Energy Disaggregation using Piecewise Affine Regression and Binary Quadratic Programming

Manas Mejari, Vihangkumar V. Naik, Dario Piga and Alberto Bemporad

***Abstract*— In this paper we consider the problem of energy disaggregation, commonly referred in the literature as "non-intrusive load monitoring". The problem is to estimate the end-use power consumption profiles of individual household appliance using only aggregated power measurements. We propose a two-stage supervised approach. At the first stage, dynamical models of individual appliances are estimated using disaggregated training data gathered over a short intrusive period. The consumption profiles of individual appliances are described by *PieceWise Affine AutoRegressive* (PWA-AR) models with multiple operating modes, which are estimated via a moving horizon PWA regression algorithm. Once the model of each appliance is identified, a binary quadratic programming problem is solved at the second stage to determine the set of active appliances which contribute to the instantaneous aggregated power, along with their operating modes. A benchmark dataset is used to assess the performance of the presented disaggregation approach.**

I. INTRODUCTION

Retrieving power consumptions at the single-appliance level provides useful information to energy suppliers, municipalities, and consumers to design and assess efficiency of energy management strategies, increase consumers' awareness on their habits, detect malfunctioning, etc. One can acquire this information via hardware, by attaching a smart meter or a smart plug to every individual appliance. However, this is not economical when there are many devices to monitor. Alternatively, a software-based solution can be used to decompose the aggregate power reading gathered from a single-point smart meter into the individual consumption of each appliance. This approach is known as *Non-Intrusive Load Monitoring* (NILM) or *energy disaggregation*. The advantages of the software-based solution are reduction of intrusiveness into consumers' houses and lower costs for installation, maintenance and replacement of the monitoring system.

A first energy disaggregation algorithm was proposed by Hart in [7], where the aggregate power signal is decomposed to match the individual appliances' typical power demand curves (commonly referred to as *signatures*). The limitation of Hart's approach is that it cannot detect appliances with multiple operating modes and it is neither able to decompose power signals made of simultaneous on/off events on multiple appliances. Thereafter, the NILM problem has been extensively studied in the literature (see [18], [19], [2] and references therein). The main idea behind most of the available methods is to characterize the typical consumption signatures of the appliances using a small set of disaggregated data gathered during a short intrusive period. Once the appliances' signatures are available, disaggregation is performed. Among the available approaches, we mention the ones based on sparse coding [5], blind identification [3], pattern recognition [4], [1], hidden Markov models and its variants [13], [8], [9], deep learning and artificial neural networks [15], [16], integer programming [17], and convex optimization [14].

In this contribution, we propose a novel approach for energy disaggregation that relies on *PieceWise Affine AutoRegressive* (PWA-AR) dynamical models to describe the behavior of the individual appliances. Using a set of disaggregated data collected over a short intrusive period, the PWA-AR models are first estimated off-line using the moving-horizon PWA regression algorithm recently proposed by the authors in [12]. Once the appliance models are estimated, energy disaggregation is formulated as an integer programming problem. Specifically, based on the measurements of the aggregated power, the active operating mode of each appliance (and thus its power consumption) is determined at each time instance in an iterative way. The developed disaggregation algorithm is tested on a benchmark dataset, using PWA-AR dynamic models and also static models defined based on the average power ratings of the devices. The obtained results show that using dynamic models for the individual appliances instead of static models significantly improves the performance.

The paper is organized as follows. In Section II we formally state the problem of energy disaggregation and present the proposed approach in Section III. Specifically, we briefly explain the moving-horizon PWA regression algorithm for estimating individual appliance models in Section III-A, the binary quadratic programming formulation to perform energy disaggregation in Section III-B. Tests on benchmark data set are discussed in Section IV. Finally, concluding remarks are given in Section V.

