Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 51-5 (2018) 108–113

Kalman filtering for energy disaggregation

V. Breschi * D. Piga ** A. Bemporad *

* IMT School for Advanced Studies Lucca, Piazza San Francesco 19, 55100 Lucca, Italy (e-mail: valentina.breschi@imtlucca.it, alberto.bemporad@imtlucca.it).

** IDSIA Dalle Molle Institute for Artificial Intelligence SUSPI-USI, 6928 Manno, Switzerland (e-mail: dario.piga@supsi.ch)

Abstract: Providing the users information on the energy consumed in the household at the appliance level is of major importance for increasing their awareness of their consumption behavior. In this paper, we propose a technique based on Kalman filters to estimate the devices' consumption patterns from aggregate readings, *i.e.*, to solve the so called *disaggregation problem.* The method is suited for on-line disaggregation and the proposed results show that it is robust against modelling errors and unmodelled appliances.

Keywords: Energy disaggregation, Kalman filter, Markov jump models, recursive estimate.

1. INTRODUCTION

Knowing the power consumed by the electronic devices in an household is of paramount importance when designing strategies to reduce the household consumption and to identify a malfunctioning or an anomaly in the appliance behavior. A possible approach to estimate the power consumed by the single devices is to install a smart sensor for each appliance. Such a solution leads to increasing costs of installation and maintenance due to the high number of sensors that have to be deployed. Alternatively, one smart meter measuring aggregate power can be installed and, then, tailored algorithms can be used to estimate the power consumed by the single appliances from the aggregate readings. In this work, we present a technique for *energy disaggregation* (or Non-Intrusive load monitoring), *i.e.*, a method for the reconstruction of the single appliances patterns from aggregate power measurements.

The first method for Non-intrusive load monitoring (NILM) was introduced in Hart (1992), where a signal processing technique based on the appliances' signatures (*i.e.*, their typical patterns) is used to identify their on/off events. However, this method does not allow one to reconstruct the consumption patterns of the devices. Since the seminal work of Hart, different approaches have been proposed to solve the disaggregation problem, which are reviewed in Zeifman and Roth (2011) and in Cominola et al. (2015), with applications also in the water sector. These methods mainly rely either in the formulation of NILM as an optimization problem, *e.g.*, Piga et al. (2016); Suzuki et al. (2008) or in the representation of the household consumption through Factorial Hidden Markov Models (FHMMs), *e.g.*, Kolter and Jaakkola (2012); Bonfigli et al. (2017); Cominola et al. (2017).

In this paper, energy disaggregation is addressed using Kalman filters. Non-intrusive load monitoring is thus reformulated as a state estimation problem and, consequently, a state-space model for the household consumption has to be identified. By approximating the behavior of the different appliances through jump models using the method proposed in Bemporad et al. (2017), we describe how a linear switching state-space model for the household consumption is obtained. Due to the switching nature of the model, we present a disaggregation method based on the use of a multi-model Kalman filter (Bar-Shalom et al., 2002, Chapter 11). At each time instant, this approach allows us to estimate the (unknown) operating condition of each appliance and to reconstruct the consumption patterns of the single devices, based on the estimated state. Thanks to the recursive nature of the Kalman filter, the method is suited for on-line disaggregation and it can potentially be used to obtain indications on the single appliances consumption in real-time.

The paper is organized as follows. The energy disaggregation problem is formulated in Section 2. Section 3 is devoted to the description of the method used to retrieve a model for the household energy consumption. The disaggregation approach is then described in Section 4, while the results obtained testing the method against the AMPDs dataset Makonin et al. (2016) are presented in Section 5.

2. PROBLEM FORMULATION

Assume N different appliances are connected to the power line of a house and let $y_i(t)$ denote the power consumed by the i-th device ($i \in \{1, \dots, N\}$) at time $t \in \mathbb{N}$, with $\mathbb{N}$ indicating the set of natural numbers.

Denote with $y(t)$ the aggregate power measured at the household level at time t, which is given by

$$y(t) = \sum_{i=1}^{N} y_i(t) + e(t), \tag{1}$$

with $e(t)$ accounting for unmodelled appliances and measurement noise on the aggregate reading. The *disaggregation problem* addressed in this paper aims at reconstructing the (unknown) appliances' consumption patterns

Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2018.06.219

$\{y_i(t)\}_{t=1}^{T}$, with $i = 1, \dots, N$, given a sequence $\mathcal{D}_T = \{y(t)\}_{t=1}^{T}$ of aggregate readings.

