GREEND: An Energy Consumption Dataset of Households in Italy and Austria

Andrea Monacchi [1], Dominik Egarter [1], Wilfried Elmenreich [1], Salvatore D'Alessandro [2], Andrea M. Tonello [2]

[1] *Institute of Networked and Embedded Systems, Alpen-Adria-Universität Klagenfurt, Austria*
[1] `andrea.monacchi@aau.at, dominik.egarter@aau.at, wilfried.elmenreich@aau.at`

[2] *WiTiKee s.r.l., via Duchi d'Aosta 2, Udine, Italy*
[2] `dalessandro@witikee.com, tonello@witikee.com`

Abstract---**Home energy management systems can be used to monitor and optimize consumption and local production from renewable energy. To assess solutions before their deployment, researchers and designers of those systems demand for energy consumption datasets. In this paper, we present the GREEND dataset, containing detailed power usage information obtained through a measurement campaign in households in Austria and Italy. We provide a description of consumption scenarios and discuss design choices for the sensing infrastructure. Finally, we benchmark the dataset with state-of-the-art techniques in load disaggregation, occupancy detection and appliance usage mining.**

I. INTRODUCTION

Stability issues arise from the progressive installation of renewable energy generation and the diffusion of electric vehicles. Demand response exploits a price signal to reflect fluctuations in the availability of energy in the grid. The price incentivizes the coordination of electrical devices in order to optimize running costs, fostering awareness to increase conservation and efficiency [1, 2] or to automatically schedule the operation of selected devices to off-peak periods [3]. Nevertheless, control strategies might somehow disrupt user's daily life routines. Collecting information of reoccurring activities can increase the effectiveness of control strategies, as they can consider necessities of inhabitants to minimize the discomfort produced. Problems such as load detection and occupancy modeling for HVAC optimization can be solved using well-established data mining and machine learning techniques, which demand suitable datasets.

To enable research on energy and sustainability, publicly available data is necessary to elaborate solutions that work in the real world. In this paper we discuss available energy consumption datasets and present a new publicly available dataset that we name GREEND[1]. The dataset contains power usage information at device level being obtained through a measurement campaign in households in Austria and Italy. We describe the hardware and software setup of the measurement campaign and elaborate the modeling of loads for energy management applications. The paper focuses on the dataset and it is meant to provide metadata describing the involved consumption scenarios. Along with the dataset we are also releasing our code base for the measurement campaign as a

[1]GREEND: GREEND Electrical ENergy Dataset.

Sourceforge project. The consumption dataset is accompanied by metadata describing consumption scenarios, which can be imported in the nilm toolkit for further analysis [4, 5]. Moreover, we benchmark the dataset by presenting experiences with load disaggregation, occupancy detection and appliance usage mining using state-of-the-art techniques.

The remainder of the paper is as follows. To identify the main requirements of an energy consumption dataset, we survey existing datasets in Section II. In Section III, we report the design of the measurement campaign. We describe our measurement infrastructure and provide a characterization of the deployed scenarios, in terms of both residents and monitored electrical devices. In Section IV-A, we apply a state-of-the art load disaggregation technique based on particle filtering to show a possible application of the dataset. Similarly, in Section IV-B and IV-C we apply the dataset respectively to observe occupancy patterns and mine appliance usage patterns. Section V concludes the paper and anticipates future directions.

II. OTHER CONSUMPTION DATASETS

To overcome the previously introduced challenges, researchers and designers of home energy management systems (HEMS) require consumption datasets where solutions can be assessed beforehand. A widely used dataset in the load disaggregation community is the reference dataset (REDD), publically released by MIT in 2011 [10]. Moved by a similar aim, various other datasets have been shared. As shown in Table I, the main classification attributes are sampling frequency and characteristics of the signal being measured, active power (P), reactive power (Q), apparent power (S), energy (E), frequency (f), phase angle (Φ), voltage (V) and current (I). Certain datasets, such as REDD and BLUED [8] monitor a small number of households at a high sampling frequency. This depends on the requirements of load disaggregation, where a higher frequency allows for extraction of more representative features capturing the transient behavior [1]. As noticeable in the table, most of datasets were collected in the USA, under a $120V$ voltage while European countries commonly work under $230V$. Moreover, certain peculiarities such as the weather and climate depend on the location of the measurement campaign. To the best of our knowledge, GREEND is the first 1Hz consumption dataset for Austria and Italy. The type and number of devices, as well as the number of monitored households