II. PROBLEM FORMULATION

Consider a household with n different electric appliances connected to the power line. The energy consumption of each appliance is described by *PieceWise Affine AutoRegressive* (PWA-AR) model with $s_i \in \mathbb{N}$ (with $i = 1, \dots, n$) operating modes. Although the appliances may have different operating modes, to simplify the notation we consider the case in which

M. Mejari, V.V. Naik and A. Bemporad are with IMT School for Advanced Studies Lucca, 55100 Lucca, Italy {manas.mejari, vihangkumar.naik, alberto.bemporad}@imtlucca.it

D. Piga is with IDSIA Dalle Molle Institute for Artificial Intelligence SUPSI-USI, 6928 Manno, Switzerland. dario.piga@supsi.ch

all appliances have equal number of operating modes $s \in \mathbb{N}$ described by PWA-AR models having the same dynamical order $n_{\mathrm{a}} \in \mathbb{N}$. More specifically, the power $y_i(k)$ consumed by the i-th appliance at time k is modeled by

$$y_i(k) = \begin{cases} \Theta_{i,1}^\top \left[\begin{smallmatrix} 1 \\ x_i(k) \end{smallmatrix}\right] & \text{if} \quad \delta_{i,1}(k) = 1, \\ \vdots & \\ \Theta_{i,s}^\top \left[\begin{smallmatrix} 1 \\ x_i(k) \end{smallmatrix}\right] & \text{if} \quad \delta_{i,s}(k) = 1, \end{cases} \tag{1}$$

where $\delta_{i,j}(k) \in \{0, 1\}$ (with $j = 1, \ldots, s$) is a binary variable which is used to characterize the active operating mode of the appliance (i.e., the j-th mode is active if and only if $\delta_{i,j}(k) = 1$), $\Theta_{i,j}$ is a set of parameters describing the behavior of the i-th appliance at the j-th operating mode and $x_i(k)$ denotes the regressor vector containing the past values of the outputs

$$x_i(k) = [y_i(k-1), \ldots,\ y_i(k-n_{\mathrm{a}})]^\top. \tag{2}$$

At a given time k, one and only one mode of the appliance is active, i.e., $\sum_{j=1}^{s} \delta_{i,j}(k) = 1$. The measured aggregated power reading is

$$y(k) = \sum_{i=1}^{n} y_i(k) + e(k), \tag{3}$$

with $e(k)$ taking into account the measurement noise and unmodeled appliances.

Problem 1 (energy disaggregation problem): Given an N-length data sequence $\mathcal{D} = \{y(k)\}_{k=1}^{N}$ of aggregated power signals $y(k)$, estimate the end-use power consumption profiles $y_i(k)$ (with, $i = 1, \ldots, n$ and $k = 1, \ldots, N$).

III. Energy Disaggregation Algorithm

We detail a supervised disaggregation algorithm which consists of two stages:

S1. The *PieceWise Affine AutoRegressive* (PWA-AR) model in (1) describing the behavior of individual appliances is estimated via the PWA regression algorithm recently developed by the authors in [12], using disaggregated training data collected over a short intrusive period. It is a common practice to use a small set of disaggregated data to learn the signature of the appliances [14], [18], [19].

S2. Using the PWA-AR models obtained from stage **S1**, an integer programming problem is solved iteratively to determine the active devices contributing to the instantaneous total power, along with their corresponding operating modes.

In the following sections, stages **S1** and **S2** are described in detail.

A. Stage *S1*: Training appliance models

In this section, we discuss the identification of PWA-AR models (1) using the regularized moving-horizon approach proposed in [12] for piecewise affine regression.

For each appliance, consider a set of training data of length $\bar{N}$ consisting of the disaggregated power consumption $\{y_i(k)\}_{k=1}^{\bar{N}}$. The training regressor/output pairs $\{x_i(k), y_i(k)\}$, with $x_i(k)$ defined in (2), are processed iteratively. At each time sample k, a moving-horizon window of length $N_p \ll \bar{N}$ containing regressor/output pairs from time $k-N_p+1$ to time k is considered. The model parameters $\Theta_{i,j}$ and the binary variables $\delta_{i,j}(k)$ (for $j = 1, \ldots, s$) at time k are estimated simultaneously by solving the mixed-integer quadratic programming problem:

$$\min_{\substack{\Theta_{i,j} \\ \delta_{i,j}(k-t)}} \sum_{j=1}^{s} \sum_{t=0}^{N_p-1} \left\| \left(y_i(k-t) - \Theta_{i,j}^\top \left[\begin{smallmatrix} 1 \\ x_i(k-t) \end{smallmatrix}\right]\right) \delta_{i,j}(k-t) \right\|_2^2 \tag{4a}$$

$$+ \sum_{t=1}^{k-N_p} \left\| y_i(t) - \Theta_{i,\sigma(t)}^\top \left[\begin{smallmatrix} 1 \\ x_i(t) \end{smallmatrix}\right] \right\|_2^2 \tag{4b}$$

$$\text{s.t. } \delta_{i,j}(k-t) \in \{0,1\},\ \sum_{j=1}^{s} \delta_{i,j}(k-t) = 1,\ t = 0, \ldots, N_p-1. \tag{4c}$$

The objective of problem (4) is to determine the optimal sequence of active modes

$$\sigma_i(k-t) = j^* \Leftrightarrow \delta_{i,j^*}(k-t) = 1, \quad t = 0, \ldots, N_p - 1,$$

within the considered time window, along with the model parameters $\Theta_{i,j}$ (for each appliance i and mode j), which best match the available power consumption up to time k. The term (4b) is a regularization term on the parameters $\Theta_{i,j}$, which takes into account the past data outside the considered horizon. More specifically, in (4b) the active mode sequence is not optimized from time 1 to $k - N_p$, but it is fixed to the estimates $\{\sigma_i(t)\}_{t=1}^{k-N_p}$ computed from the previous iterations of the moving-horizon estimation algorithm. In turn, the sequence of active modes is optimized only within the considered time horizon in (4a).

At time k, only the active mode $\sigma_i(k)$ is kept, and the N_p-length time window is shifted forward to process the next pair $\{x_i(k+1), y_i(k+1)\}$ in a moving-horizon fashion. Increasing N_p provides more data to estimate the model parameters $\Theta_{i,j}$ and the sequence of active modes $\sigma_i(k)$, which improves the accuracy of the estimates. On the other hand, increasing N_p increases the number of binary decision variables $\delta_{i,j}$. Thus, the parameter N_p acts as a tuning knob to trade off accuracy vs complexity. At the end of the training phase, the signature of the i-th appliance is captured by the estimated model parameters $\Theta_{i,j}$ for all modes $j = 1, \ldots, s$.

B. Stage *S2*: Energy disaggregation

Energy disaggregation means to determine the operating mode of each appliance. To this end we solve the following binary quadratic program

$$\min_{\{\delta_{i,j}(k)\}_{i,j=1}^{n,s}} \left\| y(k) - \sum_{i=1}^{n} \sum_{j=1}^{s} \Theta_{i,j}^\top \left[\begin{smallmatrix} 1 \\ \hat{x}_i(k) \end{smallmatrix}\right] \delta_{i,j}(k) \right\|_2^2, \tag{5a}$$

$$\text{s.t. } \delta_{i,j}(k) \in \{0,1\},\ \sum_{j=1}^{s} \delta_{i,j}(k) = 1, \tag{5b}$$

at each time instance k, where $y(k)$ is the measurement of the aggregated power, $\Theta_{i,j}$ are model parameters estimated at stage **S1**, $\hat{x}_i(k)$ is the estimated regressor vector[1] defined as

$$\hat{x}_i(k) = [\hat{y}_i(k-1), \ldots,\ \hat{y}_i(k-n_{\mathrm{a}})]^\top,$$

where $\hat{y}_i(k)$ is the estimate of the disaggregated power for the i-th appliance constructed as follows.

