

# Modeling and Analysis of Residential Flexibility: Timing of White Good Usage

N. Sadeghianpourhamami<sup>a,\*</sup>, T. Demeester<sup>a</sup>, D. F. Benoit<sup>b</sup>, M. Strobbe<sup>a</sup>, C. Develder<sup>a,\*\*</sup>

<sup>a</sup>*Ghent University – iMinds, Dept. of Information Technology – IBCN,  
Technologiepark Zwijnaarde 15, BE-9052 Ghent, Belgium*

<sup>b</sup>*Ghent University, Faculty of Economics and Business Administration,  
Tweekerkenstraat 2, BE-9000 Ghent, Belgium*

---

## Abstract

Challenges that smart grids aim to address include the increasing fraction of supply by renewable energy sources, as well as plain rise of demand, e.g., by increased electrification of transportation. Part of the solution to these challenges lies in exploiting the opportunity to steer residential electricity consumption (e.g., for flattening the peak load or balancing the supply and demand in presence of the renewable energy production). To optimally exploit this opportunity, it is crucial to have insights on how flexible the residential demand is. Load flexibility is characterized by the amount of power, time of availability and duration of deferrable consumption. Residential flexibility however, is challenging to exploit due to the variation in types of customer loads and differences in appliance usage habits from one household to the other. Existing analyses of individual customer flexibility behavior in terms of timing are often based on inferences from surveys or customer load patterns (e.g., as observed through smart meter data): there is a high level of uncertainty about customer habits in offering the flexibility. Even though some of these studies rely on real world data, only few of them have quantitative data on actual flexible appliance usage, and none of them characterizes individual user behavior. In this paper, we address this gap and contribute with: (1) a new quantitative specification of flexibility, (2) two systematic methodologies for modeling individual customer behavior, (3) evaluation of the proposed models in terms of how accurately the data they generate corresponds with real world customer behavior, and (4) a basic analysis of factors influencing the flexibility behavior based on statistical tests. Experimental results for (2)–(4) are based on a unique dataset from a real-life field trial.

*Keywords:* Generative model, Flexibility Characterization, Smart Grid

---

## 1. Introduction

The rapid integration of renewable energy sources into the power grid and their intermittent nature has created a need for flexibility in energy demand. Flexibility is generally regarded as the amount of load that is shiftable over various time scales and is quantized by 3 parameters [1]: (1) the *amount* of deferrable energy (i.e., the amount of energy that can be delayed without jeopardizing customer convenience or quality of the task to be fulfilled by a smart device), (2) the *time* of availability (i.e., the time at which a customer offers the device flexibility for exploitation), and (3) the *deadline* to exploit the offered flexibility (i.e., the maximum allowable delay for the energy consumption). Once flexibility is known and thus adequately characterized, it can be utilized by demand response (DR) algorithms to coordinate the demand-supply balance in the

---

\*Corresponding author

\*\*Principal corresponding author

*Email addresses:* [nasrin.sadeghian@intec.ugent.be](mailto:nasrin.sadeghian@intec.ugent.be) (N. Sadeghianpourhamami), [thomas.demeester@intec.ugent.be](mailto:thomas.demeester@intec.ugent.be) (T. Demeester), [dries.benoit@ugent.be](mailto:dries.benoit@ugent.be) (D. F. Benoit), [matthias.strobbe@intec.ugent.be](mailto:matthias.strobbe@intec.ugent.be) (M. Strobbe), [chris.develder@intec.ugent.be](mailto:chris.develder@intec.ugent.be) (C. Develder)

network. Various DR algorithms have already been proposed to exploit such flexibility: for an overview, we refer to [2] and [3]. Hence, proposing a new DR algorithm is not our focus.

Our main objective is to characterize and model the flexibility as DR's main asset, to improve the efficiency of DR assessment. One of the main challenges in the widespread deployment of DR algorithms (especially in the residential sector) is the uncertainty surrounding their impact [4, 5]. A poor understanding of flexibility characteristics as DR's main asset leads to inefficient DR assessment and uncertain conclusions (i.e., accurate evaluation of DR algorithms is impossible without in-depth analysis of the flexibility parameters). The outcome of our flexibility modeling and characterization (which is based on a unique dataset from a real-life field trial) can foster more realistic assessment of the potential impact of DR algorithms and pave the way to their realization in smart grid.

Flexibility of large industrial customers has already been extensively assessed and exploited by long standing programs (e.g., [6, 7, 8, 9, 10]). Since the inception of the smart grid, that industrial flexibility has increasingly been complemented by residential flexibility. Residential customers form a promising source of flexibility due to their widespread distribution and substantial share of electricity market and hence are the focus of our study here.

Residential flexibility however, is challenging to characterize due to the large variety of appliances and their diverse consumption patterns, as well as the uncertainty associated with appliance usage due to different usage habits among various households. A substantial amount of research has analyzed the flexibility potential of residential customers from various perspectives. A brief overview is presented in the next section.

### 1.1. State of the Art in Residential Flexibility Assessment

Methods to assess residential flexibility potential in literature can be categorized into two main streams, according to the objective they pursue: *DR-based* and *DR-independent* methods. The *DR-based* methods are often tailored to the underlying DR scheme (i.e., price-based or incentive-based DR) and their main objective is to model the responsiveness of customers to price signals or incentive programs. In price-based DR schemes, an elasticity matrix models customer flexibility as changes in aggregated demand in response to price changes [11, 12, 13, 14]. However, an elasticity matrix can only measure the aggregated flexibility potential and not the appliance specific flexibility. Price-demand models based on mixed integer linear programming (e.g., [15]) or probabilistic models (e.g., [11, 16]) are proposed to predict the customer consumption patterns from the appliance level perspective in response to dynamic prices. For incentive-based DR schemes, Hu *et al.* [17] propose a stochastic model to assess the probability distribution of residential demand in response to certain incentives. The proposed residential responsive demand model is formulated with consideration of the customer portfolio and household characteristics obtained from time-of-use surveys, rather than actual measurement of real behavior.

One of the limitations of DR-based approaches is that any quantization and assessment of flexibility potential is inevitably influenced by the underlying DR algorithm. Additionally, the impact of the underlying DR algorithm on the flexibility is not measurable. In other words, customers might exhibit different flexibility behavior when assessed with other DR-based methods. Hence, the outcome of the analysis of a particular DR-based method cannot be employed to reliably assess the impact of different DR algorithms. Instead, DR-independent methods (including the modeling approaches in this paper), offer an unbiased analysis where the customer flexibility behavior is not influenced by the specifics of any DR algorithm.

In *DR-independent* methods, the main objective is to model customer flexibility potential (independent of the underlying DR scheme) and subsequently use the model to assess the potential impact of DR algorithms on peak load reduction or demand-supply balancing. Some of these methods are derived merely based on appliance energy usage patterns that are either obtained from sub-metering of household appliances [18] or assumed by studying the characteristics of the various appliances [19]. Analyzing the flexibility potential based on appliance energy usage patterns provides insights about the potential amount of deferrable energy of each appliance. However, it does not completely characterize the flexibility potential because customer behavior, affecting the time of availability and deadline to exploit the offered flexibility, is not accounted for. One of the popular means to take into account customer appliance usage habits in the flexibility

model is collecting time-of-use surveys. Laicane *et al.* [20] performed a time-of-use survey on a four-person household to determine its appliance usage behaviors, particularly for washing machine and dishwasher, to quantify the flexibility potential. The model was then used to shift appliance usage accordingly for peak load reduction. Safdarian *et al.* [21] used a similar approach on 1600 Finnish households to assess the benefits of demand response on the operation of distribution networks. However, time-of-use surveys may be inaccurate in modeling the customer appliance usage habits because they indicate the self-reported behavior of the customer, which may differ from the real behavior.

Another approach taken by DR-independent methods is to obtain a time series estimate of flexibility of residential customers based on the clustering of their load profiles. Kouzelis *et al.* [22] proposed a methodology for analyzing the flexibility potential of residential heat pumps in a probabilistic way from the aggregated load profile of the customers. The proposed methodology compares the load profile of the flexible customer with electrically similar non-flexible customers by means of clustering the customer load profiles and then statistically infers the flexibility potential thereof. Labeeuw *et al.* [23] also used clustering of customer load profiles to derive a time series estimation of load curves and determine demand reduction potential of wet appliances in terms of amount of deferrable load only (without assessing the flexibility duration). They additionally incorporated attitude measurements based on questionnaires in their studies to account for customers willingness to participate in DR based on survey data. Despite valuable contributions of these approaches in terms of amount of deferrable energy and time of availability, they do not give any assessment of the deadline to exploit the flexibility due to limitations in their measurements.

In both of the aforementioned DR-based and DR-independent categories, modeling of customer responsiveness to participate in DR algorithms is not based on real-world scenarios where households are provided with smart appliances and required to configure their appliances flexibly. Hence, the uncertainty about limitations of DR algorithms due to the differences in customers' real-life (power consumption) habits remains largely unresolved. To address this gap, Kobus *et al.* [24] conducted a longitudinal study for one year over 77 Dutch households. Each household was given a smart washing machine, and an energy management system that received daily dynamic prices. The customers' behavioral changes with respect to a reference group was then studied for a full year to explore the potential role of smart appliances in shifting real electricity demand of smart washing machines in response to dynamic tariffs. Still, a limitation of this valuable work is that the analysis is tailored specifically to the underlying DR scheme.