3. MODELLING THE HOUSEHOLD CONSUMPTION

The method for energy disaggregation proposed in the paper requires the knowledge of a model for the household consumption. To retrieve it, we identify the models for the single appliances and, then, we combine them.

Each electrical device is characterized by different operating conditions. Accordingly, we model the i-th appliance using a jump model characterized by $K_i \in \mathbb{N}$ different modes. In case K_i is not known a priori, the number of operating conditions can be selected through cross-validation. At each mode, the behavior of the device is described through an affine model, *i.e.*, the power reading $y_i(t)$ is approximated as

$$y_i(t) \approx X_i(t)\theta_i^{s_i(t)}, \tag{2}$$

with $X_i(t)$ and $s_i(t) \in \{1, \dots, K_i\}$ being the feature vector and the active mode at time t, respectively. The feature vector is chosen dependently on the class of sub-models we aim at identifying. For example, $X_i(t) = 1$ if we want to estimate static sub-models ($y_i(t) \approx \theta_i^{s_i(t)}$). Instead, if dynamic models are identified, the feature vector is a collection of past power readings, *i.e.*,

$$X_i(t) = [y_i(t-1) \ \cdots \ y_i(t-n_i) \ 1], \tag{3}$$

with n_i being the order of the model.

Both the parameters $\Theta_i = (\theta_i^1, \dots, \theta_i^{K_i})$ and the active mode at each time instant $s_i(t)$ are unknown and, thus, they have to be retrieved from data.

Given a collection of power readings of the i-th appliance $\mathcal{M}_i = \{y_i(t)\}_{t=1}^{\bar{T}}$ [1], we estimate Θ_i and the active mode sequence $S_i = \{s_i(t)\}_{t=1}^{\bar{T}}$ minimizing the cost function

$$J_i(S_i, \Theta_i) = \ell(y_i(\bar{T}), s_i(\bar{T}), \Theta_i) + \sum_{t=1}^{\bar{T}-1} L(y_i(t), S_i, \Theta_i), \tag{4}$$

with

$$L(y_i(t), S_i, \Theta_i) = \ell(y_i(t), s_i(t), \Theta_i) + \mathcal{L}_i(s_i(t+1), s_i(t)). \tag{5}$$

The objective function in (4) allows us to account for the *fitting error* through $\ell(y_i(t), s_i(t), \Theta_i)$, which is chosen as

$$\ell(y_i(t), s_i(t), \Theta_i) = \frac{1}{\bar{T}} \left(y_i(t) - X_i(t)\theta_i^{s_i(t)} \right)^2, \tag{6}$$

and it enables us to weight transitions in the active mode sequence through $\mathcal{L}_i(s_i(t+1), s_i(t))$. When the measurements are taken at high sampling rates (*e.g.*, 1 min), it is reasonable to assume that the appliances seldom change their operating condition over time. We have chosen $\mathcal{L}_i$ as

$$\mathcal{L}_i(s_i(t+1), s_i(t)) = \begin{cases} \lambda_i(s_i(t)) \text{ if } s_i(t+1) \neq s_i(t) \\ 0 \text{ otherwise,} \end{cases} \tag{7}$$

accordingly. The parameters $\lambda_i(j)$, with $j = 1, \dots, K_i$ can be tuned through cross-validation or it can be selected following the guidelines provided in Bemporad et al. (2017). The cost in (4) is optimized with respect to Θ_i and S_i using the coordinate descent algorithm recently presented in Bemporad et al. (2017). Specifically, starting from an initial mode sequence S_i^{o}, we minimize the function in (4) with respect to Θ_i. Then we refine the estimate of $S_i(t)$ by minimizing (4) with respect to the mode sequence through Dynamic Programming (DP). This procedure is repeated until convergence, *i.e.*, the iterations are terminated when the same sequence is estimated at two consecutive runs.

[1] These datasets are gathered from a short intrusive period and they are disjoint from the one over which disaggregation is performed.