At each time instance $k = 1, \ldots, N$ the binary quadratic program (5) is solved iteratively using an estimate $\hat{x}_i(k)$ of the regressor obtained from the previous iterations. Specifically, at each iteration k the active operating mode j^* of each appliance is determined by the solution of problem (5), namely

$$j^* : \delta_{i,j^*}(k) = 1.$$

The power of each individual appliance is thus given by

$$\hat{y}_i(k) = \Theta_{i,j^*}^\top \left[\begin{smallmatrix} 1 \\ \hat{x}_i(k) \end{smallmatrix}\right],$$

which is used to construct the regressor $\hat{x}_i(k+1)$.

IV. APPLICATION TO REAL DATA

The proposed disaggregation algorithm is tested on a benchmark AMPds dataset [11], which consists of power consumptions of 19 appliances monitored from April 1, 2012 to March 31, 2013 at one minute read intervals in a house located in Canada. In our analysis we consider only the aggregate power consumption given by the sum of the power consumption readings of the following four electric appliances: 1) fridge (FGE); 2) dish washer (DWE); 3) heat pump (HPE); 4) clothes dryer (CDE). Moreover, for a realistic scenario the aggregated power is corrupted by a fictitious white noise $e(k)$ with Gaussian distribution $\mathcal{N}(0, \sigma_{\mathrm{e}}^2)$ having standard deviation $\sigma_{\mathrm{e}} = 4$ W.

The computations are carried out on an Intel i5 1.7 GHz running MATLAB R2016b.

A. Performance measures

The quality of the energy disaggregation results is assessed via the following performance measures [10], [14]:

1. *Energy Fraction Index* (EFI)
   The EFI index
   $$\hat{h}_i = \frac{\sum_{k=1}^{N} \hat{y}_i(k)}{\sum_{i=1}^{n} \sum_{k=1}^{N} \hat{y}_i(k)}$$
   quantifies the estimated fraction of total energy consumed by the i-th appliance. This index provides an important information to the user for potential savings. In order to asses the effectiveness of the algorithm the index $\hat{h}_i$ is compared with an analogous index defined based on the actual disaggregated profiles, namely
   $$h_i = \frac{\sum_{k=1}^{N} y_i(k)}{\sum_{i=1}^{n} \sum_{k=1}^{N} y_i(k)}.$$
   We remark that the true disaggregated power profiles y_i are not used in the disaggregation algorithm (stage **S2**) but only to evaluate estimation performance.
2. *Relative Square Error (RSE) and R^2 coefficient*
   The normalized error between the actual and the estimated power consumption is quantified for the i-th appliance by the RSE_i index defined as
   $$\mathrm{RSE}_i = \frac{\sum_{k=1}^{N} (y_i(k) - \hat{y}_i(k))^2}{\sum_{k=1}^{N} y_i^2(k)},$$
   and the R_i^2 coefficient defined as
   $$R_i^2 = 1 - \frac{\sum_{k=1}^{N} (y_i(k) - \hat{y}_i(k))^2}{\sum_{k=1}^{N} (y_i(k) - \bar{y}_i)^2},$$
   with $\bar{y}_i = \frac{1}{N}\sum_{k=1}^{N} y_i(k)$. Both RSE_i and R_i^2 measure the match between the actual and the estimated power profiles over time. Indeed, low values of RSE_i (or equivalently high values of R_i^2) imply an accurate estimate of the index $\hat{h}_i$.
3. *Total Energy Correctly Assigned (TECA)*
   The TECA index
   $$\mathrm{TECA} = 1 - \frac{\sum_{k=1}^{N} \sum_{i=1}^{n} |\hat{y}_i(k) - y_i(k)|}{2\sum_{k=1}^{N} y(k)},$$
   quantifies the percentage of energy correctly classified. A precise estimate of the power consumption profiles gives the information to the consumer about the use of household appliances over time. This is essential for potential power savings as well as monetary benefits as the consumer can differ the use of some appliances to off-peak hours.

B. Supervised training phase

At stage **S1**, *PieceWise Affine AutoRegressive* (PWA-AR) models (1) describing the behavior for each appliance are estimated as discussed in Section III-A using only 500 min data of the day 19 (April 19, 2012) for fridge, dish washer, heat pump and day 38 (May 8, 2012) for clothes dryer as an intrusive period. PWA-AR models with $s = 3$ modes and order $n_a = 2$ are considered. The moving-horizon mixed-integer quadratic programming problem (4) is solved with horizon length $N_p = 5$ using Gurobi [6]. The average computation time to solve (4) is 90 ms.