D'Hulst *et al.* [25] also have based their analysis on a real world scenario where customers are provided with a platform to operate their smart devices and offer their flexibility for DR exploitation. The probabilistic analysis of time of configuration as well as the duration of provided flexibility is performed for various appliances. Their analysis is based on the LINEAR pilot project [26] in which 15 minute based measurements were collected from 418 appliances in 186 Belgian households. The analysis in [25] was independent of the underlying DR scheme that harnessed the flexibility potential, hence inferring a better estimation. However, the analysis of the flexibility potential is performed on an aggregated level and does not model individual household flexibility.

### 1.2. Motivation and Contributions

The residential flexibility potential has been extensively analyzed in terms of the amount of shiftable energy. However, the assessment of customer appliance usage habits in terms of when and for how long the flexible device is available for DR exploitation, is usually based on survey data and this may limit the analysis accuracy. Although D'Hulst *et al.* [25] provide a probabilistic analysis of the time of the configuration and deadline to exploit the flexibility, the analysis is at the aggregated appliance level. Also, [25] is limited to analysis only: they do not offer a generative model (that could be used, e.g., to generate data that is representative of reality, for simulation of DR). Such a generative model of (individual) customer flexibility based on real world data is missing in the literature. We have addressed this gap by modeling the residential flexibility in terms of configuration time and deadline, for individual households and appliances.

The primary motivation to develop a *generative model* of customer flexibility behavior, derived from a real dataset, is that such a model sharpens the definition of flexibility and fosters a more realistic analysis of residential flexibility potential by tackling not only the appliance load patterns, but also the uncertainties due to their owner's willingness to offer the available flexibility (for exploitation by DR algorithms). Additionally,

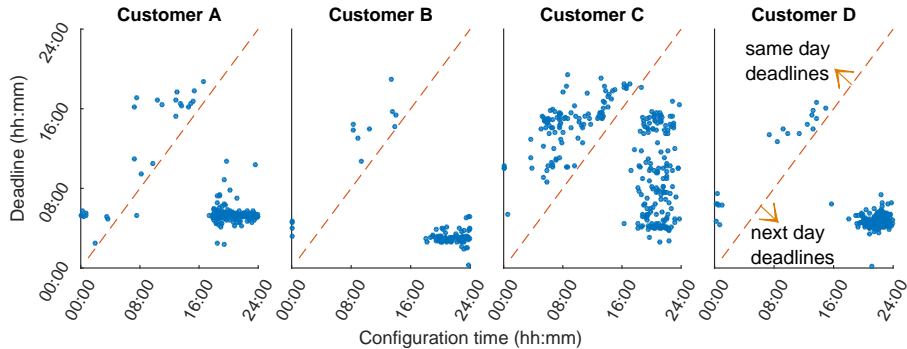


Fig. 1: Flexibility profiles of four selected customers for dishwasher usage (times on  $x$ - and  $y$ -axis are time of day). Note that the observations below the diagonal line denote deadlines on the day after the configuration time.

a parametric representation of customers allows for the comparison and selection of relevant customers for energy efficiency programs or DR participation, without jeopardizing their privacy. Finally, parametric models can be utilized for synthetic data generation, hence eliminating the need for time-of-use surveys and resulting in more realistic assessment of the DR algorithms' efficiency in harnessing the residential flexibility for various objectives.

As a secondary objective, our modeling allows for examination of factors that might influence the customer flexibility behavior. Identification of the *underlying factors* influencing the customer flexibility behavior allows for more realistic assumptions about potential flexibility. It also may help to make more accurate predictions of customer flexibility behavior.

Our analysis is based on real-world smart appliance usage (washing machines, dishwashers and tumble dryers) as collected in the LINEAR pilot project [26] and offers these contributions:

1. We take a new perspective and instead of indirectly inferring flexibility from load profiles or time-of-use surveys, we use direct measurement of flexible configuration times and deadlines to characterize flexibility (Section 2.1),
2. We propose two systematic methodologies to derive generative models for customer flexibility behavior from data (Sections 2.2 and 2.3),
3. We test the accuracy of the proposed models in representing the real world data using sampling and hypothesis tests (Section 3.3) and
4. We analyze factors influencing the flexibility behavior using reliable statistical tests (Section 3.4).

To the best of our knowledge, we are the first to propose a generative model of individual customer flexibility behavior towards a particular appliance. Note that this paper is a substantial extension of our earlier work (which proposed a two-stage model) [1], since we now present (a) a new and more general, single-stage model for the flexibility behavior, and (b) more extensive analysis. The latter includes assessing both the accuracy of the proposed algorithms in regenerating the data (based on statistical tests), and the factors influencing the flexibility behavior of customers. We also apply our models to customer data from washing machine and tumble dryers in addition to dishwashers.

## 2. Methodology

In this section, we first explain the input features used to model customer flexibility behavior. We then introduce two systematic methods to model the flexibility behavior of an individual residential household towards a particular smart appliance. Note that our first model is proposed specifically for modeling customer behavior that exhibits clear deadlines. We observed this behavior in dishwasher usage, an appliance that has been attributed a high demand response potential [21, 27, 28]. Our second model is a more general approach, which we found to be suitable for modeling all types of studied white goods (i.e., washing machine, tumble dryer and dishwasher).

### 2.1. Quantitative Specification of Flexibility

Each time a customer configures a smart appliance, (s)he sets the flexibility duration, i.e., how long the start of the device’s operation can be deferred. We therefore represent the flexibility using two parameters: (1) *Configuration time*, which is the time at which the smart appliance is configured flexibly, and (2) *Deadline*, which is the latest allowed start time of the appliance and is calculated by adding the recorded flexibility duration to the time of configuration. The flexibility profile is defined for a given customer and a particular appliance, and is a model for his/her usage of that appliance. Note that we do not parameterize flexibility by the amount of shiftable power, since this attribute has already been analyzed extensively in the literature (e.g., [28]). The flexibility data for each customer and each appliance is obtained from year-long measurements in the LINEAR pilot project [26] (see further, Section 3). Figure 1 depicts flexibility profiles of four selected customers for their dishwasher. Each point on the graph of Fig. 1 represents a single usage, with the configuration time on the  $x$ -axis and the corresponding deadline on the  $y$ -axis.

### 2.2. Model I: Two-Stage Single Variate Approach

Customer flexibility profiles for dishwashers in Fig. 1 suggest that (unlike washing machines and tumble dryers, as we’ll see later) dishwashers are usually configured with a typical deadline. For example, Customer A has 3 typical deadlines around 5 am, 10 am and 4 pm. Configuration times are different for different deadlines. For example, most of the configurations with deadline at around 5 am are set in the late evening and span until a few hours after midnight, whereas the configurations with 4 pm deadlines usually occur during daytime from 8 am onwards. Given this typical customer behavior for dishwashers, our first model proposes a two-stage univariate approach. In the first stage, we identify the typical deadlines of the customer by using a clustering algorithm and estimate the probability  $P(\text{deadline})$  proportional to the size of the cluster for the respective deadline. For each cluster of deadlines resulting from Stage I, we then use parametric probability distributions to model the corresponding configuration times in Stage II and obtain  $P(\text{configuration time} \mid \text{deadline})$ . The joint distribution of deadlines and configuration times can then be obtained by  $P(\text{configuration time}, \text{deadline}) = P(\text{deadline}) \cdot P(\text{configuration time} \mid \text{deadline})$ .

#### 2.2.1. Stage I: Identification of Typical Deadlines

The main objective of this stage is to use the deadline feature ( $y$ -axis in Fig. 1) as the input to a clustering algorithm to identify the typical deadline clusters and parameterize their distributions. For this purpose, we adopt the G-means clustering algorithm [29] and change its hypothesis test (i.e., the Anderson-Darling test of normality). For completeness, the G-means algorithm and our modifications are explained below.

A wide variety of algorithms have been proposed for clustering of load profiles (e.g., see [30]). Some of the popular ones are  $\kappa$ -means, expectation maximization, fuzzy  $\kappa$ -means, hierarchical clustering, and self-organizing maps. An overview of these algorithms and their performance comparison is presented in [31]. G-means is a clustering algorithm based on  $\kappa$ -means, with the capability of dynamically determining the number of clusters using hypothesis tests, hence eliminating the challenging task of choosing the optimal number of components in  $\kappa$ -means. G-means is an iterative approach that starts with an initial value of  $k$ , i.e., the number of clusters (starting with  $k = 1$  in absence of prior information). It then decides to split each cluster into two new clusters if the data in the current cluster does not pass the Anderson-Darling test of normality. After each iteration, the  $\kappa$ -means algorithm is executed on the entire data to refine the solution. The core assumption in G-means is that the data in each cluster is normally distributed. Since deadlines are strictly positive, we instead assumed a Gamma distribution for the data in each cluster. Hence, we replaced the Anderson-Darling test of normality with a one-tailed Kolmogorov-Smirnov ( $k$ -s) test at significance level of 1%. We refer to this adaptation as  $\Gamma$ -means.