As the disaggregation method proposed in this work relies on Kalman filters, we need a lumped state-space model for the household consumption on the bases of the models identified for the single appliances. Due to the switching nature of the devices' models, we obtain a linear switching state-space model given by

$$x(t+1) = A[s(t)]x(t) + B[s(t)]u(t) + w[s(t)], \tag{8a}$$

$$y(t) = C[s(t)]x(t) + v[s(t)], \tag{8b}$$

with $y(t)$ and $s(t) = [\, s_1(t) \ \dots \ s_N(t) \,] \in \mathcal{S}$ being the aggregate reading and the *joint active mode sequence* at time t, respectively. The state $x(t)$ and the matrices $A[s(t)]$, $B[s(t)]$, $C[s(t)]$ are defined on the basis of the devices' models, *e.g.*, when all the sub-models are static $x(t) = [\, y_1(t) \ \dots \ y_N(t) \,]$ is the collection of the power consumed by the single appliances and

$$A[h] = 0_{N \times N} \quad B[h] = \begin{bmatrix} \hat{\theta}_1^{h_1} \\ \vdots \\ \hat{\theta}_N^{h_N} \end{bmatrix} \quad C[h] = 1_N',$$

for $h = 1, \dots, |\mathcal{S}|$. The input $u(t)$ is always equal to 1, for all time instants. Under standard assumptions in Kalman filtering, the process noise $w[s(t)]$ and the measurement noise $v[s(t)]$ are assumed to be mutually independent. Furthermore, $w[s(t)]$ and $v[s(t)]$ are supposed to be generated by zero-mean Gaussian distributions with covariances $Q[s(t)]$ and $R[s(t)]$, respectively. The covariance matrices $Q[h]$ and $R[h]$, for $h = 1, \dots, \mathcal{S}$, weight the confidence in the model and the measurement, respectively, and they have thus to be tuned accordingly.

The evolution of the joint mode $s(t)$ over time is described by a stationary Markov Chain with constant transition probabilities

$$\pi_{d,h} = P(s(t+1) = h | s(t) = d), \quad h, d \in \mathcal{S}. \tag{9}$$

Under the assumption that the appliances change their operating mode independently from each other, $\pi_{d,h}$ is computed as

$$\pi_{d,h} = \prod_{i=1}^{N} P(s_i(t+1) = h_i | s_i(t) = d_i), \tag{10}$$

where $\pi_{d,h}^i = P(s_i(t+1) = h_i | s_i(t) = d_i)$ is obtained on the basis of the estimated sequence $\hat{S}_i$ as

$$\pi_{d,h}^i = \frac{\sum_{t=1}^{\bar{T}-1} \mathbf{1}(\hat{s}_i(t) = d_i, \hat{s}_i(t+1) = h_i) + 1}{\sum_{t=1}^{\bar{T}-1} \mathbf{1}(\hat{s}_i(t) = d_i) + K_i^2}, \tag{11}$$

with $\mathbf{1}(\hat{s}_i(t) = d_i, \hat{s}_i(t+1) = h_i)$ and $\mathbf{1}(\hat{s}_i(t) = d_i)$ being the indicator functions of the transition between mode d_i and h_i and the event $\hat{s}_i(t) = d_i$, respectively.

4. DISAGGREGATION WITH KALMAN FILTERS

Given a switching dynamical model for the household consumption as in (8), we address the disaggregation problem

Algorithm 1 Disaggregation method

Input: Aggregate output readings flow $y(1), y(2), \ldots$; models $A[h], B[h], C[h]$ and noise covariance matrices $Q[h], R[h]$, $h \in \mathcal{S}$; prior on the initial state $\hat{x}(0|0), P_x(0|0)$, initial mode probabilities $\alpha[h](0|0)$, $h \in \mathcal{S}$.

1. **iterate for** $t = 1, 2, \ldots$
   1.1. **update** $\hat{x}[h](t|t), P_x[h](t|t)$, $h \in \mathcal{S}$, using linear Kalman filter;
   1.2. **compute** the likelihood $p[y(t)|s(t) = h, \mathcal{I}^{t-1}]$, $h \in \mathcal{S}$, as in (16);
   1.3. **update** $\alpha[h](t|t)$, $h \in \mathcal{S}$, as in (15);
   1.4. $\hat{s}(t) \leftarrow \arg\max_{h\in\mathcal{S}} \alpha[h](t|t)$, $h \in \mathcal{S}$;
   1.5. **compute** $\hat{x}(t|t)$ as in (13a);
   1.6. **compute** $P_x(t|t)$ as in (13b);

Output: Estimated sequence of joint mode $\hat{s}(t)$ and optimal state $\hat{x}[\hat{s}(t)](t|t)$.

through multi-model Kalman filters. Specifically, we use the *first generalized pseudo-Bayesian* (GPB$_1$) algorithm (Bar-Shalom et al., 2002, Chapetr 11) to iteratively perform disaggregation over a set $D_T = \{y(t)\}_{t=1}^T$ of aggregate readings. This method allows us to estimate the state of the system $x(t)$ and to infer the unknown joint mode $s(t)$ at each time instant t.