The results of the training phase are reported in Fig. 1, which shows that the estimated PWA-AR models accurately capture the behavior of the individual appliances' power consumption profiles.

For comparison, the use of static device models for energy disaggregation is reported in the next section. Static models are a special case of PWA-AR models (1) in which the current output $y_i(k)$ does not depend on the past output values i.e. $y_i(k) = \Theta_{i,j}$. The parameter $\Theta_{i,j}$ thus models the power consumption of the i-th appliance operating at the j-th mode and is chosen via simple visual inspection of the training data. The selected values of the parameters $\Theta_{i,j}$ are

a. fridge: $[\Theta_{1,1}\ \Theta_{1,2}\ \Theta_{1,3}] = [0\ 128\ 200]$ W;

[1] The true regressor $x_i(k)$ defined in (2) can not be constructed as it depends on the past values of the individual appliance power y_i which are not available at the stage **S2**.

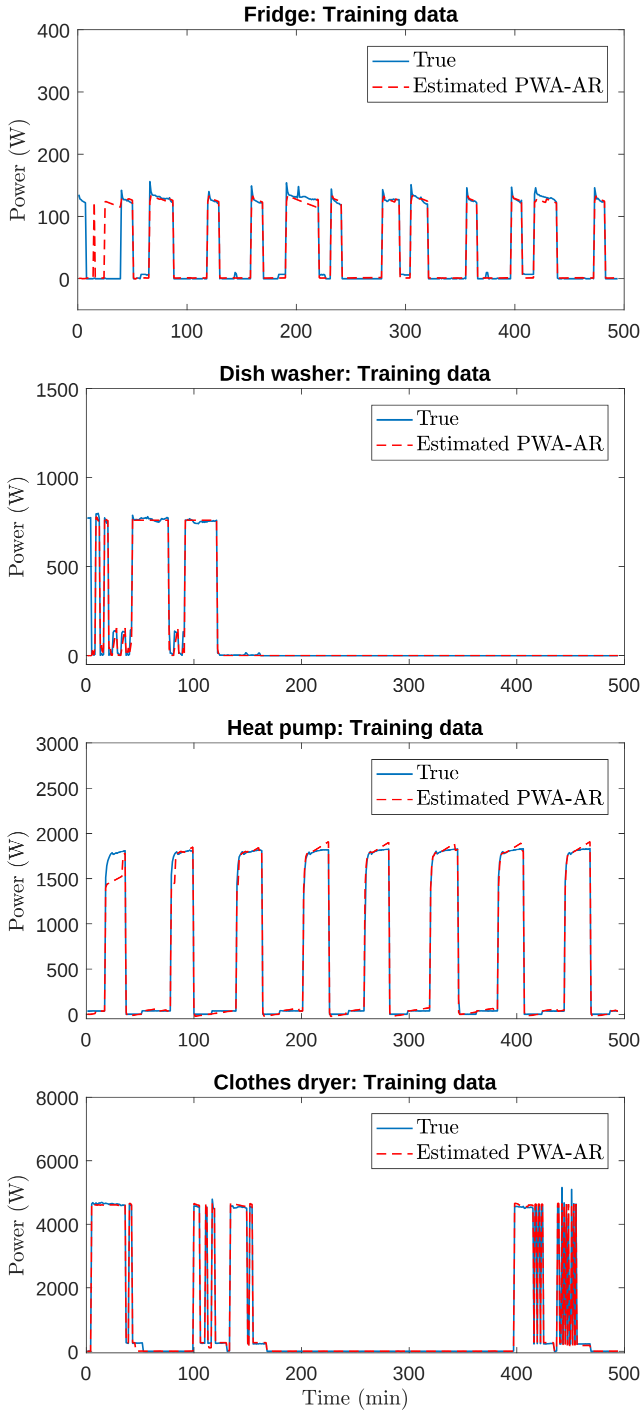

Fig. 1: Supervised learning: True vs estimated power consumption.