TABLE I: Existing datasets for energy consumption in households

Dataset	Location	Duration	#Houses	#Sensors (per house)	Features	Resolution
ACS-F1 [6]	Switzerland	1 hour session (2 sessions)	N/A	100 devices in total (10 types)	I, V, Q, f, Φ	10 secs
AMPds [7]	Greater Vancouver	1 year	1	19	I, V, pf, F, P, Q, S	1 min
BLUED [8]	Pittsburg, PA	8 days	1	Aggregated	I, V, switch events	12 Khz
GREEND	Austria, Italy	1 year (3-6 months completed)	8	9	P	1 Hz
HES	UK	1 month (255 houses) - 1 year (26 houses)	251	13-51	P	2 min
iAWE [9]	India	73 days	1	33 sensors (10 appliance level)	V, I, f, P, S, E, Φ	1 Hz
IHEPCDS[1]	France	4 years	1	3 circuits	I, V, P, Q	1 min
OCTES[2]	Finland, Iceland, Scotland	4-13 months	33	Aggregated	P, Energy price	7 secs
REDD [10]	Boston, MA	3 - 19 days	6	9-24	Aggregate: V, P; Sub-metered: P	15 Khz (aggr.), 3 sec (sub)
Sample dataset[3]	Austin, TX	7 days	10	12	S	1 min
Smart* [11]	Western Massachussets	3 months	1 Sub-metered +2 (Aggregated + Sub-metered)	25 circuits, 29 appliance monitors	P, S (circuits), P (sub-metered)	1 Hz
Tracebase [12]	Germany	N/A	15	158 devices in total (43 types)	P	1-10 sec
UK-DALE [13]	UK	499 days	4	5 (house 3) - 53 (house 1)	Aggregated P, Sub P, switch-status	16 Khz (aggr.), 6 sec (sub.)

[1] http://tinyurl.com/IHEPCDS
[2] http://octes.oamk.fi/final/
[3] http://www.pecanstreet.org/projects/consortium/

significantly constraints the final application of the dataset. A statistical analysis of consumption behavior of residents would require a high number of households, such as in HES and OCTES. Besides, seasonal consumption behaviors cannot be captured by a short-term measurement campaign, for instance lasting days or months. Also, some datasets monitor different households over different time windows, which makes a comparison of the dwellings impossible. Moreover, the context under which appliances are used over the day is an essential factor determining the complexity of the demand. Therefore, collection should take place in real environments rather than in a lab or on a device testbench. For example TraceBase [12] and ACS-F1 [6] provide consumption tracks of appliances collected in households and offices. While they constitute a great source of device signatures, at the current stage they do not provide information of consumption scenarios, which results in the impossibility to analyze device interdependence. GREEND is meant to overcome these limitations as it will be explained in the following section.

III. DATASET FOR ITALY AND AUSTRIA

A. The consumption dataset

The measurement campaign is carried out within the MON-ERGY[2] project, in which we aim at proposing solutions to reduce energy consumption in the Austrian (AT) region of Carinthia and the Italian (IT) region of Friuli-Venezia Giulia. In particular, we identified the following requirements:

- *Features:* For the selection of the features of interest we considered the requirements of load disaggregation applications, as they are generally stricter than user and appliance modeling [14]. Accordingly, we decided to collect active power measurements at 1Hz, as this allows the identification of more than 8 devices through load disaggregation algorithms [1]. Each entry is associated to a UTC Unix timestamp so that measurements can be matched to contextual data such as weather.