The disaggregation approach is summarized in Algorithm 1. It has to be underlined that the state $x(0)$ is assumed to be generated by a Gaussian distribution with mean and covariance equal to the state estimate $\hat{x}(0|0)$ and its covariance matrix $P_x(0|0)$, *i.e.*, $x(0) \sim \mathcal{N}(\hat{x}(0|0), P_x(0|0))$. As in the GPB$_1$ algorithm, we run $|\mathcal{S}|$ Kalman filters in parallel, one for each possible combination of modes of the single devices (Step 1.1). The h-th Kalman filter outputs the updated state estimate for the h-th joint mode $\hat{x}[h](t|t)$ and its covariance $P_x[h](t|t)$, based on the aggregate measurement $y(t)$ and on the model

$$x[h](t+1) = A[h]x[h](t) + B[h]u(t) + w[h], \quad (12a)$$
$$y(t) = C[h]x[h](t) + v[h]. \quad (12b)$$

However, we want to obtain an estimate for the state of the system $\hat{x}(t|t)$, its covariance $P_x(t|t)$ and to identify the joint mode $s(t)$. Based on the hypothesis made with respect to the initial state and the process and measurement noise, it can be proven (see (Bar-Shalom et al., 2002, Chapter)) that $x(t)$ is distributed as a mixture of Gaussian with mean $\hat{x}(t|t)$ and covariance $P_x(t|t)$ given by

$$\hat{x}(t|t) = \sum_{h\in\mathcal{S}} \alpha[h](t|t)\hat{x}[h](t|t), \quad (13a)$$
$$P_x(t|t) = \sum_{h\in\mathcal{S}} \alpha[h](t|t)\left[P_x[h](t|t) + \epsilon[h](t|t)\epsilon[h](t|t)'\right], \quad (13b)$$

where $\epsilon[h](t|t) = \hat{x}[h](t|t) - \hat{x}(t|t)$.

The weights $\alpha[h](t|t)$, for all $h \in \mathcal{S}$, allow to account for the uncertainty on $s(t)$ and they are equal to the probabilities associated with the different joint modes, *i.e.*,

$$\alpha[h](t|t) = P(s(t) = h|\mathcal{I}^t), \quad (14)$$

with $\mathcal{I}^t = \{y(1), \ldots, y(t)\}$ being the set of available aggregate readings up to time t. Exploiting the assumption on the evolution of the joint mode, these probabilities can be computed in a recursive fashion as

$$\alpha[h](t|t-1) = \sum_{d\in\mathcal{S}} \pi_{d,h}\alpha[d](t-1|t-1), \quad (15a)$$
$$\alpha[h](t|t) = \frac{p(y(t)|h, \mathcal{I}^{t-1})\alpha[h](t|t-1)}{\sum_{j\in\mathcal{S}} p(y(t)|j, \mathcal{I}^{t-1})\alpha[j](t|t-1)}, \quad (15b)$$

where $p(y(t)|h, \mathcal{I}^{t-1})$ is the likelihood of the aggregate reading $y(t)$ given the mode h and the information up to $t-1$. The likelihood $p(y(t)|h, \mathcal{I}^{t-1})$ is computed at Step 1.2, approximating the past through $\hat{x}(t-1|t-1)$ and its covariance $P_x(t-1|t-1)$, *i.e.*,

$$p[y(t)|h, \mathcal{I}^{t-1}] \approx p[y(t)|h, \hat{x}(t-1|t-1), P_x(t-1|t-1)]. \quad (16)$$

Once the state is estimated and the mode probabilities have been updated, the active joint mode is inferred as

$$\hat{s}(t) = \arg\min_{h\in\mathcal{S}} \alpha[h](t|t). \quad (17)$$

This implies that the active mode is chosen as the most probable mode among all the possible combinations of the appliances' operating conditions. The power consumed by each appliance can then be reconstructed from the estimate $\hat{x}[\hat{s}(t)](t|t)$.

Due to the recursive nature of the GPB$_1$ algorithm, the method is suited for real-time disaggregation.

Remark 1. To perform disaggregation, we need to know the prior probabilities $\alpha[h](0|0)$. if they are unknown a possible choice is to set them all equal to $\frac{1}{|\mathcal{S}|}$. A possible choice for the initial state estimate $\hat{x}(0|0)$ is a vector of zeros of proper dimensions, while the initial covariance $P_x(0|0)$ should be selected based on the confidence we have on the initial state estimate. ■

5. EXAMPLE

The performance of the disaggregation method presented in Section 4 are evaluated using the AMPDs dataset Makonin et al. (2016), that consists of the power readings of 19 devices collected over one year in a household in Canada. Initially, we estimate the models for 5 appliances (clothes dryer, dishwasher, fridge, heat pump and basement plugs lights) and then we reconstruct their consumption patterns form aggregate power readings.