#### 2.2.2. Stage II: Parameterizing the Distribution of Configuration Times

In this stage, we employ probability distributions to model the corresponding configuration times for each deadline cluster resulting from Stage I. Qualitative exploration of tuning data revealed that the empirical distributions of configuration times are often multi-modal, skewed, and with heavy tails. These characteristics suggested the use of finite mixture models (FMM) as parametric models to represent the unknown

distributions as a mixture of known distributions. In what follows, we present the general definition of FMMs.

Let  $X$  be an  $r$ -dimensional random variable with probability density function  $f(\mathbf{x})$ , defined in the sample space  $\mathcal{X} \subset \mathbb{R}^r$  and arising from a  $K$ -component finite mixture distribution. The probability density of  $X$  for all  $\mathbf{x} \in \mathcal{X}$  is then defined as

$$f(\mathbf{x}) = \sum_{k=1}^K \eta_k f_k(\mathbf{x}), \quad (1)$$

where  $f_k(\mathbf{x})$  is the component-wise probability density function and  $\eta_k$  is the corresponding weight of the component.  $\boldsymbol{\eta} = (\eta_1, \dots, \eta_K)$  is called the weight distribution.  $\boldsymbol{\eta}$  takes a value in the unit simplex  $\varepsilon_K$  which is a subspace of  $(\mathbb{R}^+)^K$  defined by the following constraints:

$$\eta_k \geq 0, \quad \eta_1 + \eta_2 + \dots + \eta_K = 1. \quad (2)$$

Assuming that all the component densities arise from the same distribution family,  $\mathcal{T}(\boldsymbol{\theta})$ , the mixture density function is then written as,

$$f(\mathbf{x}|\boldsymbol{\vartheta}) = \sum_{k=1}^K \eta_k f_k(\mathbf{x}|\boldsymbol{\theta}_k), \quad (3)$$

where  $f_k(\mathbf{x}|\boldsymbol{\theta}_k)$  is the probability density function of the  $k^{\text{th}}$  component indexed by the parameter  $\boldsymbol{\theta}_k$  and  $\boldsymbol{\vartheta} = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_K, \boldsymbol{\eta})$  is the parameter vector of the mixture model.

To estimate the parameter vector  $\boldsymbol{\vartheta}$ , we employ a Bayesian approach based on data augmentation and Markov-chain Monte Carlo (MCMC) proposed by Schnatter [32]. The main difference between MCMC and the classical expectation maximization (EM) algorithm based on maximum likelihood (ML) estimation is that MCMC performs integration while EM does maximization. One of the key advantages of a Bayesian approach over ML is the direct availability of confidence regions while for EM these may be inaccurate for small data sizes because of the reliance on asymptotic approximations. Also, the Bayesian approach allows including prior information in the estimation procedure.

The data augmentation and the MCMC algorithm by Schnatter [32] are briefly explained here for completeness. The main objective of the algorithm is estimating the component parameters,  $\boldsymbol{\vartheta}$ , as well as the allocation vector  $\mathbf{S} = (S_1, S_2, \dots, S_N)$  denoting the allocation of each observation to its corresponding component in the mixture ( $S_i = k$  if  $x_i$  belongs to cluster  $k$ ). This problem is interpreted as a missing data problem. The augmented parameter  $(\boldsymbol{\vartheta}, \mathbf{S})$  is estimated by sampling from the complete-data posterior distribution based on the Gibbs MCMC algorithm. Schnatter's MCMC algorithm is outlined in Algorithm 1.

First, the algorithm uses the  $K$ -means clustering method to assign each observation to one of the  $K$  mixture components and initializes the allocation vector  $\mathbf{S}$  accordingly (Line 1). The algorithm then runs for  $M + M_0$  iterations (Lines 2–6). In each iteration, the component weights are sampled from a Dirichlet distribution (Line 3) and the parameters of each component of the mixture are sampled from the complete-data posterior, given allocation vector and the observations (Lines 4–5). Once the samples of the weights and the component parameters are drawn, the probability of belonging to a particular component is calculated for each observation and each observation is then allocated to a component based on the calculated probabilities (Line 6). At the end of the algorithm, the first  $M_0$  draws are disregarded (Line 7) and the remaining  $M$  draws for  $(\boldsymbol{\vartheta}, \mathbf{S})$  are returned (Line 8).

The mixture parameters are identifiable at best up to an arbitrary permutation mainly because the component number of the mixture is not a meaningful concept. This phenomenon results in the label switching problem due to the invariance of the mixture likelihood function under relabeling the components of a mixture model [32]. To solve this problem, [32] has proposed to apply random permutations of labels on the draws of their MCMC algorithm to force balanced label switching and ensure that the sampler explores the full mixture posterior distribution.

To choose the optimum number of components, we employ marginal likelihood denoted as  $P(\mathbf{X}|\mu_k)$  where  $\mu_k$  is the  $k$ -component FMM model. The marginal likelihood measures the average fit of a model

---

**Algorithm 1:** Schnatter’s MCMC algorithm

---

**Input** : data points ( $\mathbf{X}$ , with  $|\mathbf{X}| = N$ ); number of components ( $K$ )  
**Output:** Parameter vector of the mixture model ( $\boldsymbol{\vartheta}$ ); allocation vector ( $\mathbf{S}$ ) denoting to which component each of the observations are assigned (if  $x_i$  is part of cluster  $k$ , then  $S_i = k$ ).

- 1 Initialize  $\mathbf{S}$  (based on the K-means algorithm);
- 2 repeat  $M + M_0$  times
  - /\* Given the allocation vector  $\mathbf{S}$ , estimate the mixture parameters \*/
  - 3 Sample  $\boldsymbol{\eta} = (\eta_1, \dots, \eta_K)$  from the Dirichlet distribution  $D(e_1(\mathbf{S}), \dots, e_K(\mathbf{S}))$ , where  $e_k(\mathbf{S}) = e_0 + N_k(\mathbf{S})$ ,  $k = 1, \dots, K$ , and  $N_k(\mathbf{S})$  is the number of data points allocated to component  $k$  of the mixture and  $e_0$  is the prior of the Dirichlet distribution;
  - 4 foreach  $k = 1, \dots, K$  do
    - 5 | Sample the component parameter  $\boldsymbol{\theta}_k$  from the complete-data posterior  $p(\boldsymbol{\theta}_k | \mathbf{S}, x)$ ;
    - /\* Update the allocation vector  $\mathbf{S}$  based on the sampled parameter vector  $\boldsymbol{\vartheta}$  \*/
    - 6 | Classification of each observation  $x_i$  conditional on knowing  $\boldsymbol{\vartheta}$ , by sampling  $S_i$  independently for each  $i = 1, \dots, N$  from the following discrete distribution:  $p(S_i = k | \boldsymbol{\vartheta}, x_i) \propto p(x_i | \boldsymbol{\theta}_k) \eta_k$ ;
- 7 Disregard the first  $M_0$  draws;
- 8 return  $M$  draws for  $(\boldsymbol{\vartheta}, \mathbf{S})$ ;

---

to the data, whereas likelihood based or point estimators such as Akaike information criteria (AIC) [33] or Bayesian information criteria (BIC) [34] base their decisions on the best fit of each competing model [35]. The use of the marginal likelihood also automatically penalizes the complexity of the model because complex models spread their probability mass widely by predicting various possible outcomes, hence the probability of actual data will be smaller for overly complex models.

### 2.3. Model II: Single Stage Bivariate Approach

The proposed two-stage univariate model is relatively simple since one-dimensional data is used for modeling in each stage. Yet, it is therefore not possible to model arbitrary interactions between configuration times and deadlines: Model I might not be very efficient when the customer does not make clear deadline choices (i.e., the flexibility profile does *not* amount to groups of data points with a similar  $y$ -axis value, which represents the deadline, such as Customer C in Fig. 1 or customer flexibility profiles for washing machine and tumble dryer in Fig. 7 and Fig. 8 respectively). Indeed, for such cases, the clustering in Stage I of Model I will produce multiple clusters of deadlines, each one in parallel with the  $x$ -axis, but this might not be the best representation of the underlying data. Therefore, we propose a second methodology to model the customer flexibility profile using a bivariate Gaussian mixture model (GMM).

In this scheme, a bivariate GMM is fit to the flexibility profile of the customer: both of the flexibility features (i.e., deadline and configuration time) form the input to the model, and their joint probability distribution is estimated in a single step. The same methodology based on Bayesian MCMC (explained in Section 2.2.2) is employed to fit the bivariate mixture model and select the optimal number of components.

## 3. Results and Discussion

We have applied our methodologies on the data from year-long measurements in the LINEAR pilot project [26] for 3 types of white goods: dishwasher, washing machine and tumble dryer. The flexibility data associated with each appliance is obtained from smart meters. Measurements are taken every 15 min: a day is divided into 96 time slots and any customer configuration within the 15 min interval is reported at the end of the interval, implying that the resulting customer flexibility profile is discrete. To make our data continuous, we spread the measured data over the 15 min preceding the interval end times, by adding random noise from a uniform distribution.



Fig. 2: Clusters of deadlines from Stage I in the customer flexibility profiles for dishwasher. Each cluster is shown in a different color.

For each appliance, we considered customers with at least 100 configurations to increase the reliability of the analysis, leading to a set of 15 test customers for dishwashers, 12 for washing machines and 8 for tumble dryers. It is noteworthy that the proposed methodologies are implemented in MATLAB.