b. dish washer: $[\Theta_{2,1}\ \Theta_{2,2}\ \Theta_{2,3}] = [0\ 120\ 800]$ W;
c. heat pump: $[\Theta_{3,1}\ \Theta_{3,2}\ \Theta_{3,3}] = [0\ 39\ 1900]$ W;
d. clothes dryer: $[\Theta_{4,1}\ \Theta_{4,2}\ \Theta_{4,3}] = [0\ 260\ 4700]$ W.

C. *Energy disaggregation*

Once the PWA-AR models of individual appliances are estimated, the power measurements of one month (from 1st to 30th June, 2012) are disaggregated by solving the binary quadratic program (5) iteratively using Gurobi. The average CPU time taken to solve problem (5) is 8 ms. In the case of static models (introduced in Section IV-B), iterative construction of the estimated regressor $\hat{x}_i(k)$ is not required for solving problem (5).

TABLE I: Estimated Energy Fraction Index $\hat{h}_i$ and Actual Energy Fraction Index h_i. Results obtained using PWA-AR models and static models.

	PWA-AR models $\hat{h}_i$	static models $\hat{h}_i$	ground truth h_i
Fridge	19.6 %	14.9 %	21.3 %
Dish washer	6.8 %	11.4 %	5.1 %
Heat pump	41.6 %	42.0 %	42.3 %
Clothes dryer	31.9 %	31.6 %	31.3 %

TABLE II: Relative Square Errors and R^2 coefficients. Results obtained using PWA-AR models and static models.

	PWA-AR models		static models	
	RSE$_i$	R_i^2	RSE$_i$	R_i^2
Fridge	15.5 %	76.5 %	35.9 %	45.5 %
Dish washer	12.4 %	87.3 %	38.0 %	61.4 %
Heat pump	0.6 %	99.3 %	4.0 %	95.6 %
Clothes dryer	0.1 %	99.9 %	0.3 %	99.7 %

TABLE III: Total Energy Correctly Assigned (TECA). Results obtained using PWA-AR models and static models.

	PWA-AR models	static models
TECA	95.3 %	89.4 %

The results obtained by using PWA-AR models (estimated in stage **S1**) and by using static models, are reported in Table I, Table II and Fig. 2. For visualization purpose, only a portion of the disaggregated power profiles is plotted in Fig. 2. The proposed algorithm using PWA-AR models accurately estimates the power consumption trajectories of each individual appliance over time as shown in Fig. 2. The efficiency of the method is also reflected in the performance measures reported in Table I, II and III.

We remark that for fridge and dish washer, the second operating mode of both the appliances have similar static models with parameters 128 W and 120 W respectively. From the obtained results it can be observed that using only static models it is difficult to distinguish between these two devices. On the contrary, thanks to their dynamic nature PWA-AR models are able to resolve such a conflict. Overall, PWA-AR models outperform static models in the task of energy disaggregation.

V. CONCLUSIONS

We have proposed a two-stage algorithm for energy disaggregation. In the first stage, a small set of training data consisting of disaggregated power profiles for individual appliances is used to estimate *PieceWise Affine AutoRegressive* (PWA-AR) models employing a supervised moving horizon PWA regression algorithm recently proposed in [12]. Once the model parameters are estimated for each appliance, the energy disaggregation problem is formulated as a binary