5.1 Estimating the appliances' models

The state-space models for the 5 appliances are obtained using data collected over one week ($\bar{T} = 10080$ samples). For the datasets $\{\mathcal{M}_i\}_{i=1}^5$ to capture all the operating conditions of the appliances we consider different weeks for different devices. Specifically, we use the data collected from day 74 to day 80 for the clothes dryer (CDE), from day 170 to day 176 for the dishwasher (DWE) and from day 23 to day 29 for the fridge (FGE), the heat pump (HPE) and the basement plugs lights (BME). The number of possible operating conditions $\{K_i\}_{i=1}^5$ for the devices is chosen through cross-validation. As a result of this procedure, we obtain $K_1 = K_2 = 3$ for CDE and DWE and $K_3 = K_4 = K_5 = 2$ for the other devices.

We estimate jump models with static sub-models for each of the appliances by applying the learning method summarized in Section 3, with $\lambda_i(j)$ set to 0.9, for all $j = 1, \ldots, K_i$ and $i = 1, \ldots, 5$. The initial mode sequences required to identify the model for the i-th appliance is

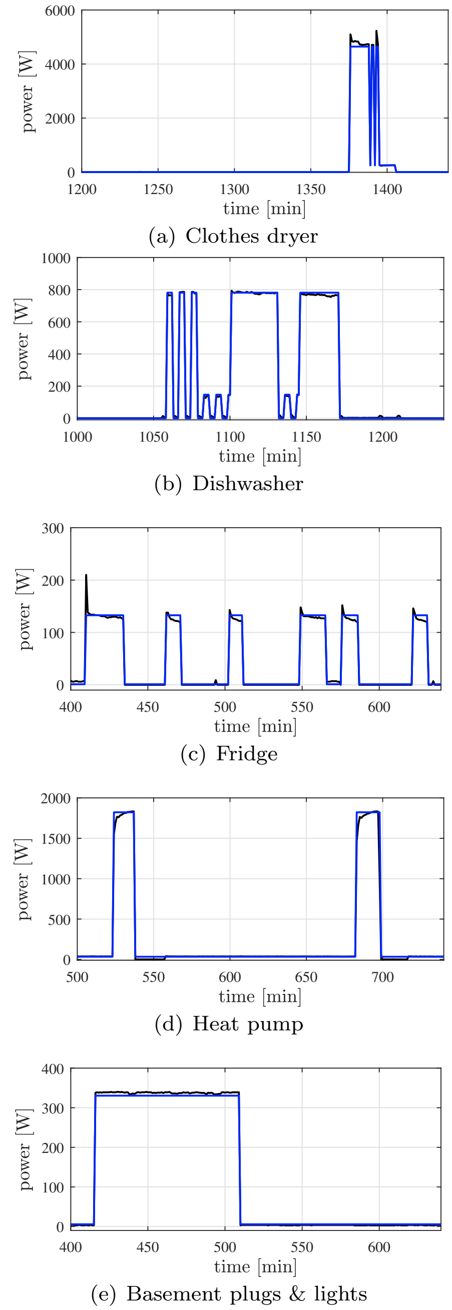
Table 1. Estimated parameters Θ_i.

Appliance	Θ_i low-power consumption	mid-power consumption	high-power consumption
CDE	0.4 W	251.8 W	4644.3 W
DWE	0.1 W	146.6 W	781.2 W
FGE	1.1 W	-	132.8 W
HPE	33.7 W	-	1822.1 W
BME	5.7 W	-	330.5 W

selected as the one minimizing the cost in (4) with respect to Θ_i among 10 randomly chosen sequences. To validate the model we consider the one-day readings collected at day 170 for CDE, day 45 for DWE, day 234 for FGE, day 80 for HPE and at day 300 for BME. The resulting parameters Θ_i and the reconstructed consumption profiles are reported in Table 1 and Fig. 1, respectively. For the sake of visualization we show only 240 samples of the one-day profiles reconstructed in validation. The quality of the identified models is quantitatively assessed through the *Best Fit Rate* (BFR)

$$BFR_i = 100 \cdot \left\{ 1 - \frac{\|y_i - \hat{y}_i\|}{\|y_i - \bar{y}_i\|}, 0 \right\} \%, \qquad (18)$$