- *Device selection:* The selection of devices followed the energy hogs identified in [2]. We favored diversity in the dataset to promote multiple applications. However, we faced issues in accessing certain devices in Austria. For instance, electric boilers are commonly connected to a separated meter to be charged with a reduced energy tariff.
- *Household selection:* The selection of householders was driven by the findings identified in [2]. In particular, we wanted to promote diversity of scenarios, for instance involving different types of dwellings and consumers. A more detailed description of the scenarios is reported in Section III-B.
- *Campaign duration:* The campaign was designed to last one year, in order to observe and be able to model seasonal consumption behavior of inhabitants. The first household is being monitored (#0) since the end of December 2013. Most of other households followed in January 2014, while the last two platforms (house #6 and #7) were deployed in May.

B. Deployment

At the time of writing, we are collecting consumption data in the following scenarios:

- *House #0* a detached house with 2 floors in Spittal an der Drau (AT). The residents are a retired couple, spending most of time at home.
- *House #1* an apartment with 1 floor in Klagenfurt (AT). The residents are a young couple, spending most of daylight time at work during weekdays, mostly being at home in evenings and weekend.
- *House #2* a detached house with 2 floors in Spittal an der Drau (AT). The residents are a mature couple (1 housewife and 1 employed) and an employed adult son (28 years).
- *House #3* a detached house with 2 floors in Klagenfurt (AT). The residents are a mature couple (1 working part-time and 1 full time), living with two young kids.
- *House #4* an apartment with 2 floors in Udine (IT). The residents are a young couple, spending most of daylight

time at work during weekdays, although being at home in evenings and weekend.

- *House #5* a detached house with 2 floors in Colloredo di Prato (IT). The residents are a mature couple (1 housewife and 1 employed) and an employed adult son (30 years).
- *House #6* a terraced house with 3 floors in Udine, (IT). The residents are a mature couple (1 working part-time and 1 full time), living with two young children.
- *House #7* a detached house with 2 floors in Basiliano (IT). The residents are a retired couple, spending most of time at home.

The device configurations for the selected households are shown in Table II.

TABLE II: Device configurations in the monitored households

House	Devices
0	Coffee machine, washing machine, radio, water kettle, fridge w/ freezer, dishwasher, kitchen lamp, TV, vacuum cleaner
1	Fridge, dishwasher, microwave, water kettle, washing machine, radio w/ amplifier, dryer, kitchenware (mixer and fruit juicer), bedside light
2	TV, NAS, washing machine, drier, dishwasher, notebook, kitchenware, coffee machine, bread machine
3	Entrance outlet, Dishwasher, water kettle, fridge w/o freezer, washing machine, hairdrier, computer, coffee machine, TV
4	Total outlets, total lights, kitchen TV, living room TV, fridge w/ freezer, electric oven, computer w/ scanner and printer, washing machine, hood
5	Plasma TV, lamp, toaster, stove, iron, computer w/ scanner and printer, LCD TV, washing machine, fridge w/ freezer
6	Total ground and first floor (including lights and outlets, with whitegoods, air conditioner and TV), total garden and shelter, total third floor.
7	TV w/ decoder, electric oven, dishwasher, hood, fridge w/ freezer, kitchen TV, ADSL modem, freezer, laptop w/ scanner and printer

C. Measurement setting

Our deployment consists of an ARM-based platform (e.g., Raspberry Pi or BeagleBone) connected to an Anker Astro E5 15000mAh external battery[3], as well as a Plugwise Basic kit[4]. The Plugwise kit consists of a Zigbee network of 9 sensing outlets, each collecting active power measurements from the connected load. The developed daemon uses the open source python-plugwise library and it is freely available both as a sourcecode and as a ready-to-use SD image[5]. The daemon is based on a collector script and a manager script. The manager is started at boot up by crontab. Its task is to make sure that a collection script is running and that the running version is the latest available. To this end, the manager periodically checks (default is 5 hours) the presence of new versions on the Monergy servers, and when necessary, it replaces the current version with the newest available. In this way we can push updates to all households without needing to access each and every unit (see Fig.1). The collection takes place in epochs. At each epoch, the system collects power measurements from each node and sleeps for the remaining time. To provide periodicity,