The results are presented in four subsections. In Sections 3.1 and 3.2, we model the distribution of the flexibility profiles based on the test data using the proposed methodologies and present the resulting clusters. In Section 3.3, we evaluate the efficiency of the derived models in reproducing the customer flexibility profile when using them to generate synthetic data. We then apply  $\chi^2$  and Fisher's exact tests in Section 3.4 to examine the dependency of cluster membership on day-of-the-week, weekends, holidays and seasons.

### 3.1. Model I Parameter Fitting and Analysis of Resulting Clusters

Model I is suitable for flexibility profiles that have typical deadlines in their configurations (i.e., the dishwasher profiles in our dataset). It is a two-stage univariate approach in which each stage models one of the two flexibility features (i.e., the configuration time or the deadline). Hence, Model I allows to analyze each of these flexibility features separately, as presented in the following subsections.

#### 3.1.1. Stage I: Analysis of Deadline Clusters

Figure 2 shows the output at Stage I of the model, where  $\Gamma$ -means clustering is applied to deadlines ( $y$ -axis), for the test customers. The corresponding configuration times ( $x$ -axis) are also shown in Fig. 2 for further analysis and comparison of differences among the configuration times of various clusters. An exploratory analysis of the outcome of Stage I on test customers reveals the following characteristics of the customers' flexibility profile for their dishwashers: (i) there are typically 2 to 3 deadline clusters, early mornings (around 4-5 am), late morning/noon (10-12 pm), and evening (around 5-6 pm), but the number of deadline clusters varies among customers; (ii) for most customers, morning deadlines are more frequent and the corresponding configuration times usually lie in the afternoon and late night. Such customers show a more deterministic behavior compared to others: for example, Customer 1 has a substantial amount of data in the cluster corresponding to the early morning deadline.

Fig. 3 shows the percentage of data in each deadline cluster of the test customers. Note that data are not evenly distributed across the clusters: for 7 out of 15 customers, one cluster contains more than 90% of the



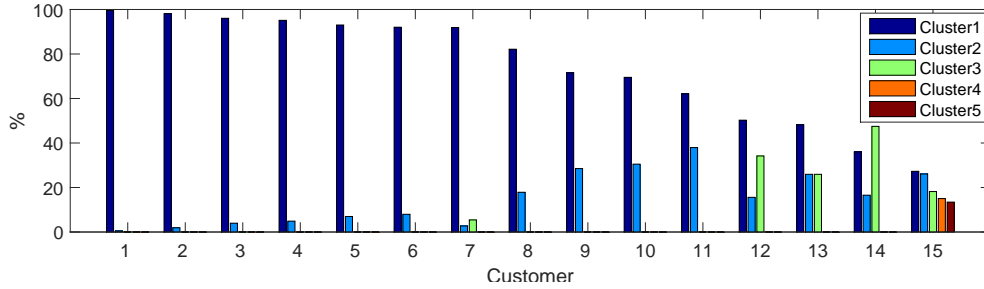


Fig. 3: Percentage of data in each deadline cluster for dishwasher. (Note that a given cluster index for one customer may relate to a different deadline than the same index for another customer.)

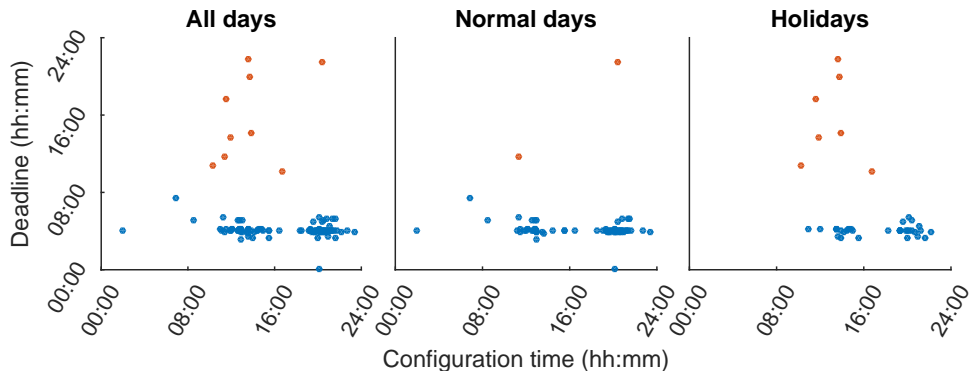


Fig. 4: Example of a customer flexible configurations for dishwasher (Customer 6) during holidays vs. normal days over the cluster of deadlines.

data (for most customers, one cluster has more than 50% of the data). This cluster represents the dominant habit of the customer while other clusters reflect his exceptional behavior. Fig. 4 depicts an exemplary customer whose flexible configurations are affected by holidays. As seen from Fig. 4, the data in the smaller cluster, representing afternoon and evening deadlines, are usually configured during holidays, although a substantial amount of configurations with early morning deadlines are still present during holidays.

However, these conclusions drawn from Fig. 4 should not be extrapolated to all the customers, as we will show in Section 3.4.

### 3.1.2. Stage II: Analysis of Distribution of Configuration Times

In this stage, the parametric model is developed to represent the distribution of the corresponding configuration times of each of the resulting clusters in the previous stage. Looking back at Fig. 2, we observe that for some clusters (e.g., those with early morning deadlines) the data in the lower left corner of the figure is related to that in the right. It is sensible to say that configurations shortly after midnight are indeed the tail of the ones at late evening and hence they are likely to belong to the same distribution. To account for this, we changed the reference point from midnight to the middle of the largest gap in the configuration time of each cluster (typically around 5-6 am). To model the configuration time distribution, we first focused on the larger clusters (with more than 100 data points) to ensure reliability of our conclusions and then applied the model to smaller clusters. In the following paragraph, we explain the three steps involved in choosing and fitting a parametric distribution to test clusters.

*Step 1 (choosing the right parametric distribution):* We initially fit all valid parametric distributions to the configuration times of test clusters and compared the fit using goodness-of-fit tests (e.g., BIC, AIC, log-likelihood) as well as hypothesis tests (one tailed k-s test). From this analysis, we concluded that there is not a unique single-component parametric distribution that can represent the data for all of the test clusters. Also, for some clusters, none of the parametric distributions provides an acceptable fit. However,

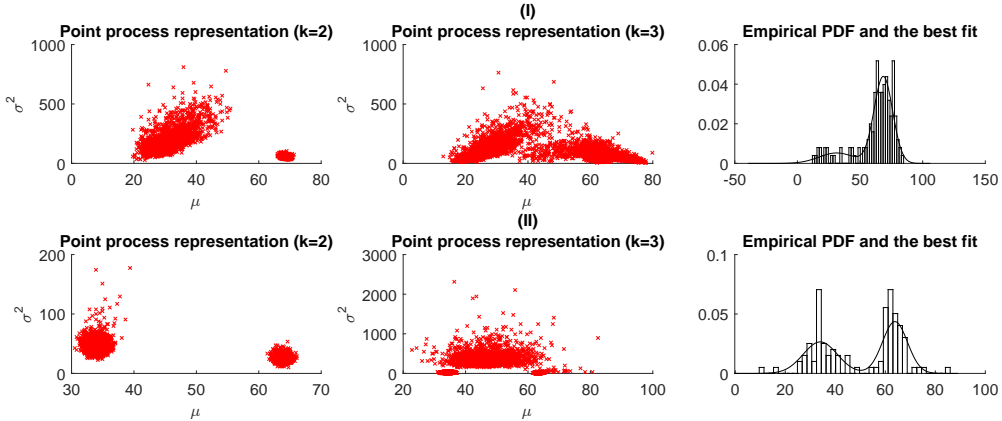


Fig. 5: Point process representation of posterior draws and PDF of the best fit for two randomly selected customers for dishwasher.

Table 1: Log-marginal Likelihood,  $p(\mu_k|\mathbf{X})$  (larger values are better).

Cluster	$k = 1$	$k = 2$	$k = 3$	$k = 4$
<b>I</b>	-553.95	-545.34	-595.17	-548.91
<b>II</b>	-595.53	-554.89	-554.84	-559.87

for those clusters that could be represented by a single-component parametric distribution, the non-central student distribution was usually the best fit. Further, the characteristics of the empirical distribution for the majority of the test clusters suggested the use of FMMs for parametric modeling. Based on the initial observation that suggested the non-central student distribution as a suitable fit, and the fact that a non-central student distribution is approximated by a Gaussian distribution for a large enough sample, we fit and compare the FMM from two families of distributions: Gaussian mixtures and student mixtures. To compare the fit of the two models, we used marginal likelihood values.

Comparing marginal likelihood values reveals that the Gaussian mixtures are a more suitable model to represent the data for all the test clusters. Hence Gaussian mixtures are employed for modeling and further analysis of the distribution of configuration times of deadline clusters.

*Step 2 (choosing the number of components):* To select the optimal number of components of the mixture, we employ the log-marginal likelihood together with a point process representation of the posterior draws to avoid overfitting. The point process representation is a viewpoint introduced by [36] which represents every component of the mixture in terms of its parameters. Next, we explain the overall procedure as well as the point process representations with the aid of an example.