with $\hat{y}_i$ being the output predicted by the model and $\bar{y}_i$ being the sample mean of the measured output. The resulting BFRs for each device are: 95.5% for the cloth dryer, 98.0 % for the dishwasher, 91.4 % for the fridge, 91.3 % for the heat pump and 97.1 % for basement plugs lights. Even though static sub-models allow us to accurately model the behavior of the 5 devices, the fridge and the heat pump present a transient that is not represented by this class of models. We thus identify second-order affine dynamic models for the high consumption states of the fridge and the heat pump, while setting the parameter characterizing the low-power consumption modes to 0 W. The dynamic models are identified minimizing the simulation error through particle swarm optimization , based on the active mode sequence estimated minimizing (4). The BFRs resulting from the substitution of the static sub-models with dynamic ones are 85.8 % and 87.3 % for FGE and HPE, respectively. Although the obtained BFRs are lower than the ones resulting static sub-models, the transient is better reconstructed introducing the dynamic models, as it can be noticed from the patterns reported in Fig. 2. The reduction of the BFRs is due to the choice of setting the low-power consumption level to 0 W.



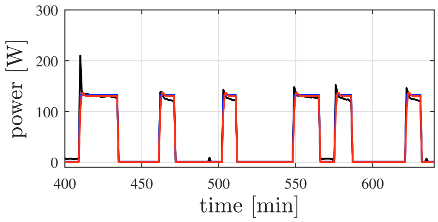
(a) Clothes dryer
(b) Dishwasher
(c) Fridge
(d) Heat pump
(e) Basement plugs & lights

Fig. 1. True (black) vs estimated (blue) consumption patterns.

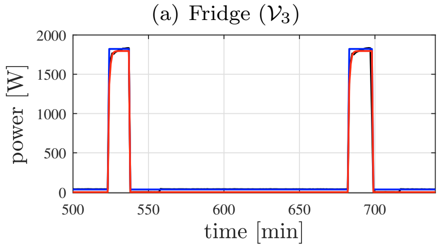
5.2 Testing the disaggregation approach

The disaggregation approach presented in Section 4 is tested against the AMPDs dataset Makonin et al. (2016). Specifically, we construct the aggregate power $y(t)$ summing up the readings of five appliances (clothes dryer, dishwasher, fridge, heat pump and basement plugs lights) collected over 12 consecutive days ($T = 17280$ samples). To assess the robustness of the disaggregation method over unmodelled appliances, we add the readings of the bedroom, the garage and the dining room collected over the same 12 days to the aggregate power. Furthermore, $y(t)$ is corrupted by a fictitious zero-mean Gaussian noise with standard deviation equal to 4 W.

The covariance matrices $Q[h]$ and $R[h]$ characterizing the noises in (12) are diagonal matrices with non-zero entries equal to 10 and 800, respectively. The initial state estimate $\hat{x}(0|0)$ is a zero vector of proper dimensions and the associated covariance $P_x(0|0)$ is a diagonal matrix with non-zero entries equal to 1000. We assume a uniform distribution for the initial mode $s(0)$ and, thus, we set $\alpha[h](0|0) = \frac{1}{|\mathcal{S}|}$. By supposing that the appliances are likely to seldom change their operating mode, we remove one-sample spikes in the reconstructed patterns through a post-processing phase. When disaggregation is performed, the single appliances profiles available in the AMPDs dataset are only used as ground-truth when assessing the performance of the method. The indexes used to quantitatively assess the performance of the proposed disaggregation method are the following.



(a) Fridge ($\mathcal{V}_3$)



(b) Heat pump ($\mathcal{V}_4$)

Fig. 2. True (black) vs estimated consumption patterns with static (blue) and dynamic (red) sub-models.

1. The F-score F_s is used to evaluate the capabilities of the approach in reconstructing the on/off states of the appliances. In particular, the F-score for the i-th appliance is calculated as

$$F_{s_i} = 2 \cdot \frac{TP_i}{2TP_i + FP_i + FN_i}, \tag{19}$$

where TP_i, FP_i and FN_i are the number of true positives (on events that are correctly classified), false positives (off events ranked as on) and false negatives (on events classified as off), respectively. As the proposed method is not limited to the identification of the appliances' states, but it allows to reconstruct the consumption pattern of each device, we classify the appliances as off if the reconstructed reading $\hat{y}_i(t)$ is smaller or equal than a threshold ε_i. Specifically, ε_i is set to 10 W for all the devices except the heat pump, for which the threshold is equal to 50 W. This choice is consistent with the low-power level associated to this appliance, equal to 33.7 W.