[3]http://www.ianker.com/support-c1-g228.html
[4]http://www.plugwise.com
[5]http://sourceforge.net/projects/monergy

in presence of failures such as mispelled packets, nodes are skipped and retried at the end of the epoch if there is time. In addition, a backoff time is used to manage faults, such as the ones resulting from temporary erroneous states or disconnection of nodes from the network. As for the data storage, the daemon can work under 4 different modalities: i) local storage as a daily comma separated value (CSV) file, ii) remote storage on a mysql server with visualization and quick download (See Fig. 2), iii) a combination of i and ii for double backup, iv) daily csv file uploaded via sftp to the Monergy server. While the latter is preferred, it was necessary to employ the first method in houses #1 and #7 due to missing internet connection.
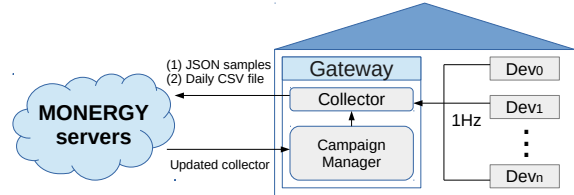

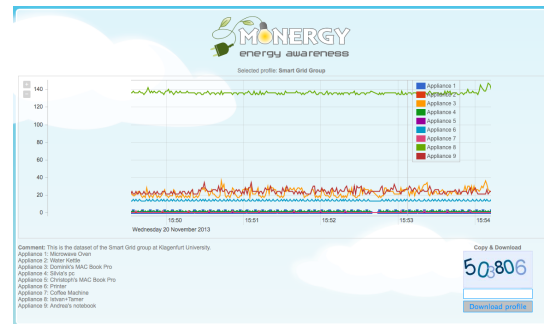Fig. 1: The sensing infrastructure

Fig. 2: The website of the testbed

IV. CASE STUDIES

The introduction of smart metering produces fine-grained energy consumption data allowing for the extraction of more valuable information of energy production and demand compared to having only an energy balance over an extended period of time. Energy management systems (EMS) can be used to better manage local consumption and production from renewable sources. Beside providing consumers with feedback regarding used appliances and their operational costs, this might also support appliance scheduling in order to minimize costs. In this section, we report 3 case studies dealing with consumption information: i) load disaggregation, ii) occupancy detection, and iii) appliance usage modeling.

A. Non-Intrusive appliance load monitoring

Non-intrusive load monitoring (NILM), also known as load disaggregation, is the problem to detect and to classify appliances from the total household power draw, based on specific characteristics of electrical devices. Specifically, NILM tries to solve the disaggregation of aggregated appliance power loads under the influence of noise in measurements and models.

TABLE III: Accuracy of load disaggregator for the aggregated power draws of houses #0 and #1

House ID 0		House ID 1	
Type	ACC	Type	ACC
TV	0.91	hair dryer	0.99
coffee machine	0.99	light	0.97
dishwasher	0.96	dishwasher	0.46
fridge	0.93	fridge	0.90
vacuum cleaner	0.99	water kettle	0.99
water kettle	0.99	washing machine	0.80

NILM was initially presented by Hart [15], although various approaches were later presented. It is generally possible to distinguish between supervised and unsupervised techniques, depending on the necessity of labeled data to train the classifier. Nevertheless, there exists no NILM algorithm able to solve the problem in all its aspects.

Our evaluation is based on the load disaggregator presented in [16], which uses particle filtering (PF) to estimate the appliance state of all used appliances. In detail, appliances are model as a hidden Markov model (HMM), whose states define device states and are associated to the respective power demand. Furthermore, the appliance models are combined into a factorial hidden Markov model (FHMM) modeling the total household power demand. The PF is used to estimate the household power demand according to the given appliance models. Thereafter, a simple decision marker determines the appliance condition and operational state, based on thresholding and the given appliance models. For the evaluation we used the data of house #0 and #1 for 7 consecutive days. As the PF is mainly dependent the number of used particles, we empirically identified 1000 as the appropriate number. The aggregated power draw is composed by 6 different appliances which are listed in Table III. Beside the used appliances also the reached accuracy of the proposed load disaggregator is presented. In this context, the accuracy (ACC) is defined as:

$$ACC = \frac{TP + TN}{N}, \qquad (1)$$

where TP (true positives) is the number of times an appliance is correctly detected as ON, TN (true negatives) is the number of times an appliance is correctly detected as OFF, while N is the number of samples in the observation window. The results in Table III show that the presented dataset can be applied as a reference dataset to load disaggregation problems. The complexity of the problem can be adjusted according to the used appliances and corresponding appliance model (on/off device such as the water kettle or multi-state devices such as the dishwasher) and the number of aggregated power loads.
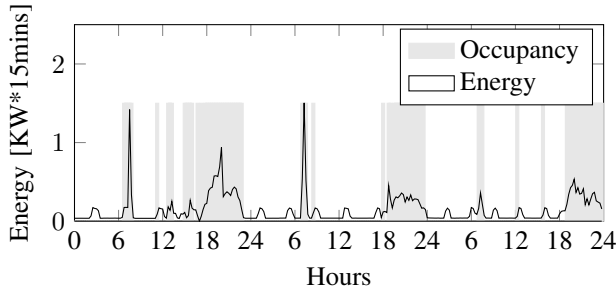
B. Occupancy detection

Occupancy detection is the problem of inferring presence of people in environments. Many different approaches have been considered in the literature: motion detection, doors opening, use of acoustic sensors and/or cameras, use of GPS and localization systems through smart phones. Nevertheless, only a few works propose to use energy consumption data to develop occupancy detection techniques [17]. To this end, the assessment of occupancy detection requires suitable datasets, offering either an aggregated or a disaggregated power draw, along with a description of consumption scenarios. In particular, the presence of user-driven devices is relevant to assume presence in the environment. It is important to specify whether residents can postpone the operation of user-driven devices, as this will condition consumption-based approaches. A simple approach based on the use of energy consumption data to detect occupancy was presented in [17]. Therein, the authors developed an algorithm called non-intrusive occupancy monitoring (NIOM). In particular, NIOM guesses the presence of people at home by comparing the average, the variance and the maximum value of the current power consumption with threshold values. The threshold values are computed during inactivity periods, namely when the activity of residents does not add power consumption to the baseline consumption, e.g., the consumption of devices such as fridge, HVAC systems etc. Accordingly, the thresholds are computed every night when people are supposed to sleep. As both power values and threshold values are computed over time slots of fixed duration, the duration of the slot is an important parameter affecting the estimation. Although occupancy detection through energy consumption observation is a cheap and non-intrusive solution, it is shown being rather inaccurate in determining the occupancy during periods of inactivity.

In order to give an example on the use of the GREEND dataset to study occupancy, we consider the house #4. We chose this scenario because we have measured the total power consumption, which allows for better estimating the baseline during inactivity periods. We consider two months of measurements (February and March 2014) and we apply the NIOM algorithm as described in [17]. The time slot duration is chosen equal to 15 minutes. We consider two cases, the weekdays (WDs) and the weekend days (WEs). This is because the analyzed power consumption greatly differs in these two circumstances as it can be seen in Fig. 3c and Fig. 3e, where the 99th percentile of the energy consumption during WDs and WEs are shown. The baseline, within which the baseline threshold values are computed, was chosen at night between midnight and 6.30 AM. Fig. 3a reports the occupancy obtained applying the NIOM algorithm over three consecutive days. Fig. 3b and Fig. 3d respectively show the probability of occupancy obtained for WDs and WEs. We first notice that, as expected, the results obtained between midnight and 6.30 AM do not indicate any activity.We notice a good agreement between the derived probability of occupancy and the habits of residents. Residents stated that during WDs, they usually wake up at 6.45 AM, leave around 8.00 AM to go to work and are back home around 6.00 PM. During WEs residents tend to spend more time at home during the day and they go out at night coming back home around midnight.