Table 1 shows the log-marginal likelihood for fitting Gaussian mixtures with different values for  $k$ , i.e., the number of components of FMMs for test clusters. According to Table 1, the largest log-marginal likelihood corresponds to  $k = 2$  for Cluster I and  $k = 3$  for Cluster II. The corresponding point process representation of posterior draws for  $k = 2$  and 3, as well as the probability density function (PDF) of the best fit is shown in Fig. 5. According to Fig. 5(I), the point process representations of posterior draws for Cluster I form well-separated clusters around the parameters of the mixture components for  $k = 2$  but increasing the number of components indicates overfitting because well-separated clusters corresponding to parameters of each component are no longer seen in the point process representation plot. Hence,  $k = 2$  is the optimum choice as suggested by log-marginal likelihood values. However, for Cluster II, the  $k = 3$  suggested by Table 1 overfits the data according to the point process representations shown in Fig. 5(II). Therefore,  $k = 2$  is selected as the optimum number of mixture components for Cluster II. It is noteworthy that in Fig. 5, the  $x$ -axis is changed from time to timeslot of a day (i.e., a value between 0 and 95) to ease the comparison between point process representation and PDF of the best fits. Experimenting with test clusters suggests

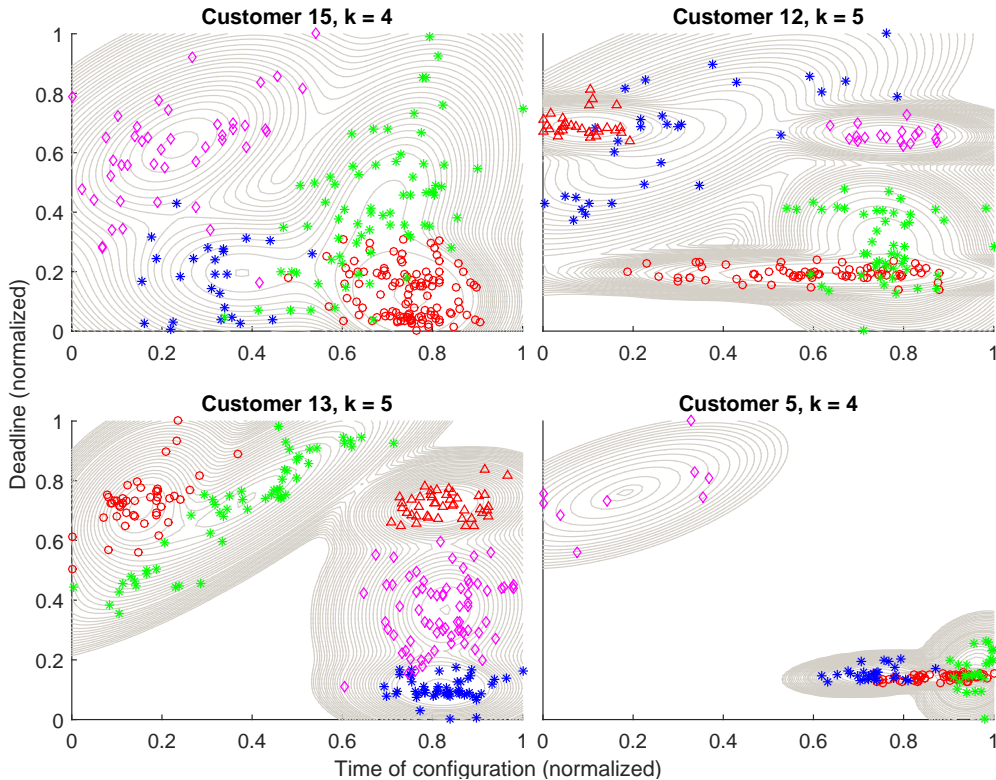


Fig. 6: Bivariate Gaussian mixture fits and the resulting MAP clusters for *dishwasher* (in distinct symbols and colors; contour plots indicate the fitted distributions).

that when the difference in log-marginal likelihood of  $k$  and  $k + 1$  components is smaller than 1, the smaller number of components should be chosen to avoid overfitting.

*Step 3 (determining the mixture component membership)*: Any data that is represented best with FMMs, can be clustered based on the mixture component membership and hence be analyzed to identify the potential factors influencing the component membership. We use the maximum-a-posteriori (MAP) clustering algorithm to identify, with a certain probability, the component membership of each data point. MAP identifies the allocation vector that yields the largest posterior probability. The optimum allocation vector is easily obtained by keeping track of the allocation vectors and corresponding posterior probability at each iteration of Schnatter’s algorithm (explained in Section 2.2.2).

Application of the outlined steps to deadline clusters of the test customers reveals that: (i) the majority of the clusters are best represented with a 2-component Gaussian mixture model (and the optimum number of components never exceeds 3), and (ii) the component weights are not uniformly distributed in the majority of the clusters, meaning that there is typically one dominant cluster.

### 3.2. Model II Parameter Fitting and Analysis of Resulting Clusters

In this section, we fit bivariate Gaussian mixture models to our test data. Similar to Stage 2 of Model I, we change the reference point to the middle of the largest gap seen in the configuration times ( $x$ -axis) of the customer’s data, in order to ensure an acceptable continuity of the distribution of configuration times over midnight. The summary of the estimated model parameters are tabulated in the Appendix.

The results of fitting Model II to flexibility profiles for dishwashers are depicted in Fig. 6 for 4 example customers. The contour plots in Fig. 6 indicate the bivariate Gaussian mixture fits. We have also employed MAP clustering to assign each point to a single component (as explained in Subsection 3.1.2, Step 3). The resulting MAP clusters are shown using distinct colors and symbols in Fig. 6. Unlike clusters in Model I, the

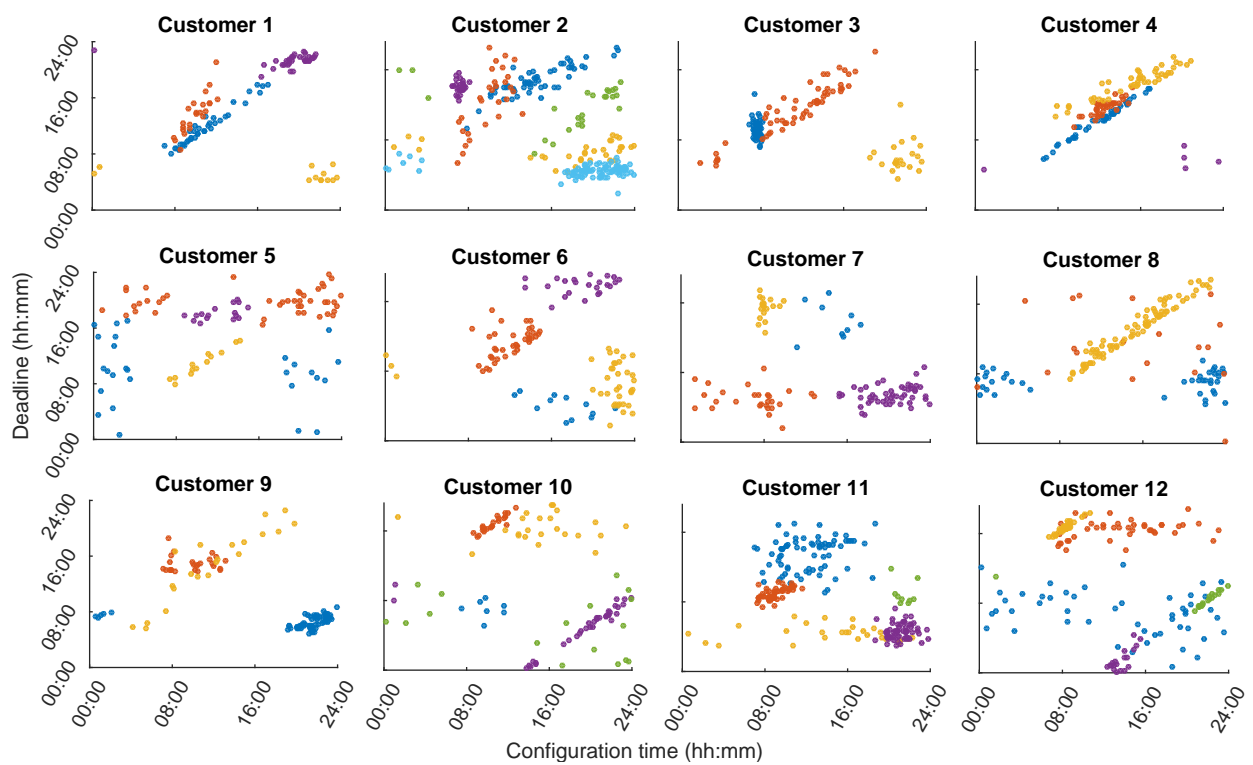


Fig. 7: Bivariate Gaussian mixture fits and the resulting MAP clusters for *washing machine* flexible usage (in distinct symbols and colors; contour plots omitted for clarity).