2. The *Estimated Energy Fraction Index* (EEFI), computed as

$$\text{EEFI}_i = \frac{\sum_{t=1}^{T} \hat{y}_i(t, \Theta_i, \hat{s}_i(t))}{\sum_{i=1}^{N} \sum_{t=1}^{T} \hat{y}_i(t, \Theta_i, \hat{s}_i(t))}, \tag{20}$$

indicates the fraction of energy assigned to the i-th appliances on the basis of the reconstructed patterns. This index is compared with the *Actual Energy Fraction Index* (AEFI), given by

$$\text{AEFI}_i = \frac{\sum_{t=1}^{T} y_i(t)}{\sum_{i=1}^{N} \sum_{t=1}^{T} y_i(t)}, \tag{21}$$

that provides the actual fraction of energy consumed by the i-th appliance.

3. The *Relative Square Error* (RSE) and the R^2 coefficient measure the quality of the reconstructed consumption patterns and are computed as

$$\text{RSE}_i = \frac{\sum_{t=1}^{T} (y_i(t) - \hat{y}_i(t, \Theta_i, \hat{s}_i(t)))}{\sum_{t=1}^{T} y_i^2(t)}, \tag{22}$$

$$R_i^2 = 1 - \frac{\sum_{t=1}^{T} (y_i(t) - \hat{y}_i(t, \Theta_i, \hat{s}_i(t)))^2}{\sum_{t=1}^{T} (y_i(t) - \bar{y}_i)}, \tag{23}$$

respectively.

Table 2. F-score.

	Static	Dynamic
Clothes dryer	90.1%	90.5%
Dishwasher	90.4%	90.8%
Fridge	76.3%	76.1%
Heat Pump	81.0%	81.6%
Basement	97.3%	98.1%

Table 3. Indexes AEFI vs EEFI.

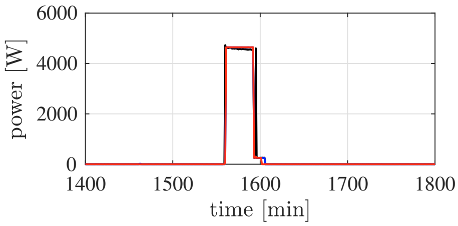
	Static		Dynamic	
	AEFI$_i$	EEFI$_i$	AEFI$_i$	EEFI$_i$
Clothes dryer	12.0%	13.5%	12.0%	13.4%
Dishwasher	2.2%	3.3%	2.2%	2.9%
Fridge	8.0%	9.2%	8.0%	9.1%
Heat Pump	66.3%	73.9%	66.3%	63.8%
Basement	11.5%	11.2%	11.5%	10.5%

Table 4. RSE and R^2 coefficient.

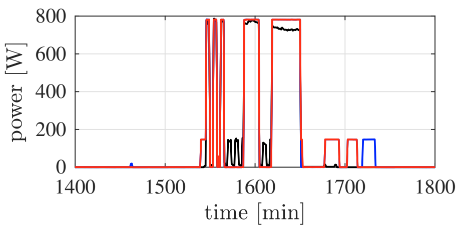
	Static		Dynamic	
	RSE$_i$	R_i^2	RSE$_i$	R_i^2
Clothes dryer	9.6%	90.2%	3.5%	96.5%
Dishwasher	8.4%	91.5%	8.2%	91.6%
Fridge	31.3%	55.5%	31.7%	55.0%
Heat Pump	4.3%	94.5%	0.6%	99.2%
Basement	17.8%	77.4%	15.5%	80.3%

The F-scores obtained using static sub-models for all the appliances and the ones resulting from the description of the behavior of the fridge and the heat pump through dynamic sub-models are reported in Table 2. These results show that a slight improvement in the classification of on/off events can be generally achieved replacing the static sub-models for the fridge and the heat pump with dynamic ones. The lower F-score is obtained for the fridge, independently on the considered sub-models. Such result can be related to the similarities between the mid-power state of the dishwasher (146.6 W) and the high-power mode of the fridge (132.8 W). The AEFI and EEFI for each of the 5 considered devices are compared in Table 3, showing that the introduction of the dynamic models for the fridge and the heat pump leads to an increase in the accuracy of the reconstructed fraction of energy of the appliances. The advantages introduced by replacing the static sub-models for the fridge and the heat pump with dynamic ones can be further noticed looking at the RSE and R^2 coefficients reported in Table 4. The improvement is particularly significant for the indexes associated with the clothes dryer and the heat pump, while the RSE and R^2 coefficients obtained for the fridge seems to indicate that its reconstructed consumption pattern is not accurate. Nonetheless, the estimated profiles reported in Fig. 3[2] show that the approach allows us to accurately estimate the power consumed by the fridge most of the time. We can thus conclude that the proposed approach allows to accurately reconstruct the patterns of the single appliances from aggregated power readings, even if corrupted by the power consumed by unmodelled appliances and measurement noise. Furthermore, the introduction of dynamic sub-models for the appliances showing a significant transient behavior leads to an improvement in the quality of the disaggregated profiles.