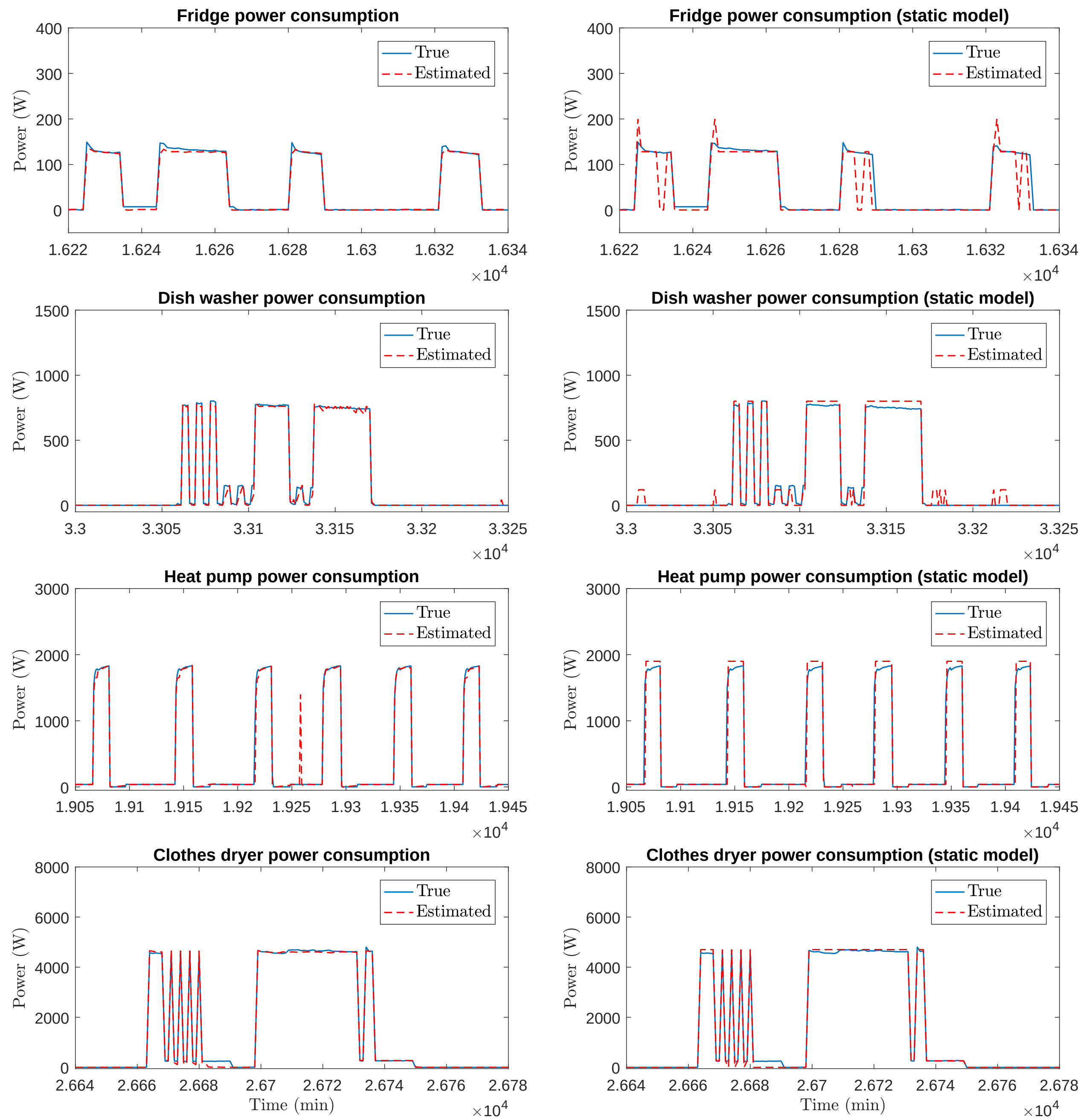

Fig. 2: Disaggregated power consumption profiles. Results obtained using PWA-AR models (left panels) and static models (right panels).

quadratic program. The dynamic modeling of the power profiles of individual appliances leads to better energy disaggregation results compared to the same approach relying on static models. This is due to the fact that the dynamic models are able to capture the transient behavior, thus providing vital information to distinguish between the appliances having similar power signatures.

The proposed method is computationally efficient as the appliance models can be estimated off-line only once, while energy disaggregation is performed online with low computational complexity. Thus, the approach proposed in this paper is promising for embedded implementation in smart meters.