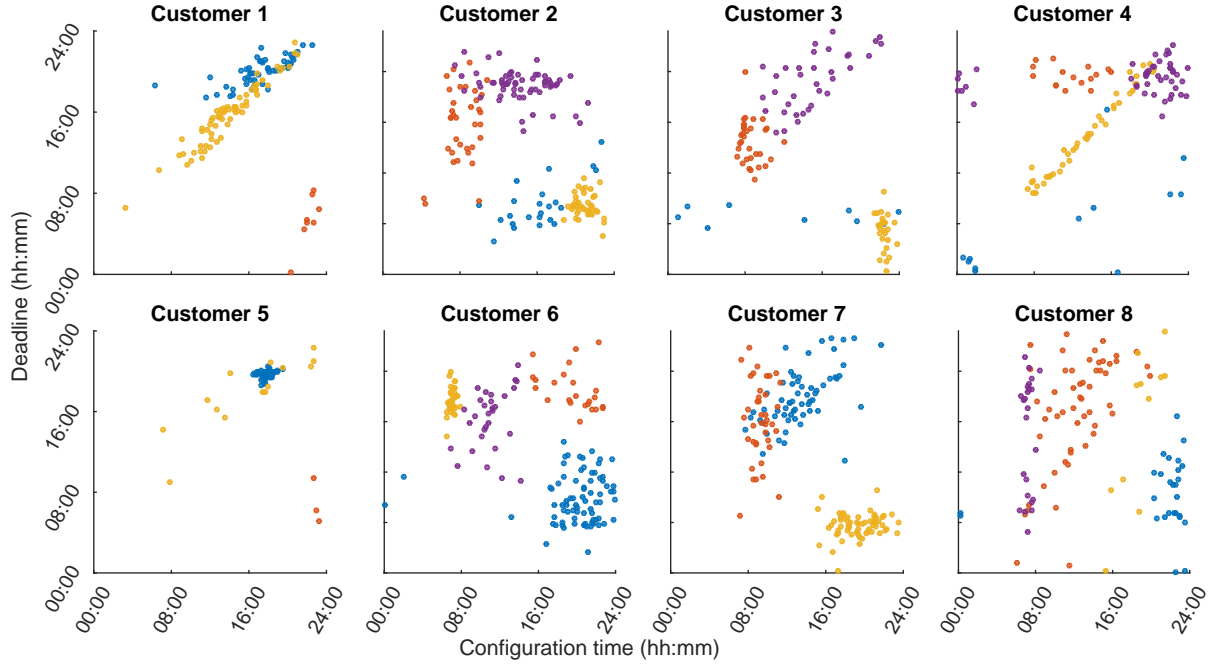


Fig. 8: Bivariate Gaussian mixture fits and the resulting MAP clusters for *tumble dryer* flexible usage (in distinct colors; contour plots omitted for clarity).

clusters in Model II are not always aligned in parallel with the  $x$ -axis (i.e., do not have similar deadlines). Two categories of clusters are observed in Model II: (i) clusters with similar deadlines (i.e., horizontal ones), and (ii) those in which the customers configure the device with similar flexibility duration (i.e., clusters in parallel with the  $x = y$  diagonal line). This shows the versatility of Model II in accommodating various types of customer behaviors (i.e., configurations with similar deadlines or similar flexibility durations).

We also applied Model II on the customer flexibility profiles for washing machines and tumble dryers (refer to the tables in Appendix for a summary of the model parameters). The resulting clusters are shown in Fig. 7 and Fig. 8 for washing machine and tumble dryer respectively. The aforementioned categories of clusters are also observed in the flexibility profiles for washing machines and tumble dryers. As seen from Fig. 7 and Fig. 8, for the majority of customers using a washing machine or tumble dryer, their configurations with similar deadlines (early morning) usually are in the late evening and span to few hours after the midnight. The configurations throughout the day all have a similar flexibility duration (resulting in clusters roughly parallel to the  $x = y$  diagonal). In addition, there are clusters with similar configuration times but different deadlines (i.e., vertical ones) for some customer flexibility profiles for washing machine and tumble dryers (e.g., Customers 2, 3, 6, 7 and 8 in Fig. 8 and Customers 2, 3 and 7 in Fig. 7). Such vertical clusters are not observed in customer flexibility profiles for dishwashers.

### 3.3. Analysis of Model Efficiency for Data Regeneration

In this section we analyze the efficiency of the models in regenerating the customers flexibility profile. To achieve this objective, we first model a customer using our proposed methods. Each method yields a probabilistic model of the customer's data. We note that in Stage I of Model I, we model the data in each deadline cluster using a gamma distribution. Hence, we represent the overall data as mixture of gamma distributions where the parameters of each component of the mixture are equal to the underlying gamma distribution in each deadline cluster. The weight of each component in the mixture is calculated by taking the ratio of the corresponding deadline cluster size to the overall data size. We then generate 500 sample sets of the same size as the original data (from which the models are derived). Each sample set is compared to the original data using a two tailed k-s test to determine whether the sample and the original data come

Table 2: Analyses of model efficiency in regenerating the customer flexibility profiles: percentage of times the null hypothesis is not rejected. (Note that the same customer numbers across devices do not necessarily correspond to the same users.)

Dishwasher			Washing Machine		Tumble Dryer		
Customer	Model I	Model II	Customer	Model II	Customer	Model II	
1	0.0 %	90.4 %	1	100 %	1	100	%
2	83.2 %	95.2 %	2	100 %	2	100	%
3	0.2 %	99.2 %	3	100 %	3	100	%
4	7.8 %	99.2 %	4	98 %	4	91	%
5	31.4 %	89.8 %	5	100 %	5	100	%
6	0.0 %	97.2 %	6	100 %	6	100	%
7	0.0 %	61.2 %	7	100 %	7	100	%
8	0.6 %	99.8 %	8	100 %	8	100	%
9	48.0 %	95.4 %	9	100 %			
10	99.6 %	99.6 %	10	100 %			
11	77.2 %	100.0 %	11	100 %			
12	99.2 %	99.8 %	12	100 %			
13	99.6 %	99.8 %					
14	98.2 %	100.0 %					
15	98.2 %	99.0 %					

from the same distribution. The null hypothesis is that data and sample come from the same distribution and the alternative hypothesis is that they do not. It should also be noted that since the original k-s test is developed for one-dimensional data, we have applied the extension of the k-s test to bivariate data, as proposed by [37].

Table 2 shows the percentages of times the test fails to reject the null hypothesis at 5% significance level. As seen from Table 2, for most of the customer flexibility profiles (of all 3 categories of white goods) more than 90% of the samples generated by Model II are not rejected by the hypothesis test and hence, may be representative samples.

For dishwashers, Model II outperforms or matches Model I in re-generating the data, although for some customers they perform equally well. Yet, for some customers, Model I never generates a sample that can be considered to be representative of the original data. Investigating these cases, analysis reveals that although data in each cluster follows a Gamma distribution (due to  $\Gamma$ -means clustering), for customers with very deterministic behavior (i.e., very narrow deadline windows in each deadline cluster), representing hard clusters of deadlines with a mixture of Gamma distributions does not regenerate a sample representative of deadline distributions. We conclude this from a one dimensional two tailed k-s test by considering deadline feature of customer flexibility data: the null hypothesis (i.e., data arising from the mixture of Gamma distribution) is always rejected for customers with very narrow spread in their deadlines.

Despite the fact that Model I often does not represent the data as successfully as Model II, it still has the benefit of simplicity, both in constructing the model and in interpreting it (e.g., clear identification of typical deadlines, which is characteristic of several dishwasher flexibility profiles).

### 3.4. Dependency Analysis Using Statistical Tests

To investigate the dependency of resulting clusters of each model on potential determining factors, including weekends, day-of-a-week, holidays and seasons, we employ  $\chi^2$  and Fisher’s exact tests. Fisher’s exact test is only employed when the conditions of the  $\chi^2$  test are not met, mainly due to small data sizes in some cells of the contingency table. Additionally, Fisher’s exact test is only applicable for  $2 \times 2$  contingency tables: for larger tables, extensions of Fisher’s test are employed (e.g., Freeman-Halton’s extension [38] for  $2 \times 3$ ). We define our null and alternative hypothesis for a ‘factor’ as follows:

- *Null hypothesis*: cluster membership and ‘factor’ are independent,
- *Alternative hypothesis*: cluster membership and ‘factor’ are dependent.

Table 3:  $p$ -Values of dependency tests in Stage I of Model I for dishwasher usage. (Bold values indicate a  $p$ -value less than 0.05 and shaded rows indicate customers affected by at least one ‘factor’.)