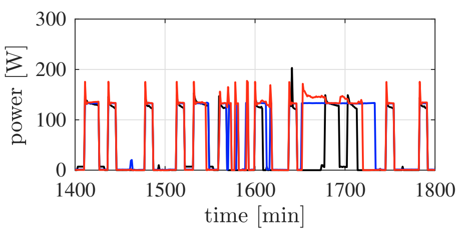
[2] For the sake of visualization only 400 samples are reported.



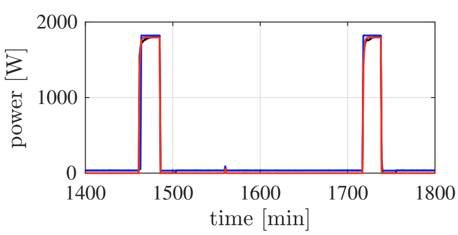
(a) Clothes dryer



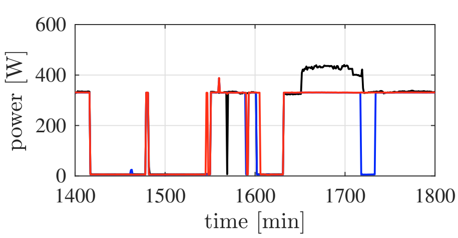
(b) Dishwasher



(c) Fridge



(d) Heat pump



(e) Basement plugs & lights

Fig. 3. True (black) vs estimated power demands obtained using static (blue) or dynamic sub-models (red) for the fridge and the heat pump.

The average CPU time required to perform energy disaggregation at a given instant is 1.1 ms and 1.2 ms when using static and dynamic sub-modes for the fridge and the heat pump, respectively. This slight difference is due to the increased dimension of the state when dynamic sub-models are used. Nonetheless, the obtained CPU times show that the proposed approach is suited for real-time disaggregation.

6. CONCLUSIONS

In this paper we have proposed an approach for energy disaggregation based on multi-model Kalman filters. The presented approach allows us to iteratively process the data and it has been proven to be quite robust to noises on the aggregate readings. The proposed method requires to model the behavior of the single appliances, which is estimated with the jump model identification method recently proposed in Bemporad et al. (2017).

Future research will be devoted to the integration of the identification method into the disaggregation procedure with the aim of reducing the intrusiveness of the method. We will also study strategies to reduce the number of Kalman filters to be run at each instant while being able to follow changes in the appliances' behavior over time.

ACKNOWLEDGEMENT

This work was supported in part by the European Commission under the Project H2020-723248 *DAEDALUS - Distributed control and simulation platform to support an ecosystem of digital automation developer*.

REFERENCES

Bar-Shalom, Y., Kirubarajan, T., and Li, X.R. (2002). *Estimation with Applications to Tracking and Navigation.* John Wiley & Sons, Inc., New York, NY, USA.

Bemporad, A., Breschi, V., Piga, D., and Boyd, S. (2017). Fitting jump models. `https://arxiv.org/abs/1711.09220`.

Bonfigli, R., Principi, E., Fagiani, M., Severini, M., Squartini, S., and Piazza, F. (2017). Non-intrusive load monitoring by using active and reactive power in additive Factorial Hidden Markov Models. *Applied Energy*, 208, 1590–1607.

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

Cominola, A., Giuliani, M., Piga, D., Castelletti, A., and Rizzoli, A. (2017). A hybrid signature-based iterative disaggregation algorithm for non-intrusive load monitoring. *Applied Energy*, 185(P1), 331–344.

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

Kolter, J.Z. and Jaakkola, T. (2012). Approximate inference in additive factorial hmms with application to energy disaggregation. In *Proc. of the 15th Int. Conf. on Artificial Intelligence and Statistics*, 1472–1482.

Makonin, S., Bradley, E., Bajic, I., and Popowich, F. (2016). Electricity, water, and natural gas consumption of a residential house in canada from 2012 to 2014. *Scientific Data*, 3(160037), 1–12.

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

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

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