C. Appliance usage modeling

Mining appliance usage patterns concerns the extraction of models describing how appliances are used by residents, given a sequence describing changes on their operational state. Given log data describing status changes of electrical devices, the extraction of models describing the usage of devices can
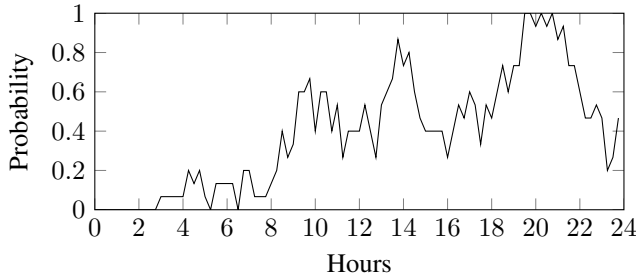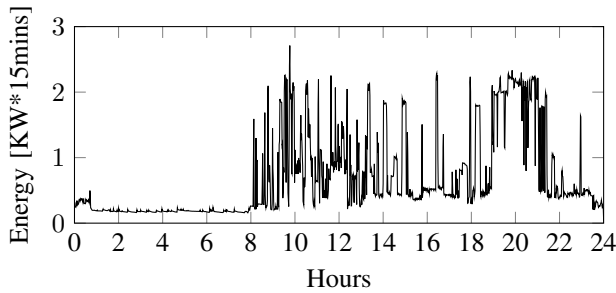
(a) Occupancy over 3 days
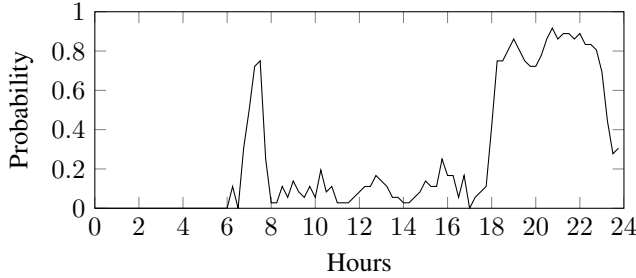
(b) Weekday occupancy probability
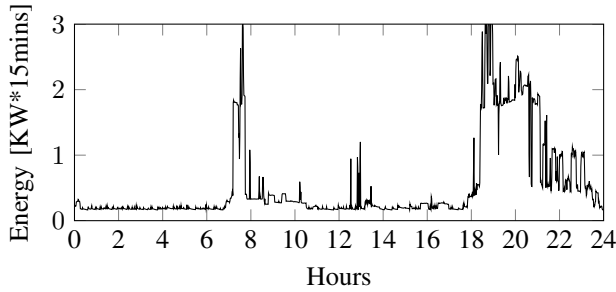
(c) Weekday energy

(d) Weekend occupancy probability

(e) Weekend energy

Fig. 3: Occupancy detection for house #4

be achieved using different approaches, such as association rule mining, artificial neural networks, hidden Markov models and Bayesian networks. To show a possible application of the dataset, we propose to use a Bayesian network (BN) to predict the usage of user-driven devices. A BN is a probabilistic graphical model encoding the joint probability distribution of a set of random variables. A BN $G = (V, E)$ is a directed acyclic graph (DAG) whose nodes $V = \{X_1, \ldots, X_n\}$ are random variables, while dependency between variable is represented by edges such as $E = X_i \rightarrow X_j$, and quantified by the conditional probability $P(X_j | X_i)$. In particular, each node is associated to a conditional probability distribution (CPD) quantifying the effect the parents have on the node, that is

$$P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_i | pa(X_i)) \text{ [18]}.$$