References

[1] T. R. Camier, S. Giroux, B. Bouchard, and A. Bouzouane. Designing a NIALM in smart homes for cognitive assistance. *Procedia Computer Science*, 19:524–532, 2013.

[2] A. Cominola, M. Giuliani, D. Piga, A Castelletti, and A. E. Rizzoli. Benefits and challenges of using smart meters for advancing residential water demand modeling and management: a review. *Environmental Modelling & Software*, 72:198–214, 2015.

[3] R. Dong, L. Ratliff, H. Ohlsson, and S. Sastry. A dynamical systems approach to energy disaggregation. In *Proc. of the 52nd IEEE Conference on Decision and Control*, pages 6335–6340, Florence, Italy, 2013.

[4] L. Farinaccio and R. Zmeureanu. Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses. *Energy and Buildings*, 30(3):245–259, 1999.

[5] M. Figueiredo, B. Ribeiro, and A.M. de Almeida. On the regularization parameter selection for sparse code learning in electrical source separation. In *Adaptive and Natural Computing Algorithms*, pages 277–286. Springer, 2013.

[6] Gurobi Optimization, Inc. *Gurobi Optimizer Reference Manual*, 2014.

[7] G.W. Hart. Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, 80(12):1870–1891, 1992.

[8] M.J. Johnson and A.S. Willsky. Bayesian nonparametric hidden semi-Markov models. *The Journal of Machine Learning Research*, 14(1):673–701, 2013.

[9] J.Z. Kolter and T. Jaakkola. Approximate inference in additive factorial HMMs with application to energy disaggregation. In *Int. Conference on Artificial Intelligence and Statistics*, pages 1472–1482, 2012.

[10] J.Z. Kolter and M.J. Johnson. REDD: a public data set for energy disaggregation research. In *Proc. of the SustKDD workshop on Data Mining Applications in Sustainability*, San Diego, CA, USA, 2011.

[11] S. Makonin, F. Popowich, L. Bartram, B. Gill, and I.V. Bajic. AMPds: a public dataset for load disaggregation and eco-feedback research. In *Electrical Power and Energy Conference (EPEC)*, pages 1–6, 2013.

[12] V. V. Naik, M. Mejari, D. Piga, and A. Bemporad. Regularized moving-horizon piecewise affine regression using mixed-integer quadratic programming. In *Proc. 25th Mediterranean Conference on Control and Automation*, pages 1349–1354, Valletta, Malta, 2017.

[13] O. Parson, S. Ghosh, M. Weal, and A. Rogers. An unsupervised training method for non-intrusive appliance load monitoring. *Artificial Intelligence*, 217:1–19, 2014.

[14] D. Piga, A. Cominola, M. Giuliani, A. Castelleti, and A. E. Rizzoli. Sparse optimization for automated energy end use disaggregation. *IEEE Transactions on Control Systems Technology*, 24(3):1044–1051, 2016.

[15] A. G. Ruzzelli, C. Nicolas, A. Schoofs, and G.M.P. O'Hare. Real-time recognition and profiling of appliances through a single electricity sensor. In *IEEE Communications Society Conference on Sensor Mesh and Ad Hoc Communications and Networks (SECON)*, pages 1–9, 2010.

[16] D. Srinivasan, W.S. Ng, and A.C. Liew. Neural-network-based signature recognition for harmonic source identification. *IEEE Transactions on Power Delivery*, 21(1):398–405, 2006.

[17] K. Suzuki, S. Inagaki, T. Suzuki, H. Nakamura, and K. Ito. Nonintrusive appliance load monitoring based on integer programming. In *SICE Annual Conference, 2008*, pages 2742–2747, 2008.

[18] M. Zeifman and K. Roth. Nonintrusive appliance load monitoring: Review and outlook. *IEEE Transactions on Consumer Electronics*, 57(1):76–84, 2011.

[19] A. Zoha, A. Gluhak, M.A. Imran, and S. Rajasegarar. Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey. *Sensors*, 12(12):16838–16866, 2012.