Baumann, Daniel Burgstahler, Matthias Hollick, Hristo Chonov, Marc Werner, and Ralf Steinmetz. On the accuracy of appliance identification based on distributed load metering data. In *2012 Sustainable Internet and ICT for Sustainability (SustainIT)*, pages 1–9. IEEE, 2012.

[66] Joshua D Rhodes, Charles R Upshaw, Wesley J Cole, Chris L Holcomb, and Michael E Webber. A multi-objective assessment of the effect of solar pv array orientation and tilt on energy production and system economics. *Solar Energy*, 108:28–40, 2014.

[67] Mardavij Roozbehani, Munther A Dahleh, and Sanjoy K Mitter. Volatility of power grids under real-time pricing. *IEEE Transactions on Power Systems*, 27(4):1926–1940, 2012.

[68] James R Schofield, Richard Carmichael, S Tindemans, Mark Bilton, Matt Woolf, Goran Strbac, et al. Low carbon london project: Data from the dynamic time-of-use electricity pricing trial, 2013. *[data collection]. UK Data Service*, 2015.

[69] GM Shafiullah, Amanullah MT Oo, ABM Shawkat Ali, and Peter Wolfs. Smart grid for a sustainable future. *Smart Grid and Renewable Energy*, 4(1):23–34, 2013.

[70] Mohammadhadi Shateri, Francisco Messina, Pablo Piantanida, and Fabrice Labeau. Deep directed information-based learning for privacy-preserving smart meter data release. In *2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pages 1–7. IEEE, 2019.

[71] Marnie Shaw, Bjorn Sturmberg, Lin Guo, Xinyu Gao, Elizabeth Ratnam, and Lachlan Blackhall. The nextgen energy storage trial in the act, australia. In *Proceedings of the Tenth ACM International Conference on Future Energy Systems*, pages 439–442, 2019.

[72] Pecan Street. Dataport: the world's largest energy data resource. *Pecan Street Inc*, 2015.

[73] Seyed Mostafa Tabatabaei, Scott Dick, and Wilsun Xu. Toward non-intrusive load monitoring via multi-label classification. *IEEE Transactions on Smart Grid*, 8(1):26–40, 2016.

[74] Wen-Jun Tang, Yi-Syuan Wu, and Hong-Tzer Yang. Adaptive segmentation and machine learning based potential dr capacity analysis. In *2018 18th International Conference on Harmonics and Quality of Power (ICHQP)*, pages 1–5. IEEE, 2018.

[75] Akshay SN Uttama Nambi, Antonio Reyes Lua, and Venkatesha R Prasad. Loced: Location-aware energy disaggregation framework. In *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*, pages 45–54, 2015.

[76] Thijs van der Klauw, Gerwin Hoogsteen, Marco ET Gerards, Johann L Hurink, Xianyong Feng, and Robert E Hebner. Assessing the potential of residential hvac systems for demand-side management. In *2016 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–5. IEEE, 2016.

[77] Andreas Veit, Christoph Goebel, Rohit Tidke, Christoph Doblander, and Hans-Arno Jacobsen. Household electricity demand forecasting: benchmarking state-of-the-art methods. In *Proceedings of the 5th international conference on Future energy systems*, pages 233–234, 2014.

[78] Sadegh Vejdan, Adam Kline, Mason Totri, Santiago Grijalva, and Richard Simmons. Behind-the-meter energy storage: Economic assessment and system impacts in georgia. In *2019 North American Power Symposium (NAPS)*, pages 1–6. IEEE, 2019.

[79] Marcus Voß, Asmaa Haja, and Sahin Albayrak. Adjusted feature-aware k-nearest neighbors: Utilizing local permutation-based error for short-term residential building load fore-

casting. In *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pages 1–6. IEEE, 2018.

[80] Evangelos Vrettos, KuanLin Lai, Frauke Oldewurtel, and Göran Andersson. Predictive control of buildings for demand response with dynamic day-ahead and real-time prices. In *2013 European Control Conference (ECC)*, pages 2527–2534. IEEE, 2013.

[81] Zhiqi Wang, Kui Wang, Yang Yu, and Chenye Wu. An algorithmic view on optimal storage sizing. *arXiv preprint arXiv:1911.03684*, 2019.

[82] Marc Wenninger, Jochen Schmidt, and Toni Goeller. Appliance usage prediction for the smart home with an application to energy demand side management. In *Proceedings of the 6th International Conference on Smart Cities and Green ICT Systems*, pages 143–150. SCITEPRESS-Science and Technology Publications, Lda, 2017.

[83] Bainan Xia, Hao Ming, Ki-Yeob Lee, Yuanyuan Li, Yuqi Zhou, Shantanu Bansal, Srinivas Shakkottai, and Le Xie. Energycoupon: A case study on incentive-based demand response in smart grid. In *Proceedings of the Eighth International Conference on Future Energy Systems*, pages 80–90, 2017.

[84] Xiaochen Zhang and Santiago Grijalva. An advanced data driven model for residential plug-in hybrid electric vehicle charging demand. In *IEEE PES 2015 General Meeting*, pages 26–30.

[85] Jing Zhao, Taeho Jung, Yu Wang, and Xiangyang Li. Achieving differential privacy of data disclosure in the smart grid. In *IEEE INFOCOM 2014-IEEE Conference on Computer Communications*, pages 504–512. IEEE, 2014.

[86] Ahmed Zoha, Alexander Gluhak, Muhammad Ali Imran, and Sutharshan Rajasegarar. Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey. *Sensors*, 12(12):16838–16866, 2012.