The approach undertaken follows the work presented in [19, 20] from which we derived the network structure. Because of the limited length of the dataset at time of writing we decided to omit seasonal information expressed through the month (see Fig. 4). As a first step, we processed the daily CSV files collected in the household ID#0 and extracted starting events for each monitored device. The BN was implemented using the Netica[6] tool. In particular, we used expectation maximization (EM) to perform parameter learning for the given network and we used the junction tree (JT) algorithm to perform exact inference. Netica provides APIs for various programming languages, making the learned model applicable in working solutions. The plot in Fig.5 reports the posterior probability of starting the coffee machine of household ID#0, given the observations day of the week (i.e. weekday, Saturday and Sunday) and hour of the day. As noticeable, the patterns of coffee making tend to be regular over the day. The modeled couple tends to wake up earlier during weekdays, which was confirmed by an interview with the householders.
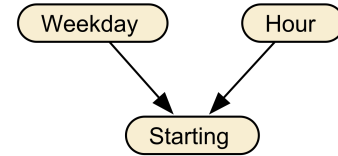

Fig. 4: Bayesian network for the coffee machine of House #0

V. CONCLUSIONS AND FUTURE WORK

We have presented the GREEND consumption dataset, obtained through a measurement campaign in households in Austria and Italy. The dataset is available for open use and consists of consumption data of selected devices. Along with the dataset we have also released our code base for the campaign as a Sourceforge project. This includes the scripts to collect consumption data and a ready-to-use SD image for the Raspberry Pi, as well as processing scripts that can be used to extract edges and consumption events, based on configurable device-specific thresholds. The consumption data and the metadata modeling the consumption scenarios can
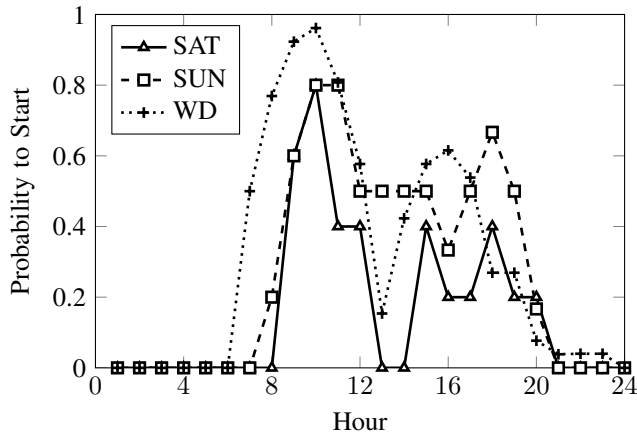
[6]https://www.norsys.com

Fig. 5: Usage forecasting for the coffee machine of House #0

be directly imported into the nilm toolkit. We expect our contribution to be of value for researchers and engineers dealing with domestic energy management systems. To this end, we showed the use of GREEND to develop and assess techniques for load disaggregation, occupancy detection and appliance usage mining. We expect to extend the monitoring campaign to collect aggregated consumption data at a higher frequency, as this allows for modeling the whole demand, which is a valuable information when emulating microgrids and assessing control strategies at a wider scope. Similarly, we plan to include data concerning production from renewable energy, as it can be of value to forecast energy generation. The presence of temporal information, along with the locality and therefore the weather, will allow for the extraction of seasonal consumption patterns. The long term measurement campaign carried out within MONERGY will provide the time span to allow this type of studies. We expect to ultimately provide ready-to-use models of the regions, that can be imported in simulation tools for smart grid applications.

REFERENCES

[1] K. C. Armel, A. Gupta, G. Shrimali, and A. Albert, "Is disaggregation the holy grail of energy efficiency? The case of electricity," *Energy Policy*, vol. 52, no. 0, pp. 213 -- 234, 2013.

[2] A. Monacchi, W. Elmenreich, S. D'Alessandro, and A. M. Tonello, "Strategies for Domestic Energy Conservation in Carinthia and Friuli-Venezia Giulia," in *Proc. of the 39th Annual Conference of the IEEE Industrial Electronics Society*, Vienna, Austria, 2013.

[3] P. Palensky and D. Dietrich, "Demand side management: Demand response, intelligent energy systems, and smart loads," *IEEE Trans. on Ind. Informat.*, vol. 7, no. 3, pp. 381--388, Aug 2011.