(a) Holidays, weekends and seasons

Customer	No. of clusters	Test type	$p$ -Value (holiday)	$p$ -Value (week-end)	$p$ -Value (spring)	$p$ -Value (summer)	$p$ -Value (autumn)	$p$ -Value (winter)
1	2	Fisher	2.9E-01	1.0E+00	6.2E-01	8.8E-02	5.7E-01	5.1E-01
2	2	Fisher	5.4E-01	5.3E-01	4.6E-01	3.4E-01	4.0E-01	3.8E-01
3	2	Fisher	2.1E-01	1.0E+00	5.6E-01	1.7E-01	5.8E-01	2.6E-01
4	2	Fisher	5.4E-01	1.0E+00	6.5E-01	5.5E-01	1.0E-01	<b>2.8E-02</b>
5	2	Fisher	1.0E+00	1.2E-01	<b>6.2E-03</b>	5.3E-01	<b>3.8E-02</b>	3.9E-01
6	2	Fisher	<b>1.1E-02</b>	4.4E-01	1.7E-01	7.6E-02	<b>3.7E-02</b>	5.9E-01
7	3	Fisher	5.2E-01	<b>5.2E-04</b>	7.8E-01	8.4E-02	<b>6.4E-03</b>	2.0E-01
8	2	$\chi^2$	8.8E-02	<b>1.2E-04</b>	<b>4.3E-02</b>	1.6E-01	9.0E-01	5.2E-01
9	2	$\chi^2$	3.9E-01	2.8E-01	6.32E-02	6.9E-01	2.5E-01	1.2E-01
10	2	$\chi^2$	<b>2.8E-03</b>	3.2E-01	8.5E-02	<b>5.0E-02</b>	<b>4.7E-03</b>	5.2E-01
11	2	$\chi^2$	9.5E-01	3.0E-01	1.03E-01	9.6E-01	1.4E-01	9.9E-01
12	3	$\chi^2$	3.5E-01	<b>1.8E-02</b>	9.8E-01	1.8E-01	<b>3.1E-04</b>	8.6E-02
13	3	$\chi^2$	<b>5.1E-03</b>	2.0E-01	<b>2.1E-04</b>	<b>2.2E-02</b>	<b>5.0E-03</b>	7.1E-02
14	3	$\chi^2$	3.7E-01	7.7E-02	3.1E-01	2.4E-01	2.0E-01	2.2E-01
15	5	$\chi^2$	7.7E-02	6.0E-02	9.8E-02	3.7E-01	<b>4.9E-02</b>	<b>2.6E-02</b>

(b) Day-of-the-week

Customer	Sun	Mon	Tue	Wed	Thur	Fri	Sat
1	1.0E+00	1.0E+00	1.6E-01	1.0E+00	1.0E+00	1.0E+00	1.0E+00
2	1.0E+00	1.0E+00	1.0E+00	2.3E-01	3.1E-01	1.0E+00	1.0E+00
3	3.6E-01	1.0E+00	6.3E-01	6.0E-01	6.1E-01	9.1E-02	2.3E-01
4	1.0E+00	7.0E-01	6.7E-01	3.7E-01	4.1E-01	3.9E-01	1.0E+00
5	6.9E-01	6.9E-01	3.5E-01	1.0E+00	3.6E-01	5.6E-01	<b>3.8E-02</b>
6	1.0E+00	5.9E-01	1.0E+00	3.5E-01	1.0E+00	9.8E-02	2.1E-01
7	7.9E-01	2.9E-01	1.6E-01	5.0E-01	6.5E-01	1.7E-01	<b>7.9E-05</b>
8	<b>8.8E-03</b>	5.3E-01	1.1E-01	<b>2.2E-02</b>	5.9E-01	7.2E-01	<b>2.2E-02</b>
9	1.7E-01	3.4E-01	8.9E-01	1.3E-01	7.9E-02	6.4E-01	9.5E-01
10	3.6E-01	1.3E-01	8.4E-01	4.7E-01	6.1E-01	1.5E-01	8.0E-02
11	8.6E-01	9.3E-01	9.9E-01	2.8E-01	6.9E-01	8.1E-01	2.5E-01
12	6.6E-01	4.1E-01	9.0E-01	6.8E-01	8.3E-01	6.5E-02	<b>4.8E-04</b>
13	9.5E-01	<b>2.7E-02</b>	6.2E-01	8.5E-01	<b>1.6E-02</b>	<b>6.2E-03</b>	<b>1.7E-02</b>
14	9.1E-02	3.0E-01	3.2E-01	1.1E-01	3.1E-01	4.8E-01	5.6E-01
15	3.6E-01	3.9E-01	2.0E-01	2.7E-01	<b>7.1E-04</b>	7.2E-02	1.1E-01

The dependency tests are conducted at the 5% significance level. In other words, for  $p$ -Values larger than 0.05, the test fails to reject the null hypothesis and hence does not conclude a dependency among the considered ‘factor’ and clusters under analysis. The number of columns in the contingency table corresponds to the number of deadline clusters for each customer. The contingency tables always consists of two rows. For example, to test the dependency of cluster membership on weekends, the first row of the corresponding table would contain count data for configurations on weekend days and the second row would contain the configurations on all the other days of the week. Similarly, for a day-of-the-week factor, the contingency table of each day would have count data for the particular day in one row and the configurations of the remaining six days of the week on another row.

The outcome of the tests (for dishwashers only) for clusters in Stage I and Stage II of Model I are summarized in Table 3 and Table 4 respectively.

Table 3(a) shows that weekends, holidays and seasons may influence the behavior of some customers in setting deadlines for their dishwasher. However, as mentioned earlier, the aforementioned factors do not affect all customers behaviors similarly. Additionally, autumn seems to have a more pronounced effect on customer deadlines than the other seasons. From Table 3(b), we note that 6 out of the 15 customers (shaded rows) have a dependency of dishwasher configuration times on at least one day of the week. While some customers remain unaffected by day-of-the-week, some others are greatly influenced by this factor (e.g., Customer 15, for whom the null hypothesis is rejected for 4 days of the week).

Table 4, shows that holidays and seasons do not significantly affect the component memberships in Stage II for the test deadline clusters from Stage I. However, weekends influence component membership for 5 out of the 10 test clusters. Additionally, 7 out of 10 test clusters have at least one day that influences the component membership of their data points.

Table 4:  $p$ -Values of dependency tests in Stage II of Model I for dishwasher usage. (Bold values indicate a  $p$ -value less than 0.05 and shaded rows indicate customers affected by at least one ‘factor’.)

(a) Holidays, weekends and seasons

Cluster	No. of clusters	Test type	$p$ -Value (holiday)	$p$ -Value (week-end)	$p$ -Value (Spring)	$p$ -Value (Summer)	$p$ -Value (Autumn)	$p$ -Value (Winter)
A	3	$\chi^2$	6.9E-01	3.1E-01	3.0E-01	2.5E-01	2.1E-01	3.7E-01
B	3	Fisher	2.1E-01	<b>2.5E-03</b>	4.6E-01	4.6E-01	<b>2.7E-02</b>	1.7E-01
C	3	$\chi^2$	4.9E-01	<b>5.2E-03</b>	1.0E+00	2.1E-01	8.4E-01	5.0E-01
D	3	$\chi^2$	4.0E-01	<b>3.4E-02</b>	4.5E-01	4.3E-01	1.4E-01	5.9E-01
E	2	$\chi^2$	6.8E-01	<b>7.6E-04</b>	2.2E-01	2.6E-01	6.0E-01	7.5E-02
F	2	$\chi^2$	4.2E-01	6.5E-01	1.2E-01	5.0E-01	5.6E-02	6.3E-01
G	2	$\chi^2$	3.7E-01	6.2E-02	9.5E-01	8.8E-01	3.4E-01	4.6E-01
H	3	Fisher	3.6E-01	4.7E-01	5.5E-01	6.4E-01	6.3E-01	4.4E-01
I	3	Fisher	1.7E-01	1.9E-01	<b>3.9E-03</b>	7.7E-02	<b>1.5E-02</b>	1.8E-01
J	3	$\chi^2$	<b>3.4E-06</b>	<b>8.7E-04</b>	7.0E-01	<b>2.7E-02</b>	5.5E-01	1.5E-01

(b) Day-of-the-week

Cluster	Sun	Mon	Tue	Wed	Thur	Fri	Sat
A	9.9E-01	3.3E-01	4.4E-01	1.2E-01	3.5E-01	<b>2.9E-02</b>	6.9E-02
B	<b>6.7E-03</b>	8.1E-01	8.6E-02	4.2E-01	7.2E-01	8.3E-01	3.8E-01
C	6.8E-02	9.2E-01	<b>1.2E-04</b>	9.6E-01	6.9E-01	7.2E-01	5.5E-02
D	2.4E-01	<b>3.3E-02</b>	6.3E-01	2.4E-01	<b>2.9E-02</b>	1.0E+00	1.1E-01
E	<b>7.8E-05</b>	6.9E-01	<b>3.5E-02</b>	1.6E-01	2.7E-01	9.4E-01	9.4E-01
F	2.6E-01	<b>3.0E-02</b>	7.8E-01	4.0E-01	3.9E-01	<b>3.8E-02</b>	3.5E-01
G	3.1E-01	3.0E-01	2.5E-01	9.6E-01	1.6E-01	2.8E-01	1.9E-01
H	2.4E-01	6.9E-01	5.2E-01	7.8E-01	2.8E-01	6.9E-01	3.1E-01
I	1.0E-01	5.3E-01	3.1E-01	9.2E-01	1.2E-01	7.0E-01	6.5E-01
J	<b>1.7E-04</b>	8.8E-01	<b>6.1E-03</b>	<b>2.3E-02</b>	2.3E-01	<b>2.9E-03</b>	<b>3.9E-05</b>

We have also applied the dependency tests to the clusters of Model II to identify factors influencing the cluster membership of a customer’s data points. The  $p$ -values of the dependency tests on dishwashers are summarized in Table 5. We constructed similar tables for customer flexibility profiles for washing machines and tumble dryers, however, we omit them from the paper to save space: below we summarize the main observations.

For *customer flexibility profiles for dishwashers*: from Table 5, 5 out of 15 customers’ cluster membership might be dependent on holidays, whereas more than 50% of the customers behavior is affected by weekends, seasons and day-of-the-week factors. Among seasons, winter least influences the cluster memberships. The results of Table 5 also indicate that while not all customers are affected in a similar manner, from an overall perspective, weekend, seasons and day-of-the-week factors influence more customers than the holiday factor.