[4] N. Batra, J. Kelly, O. Parson, H. Dutta, W. Knottenbelt, A. Rogers, A. Singh, and M. Srivastava, "Nilmtk: An open source toolkit for non-intrusive load monitoring," in *Proceedings of the 5th International Conference on Future Energy Systems*, ser. e-Energy '14. ACM, 2014, pp. 265--276.

[5] J. Kelly and W. Knottenbelt, "Metadata for Energy Disaggregation," in *The 2nd IEEE International Workshop on Consumer Devices and Systems (CDS 2014)*, Västerås, Sweden, Jul. 2014.

[6] C. Gisler, A. Ridi, D. Zujferey, O. Khaled, and J. Hennebert, "Appliance consumption signature database and recognition test protocols," in *8th International Workshop on Systems, Signal Processing and their Applications (WoSSPA)*, May 2013.

[7] S. Makonin, F. Popowich, L. Bartram, B. Gill, and I. V. Bajic, "AMPds: A Public Dataset for Load Disaggregation and Eco-Feedback Research," in *Proceeding of the IEEE Conference on Electrical Power and Energy (EPEC)*, 2013.

[8] K. Anderson, A. Ocneanu, D. Benitez, D. Carlson, A. Rowe, and M. Berges, "BLUED: a fully labeled public dataset for Event-Based Non-Intrusive load monitoring research," in *Proc. of the 2nd KDD Workshop on Data Mining Applications in Sustainability (SustKDD)*, 2012.

[9] N. Batra, M. Gulati, A. Singh, and M. B. Srivastava, "It's Different: Insights into Home Energy Consumption in India," in *Proc. of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, ser. BuildSys'13. New York, NY, USA: ACM, 2013.

[10] J. Z. Kolter and M. J. Johnson, "REDD: A Public Data Set for Energy Disaggregation Research," in *Procddings of the SustKDD Workshop on Data Mining Applications in Sustainability*, 2011.

[11] S. Barker, A. Mishra, D. Irwin, E. Cecchet, P. Shenoy, and J. Albrecht, "Smart*: An open data set and tools for enabling research in sustainable homes," *SustKDD, August*, 2012.

[12] A. Reinhardt, P. Baumann, D. Burgstahler, M. Hollick, H. Chonov, M. Werner, and R. Steinmetz, "On the Accuracy of Appliance Identification Based on Distributed Load Metering Data," in *Proc. of the 2nd IFIP Conference on Sustainable Internet and ICT for Sustainability (SustainIT)*, 2012.

[13] J. Kelly and W. Knottenbelt, ""UK-DALE": A dataset recording UK Domestic Appliance-Level Electricity demand and whole-house demand," *ArXiv e-prints*, Apr. 2014.

[14] M. Zeifman, "Disaggregation of home energy display data using probabilistic approach," *IEEE Trans. Consum. Electron.*, vol. 58, no. 1, pp. 23--31, 2012.

[15] G. Hart, "Nonintrusive appliance load monitoring," *Proc. of the IEEE*, vol. 80, no. 12, pp. 1870--1891, 1992.

[16] D. Egarter, V. P. Bhuvana, and W. Elmenreich, "PALDi: Online load dis-aggregation via particle filtering," *IEEE Transactions on Instrumentation and Measurement (to appear)*, 2014.

[17] D. Chen, S. Barker, A. Subbaswamy, D. Irwin, and P. Shenoy, "Non-intrusive occupancy monitoring using smart meters," in *Proc. of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, ser. BuildSys'13, 2013.

[18] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques - Adaptive Computation and Machine Learning*. The MIT Press, 2009.

[19] L. Hawarah, S. Ploix, and M. Jacomino, "User behavior prediction in energy consumption in housing using bayesian networks," in *Artificial Intelligence and Soft Computing*, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2010, vol. 6113, pp. 372--379.

[20] C. Chang, P.-A. Verhaegen, J. R. Duflou, M. M. Drugan, and A. Nowe, "Finding Days-of-week Representation for Intelligent Machine Usage Profiling," *Journal of Industrial and Intelligent Information*, vol. 1, no. 3, pp. 148--154, 2010.