For *customer flexibility profiles for washing machines*: only 3 out of 12 customers are influenced by holidays, whereas, weekends, seasons and day-of-the-week factors affect 50% of the customers. Among seasons, winter has the highest influence (6 out of 12) and autumn has the least (2 out of 12) influence on cluster membership.

For *customer flexibility profiles for tumble dryers*: none of the 8 test customers are affected by holiday, and only 1 customer is affected by seasons. On the other hand, weekend and day-of-the-week factors influence more than 50% of the customers.

The main conclusion drawn from the dependency tests above is that customers not only exhibit different behavior in offering the flexibility of their smart device, but also that they are affected differently by the aforementioned factors. Hence, the assumption that flexibility will be offered by all customers in a similar manner, is not supported by our data. It hence seems crucial to individually model customer behavior in order to incorporate the uncertainties due to their distinct habits, and thus allow for a more realistic assessment of DR impact.

## 4. Conclusion

Flexibility is characterized by (1) amount of deferrable load, (2) time of availability and (3) deadline to exploit. Existing works have extensively studied (1). However, (2) and (3) are challenging to analyze



Table 5:  $p$ -Values of dependency tests for Model II for dishwasher usage. (Bold values indicate a  $p$ -value less than 0.05 and shaded rows indicate customers affected by at least a ‘factor’.)

(a) Holidays, weekends and seasons

Customer	No. of clusters	Test type	$p$ -Value (holiday)	$p$ -Value (week-end)	$p$ -Value (spring)	$p$ -Value (summer)	$p$ -Value (autumn)	$p$ -Value (winter)
1	3	Fisher	3.8E-01	7.4E-01	5.1E-01	2.0E-01	8.3E-01	3.5E-01
2	4	$\chi^2$	6.4E-01	1.8E-01	<b>7.9E-03</b>	6.2E-01	2.9E-01	6.7E-01
3	5	$\chi^2$	<b>3.8E-07</b>	<b>8.1E-03</b>	3.4E-01	1.6E-01	8.6E-02	4.1E-01
4	4	$\chi^2$	<b>1.5E-03</b>	<b>1.7E-03</b>	7.0E-01	<b>4.0E-03</b>	1.1E-01	9.0E-02
5	4	$\chi^2$	8.9E-01	1.2E-01	<b>4.9E-02</b>	8.7E-01	1.6E-01	4.1E-01
6	5	$\chi^2$	<b>8.8E-03</b>	2.1E-01	<b>3.8E-02</b>	<b>1.3E-04</b>	7.2E-02	1.6E-01
7	4	$\chi^2$	7.0E-01	<b>2.0E-02</b>	1.6E-01	<b>2.8E-02</b>	<b>1.7E-04</b>	3.4E-01
8	4	$\chi^2$	1.6E-01	<b>1.2E-05</b>	3.7E-01	3.1E-01	3.7E-01	8.0E-01
9	5	$\chi^2$	5.1E-01	3.8E-01	2.6E-01	1.2E-01	<b>1.7E-02</b>	8.5E-02
10	4	$\chi^2$	8.2E-02	4.5E-01	<b>5.5E-03</b>	<b>4.8E-02</b>	<b>1.3E-02</b>	9.6E-02
11	4	$\chi^2$	5.5E-01	1.6E-01	7.7E-02	1.1E-01	6.5E-02	3.4E-01
12	5	$\chi^2$	<b>5.0E-02</b>	<b>3.3E-03</b>	8.0E-01	2.0E-01	<b>8.7E-04</b>	1.3E-01
13	5	$\chi^2$	<b>9.1E-05</b>	<b>8.5E-04</b>	<b>4.0E-05</b>	<b>2.8E-03</b>	<b>2.5E-03</b>	5.6E-02
14	5	$\chi^2$	6.7E-01	<b>5.7E-03</b>	3.5E-01	5.4E-01	1.2E-01	2.8E-01
15	4	$\chi^2$	1.2E-01	8.1E-02	8.6E-01	5.0E-01	3.0E-01	<b>4.4E-02</b>

(b) Day-of-the-week

Customer	Sun	Mon	Tue	Wed	Thur	Fri	Sat
1	8.6E-01	8.5E-01	<b>3.1E-02</b>	8.8E-01	<b>1.6E-02</b>	3.6E-01	4.4E-01
2	1.0E-01	4.7E-01	6.1E-01	2.3E-01	3.6E-01	6.9E-01	2.6E-01
3	<b>3.9E-02</b>	8.2E-01	7.8E-01	6.0E-01	2.3E-01	<b>2.4E-02</b>	1.0E-01
4	8.6E-02	5.0E-02	<b>1.4E-02</b>	7.4E-02	6.5E-01	<b>4.7E-04</b>	<b>6.7E-06</b>
5	3.2E-01	1.9E-01	2.3E-01	3.7E-01	<b>3.9E-02</b>	8.9E-02	<b>2.8E-02</b>
6	3.9E-01	7.7E-01	3.1E-01	1.4E-01	8.0E-01	5.5E-01	2.5E-01
7	9.2E-01	3.4E-01	2.2E-01	3.4E-01	6.3E-01	9.7E-02	<b>6.6E-03</b>
8	6.1E-02	8.2E-01	2.7E-01	<b>2.5E-02</b>	7.0E-01	6.0E-02	<b>9.2E-04</b>
9	2.9E-01	<b>3.2E-02</b>	8.7E-01	6.7E-01	5.3E-01	5.7E-01	9.9E-01
10	3.5E-01	2.9E-01	9.8E-01	9.9E-01	9.5E-01	2.4E-01	1.5E-01
11	5.2E-01	5.6E-01	4.4E-01	3.6E-01	7.6E-01	8.9E-01	3.0E-01
12	4.7E-01	8.0E-01	5.8E-01	8.9E-01	1.1E-01	5.8E-01	<b>8.3E-05</b>
13	<b>7.6E-05</b>	3.5E-01	<b>5.1E-03</b>	6.9E-02	5.5E-02	<b>7.0E-04</b>	<b>2.0E-04</b>
14	1.7E-01	7.3E-01	3.1E-01	3.8E-01	5.4E-01	4.9E-01	<b>1.6E-02</b>
15	1.1E-01	<b>8.2E-03</b>	1.1E-01	5.3E-01	9.2E-01	3.6E-01	5.4E-01

because they depend on customer appliance usage habits and may substantially differ among households. Hence, state-of-the-art lacks the quantitative analysis on (2) and (3): they are inferred from either the load patterns or time-of-use surveys which do not depict a realistic customer behavior (i.e., inferences are not based on real scenarios where customers configure their smart appliances flexibly). This distorts the realistic assessment of flexibility and hinders their efficient exploitation.

In this paper, we took the first step to tackle this issue and aimed to sharpen the analysis of flexibility by characterizing the individual customer flexibility profile in terms of time of configuration and deadline to exploit the offered flexibility. We proposed a systematic approach to derive a generative statistical model of an individual customer’s flexibility behavior in offering smart devices for DR exploitation, based on real-world data. Validating two proposed models with statistical tests, we found that especially Model II (bivariate Gaussian mixture models) was efficient in re-generating the customer flexibility profile. Finally, we evaluated the dependency of customer behavior on various factors using statistical tests, to conclude that not only do different users exhibit potentially substantially different flexibility behavior, but also that such flexibility could be influenced differently by factors such as day-of-the-week or seasons, etc. More details on the outcome of our contributions are explained below.

Our Model I is a two-stage univariate approach that is suitable for flexibility profiles that exhibit typical deadlines (e.g., dishwasher usages as observed in our real-world dataset). The first stage employs a  $\Gamma$ -means clustering algorithm to identify the typical deadlines and the second stage uses a Bayesian model and employs the MCMC algorithm to obtain the distribution of the corresponding configuration times for deadline clusters from Stage I. The analysis of customer flexibility profiles for dishwasher based on Model I reveals the following: (i) the majority of the customers tend to configure their smart dishwasher to meet a limited set of typical deadlines and the data in each cluster follows a Gamma distribution (based on k-s

hypothesis test), (ii) for most customers, the size of deadline clusters were unbalanced with one cluster representing the dominant habit of the customer, (iii) the distribution of configuration times in each cluster resulting from Stage I was best modeled by a mixture of Gaussian distributions (based on marginal likelihood criteria).

Model II is a more general approach that takes both flexibility features as the input and models their joint distribution in a single step. The proposed model fits bivariate Gaussian mixtures on the flexibility data of each customer using a Bayesian MCMC algorithm. MAP clustering is used to assign each observation to a single component. Three categories of MAP clusters are observed when modeling the customer flexibility profiles using Model II: (i) the ones with similar deadlines and different configuration times, (ii) the ones with similar configuration times but different deadlines and (iii) the ones with similar flexibility duration. Hence, this approach is more versatile in modeling various customer behaviors.

To validate the efficiency of Model I and Model II in regenerating the customer behavior for synthetic data generation purposes, we proposed a systematic approach based on a Kolmogorov-Smirnov (k-s) test. Based on our validations, Model II was identified to be an appropriate regenerative model for all 3 types of white goods in our analysis. The summary of Model II parameters fit to our real-world dataset are given in the Appendix.

Finally, to study factors influencing the customer flexibility behavior, we used  $\chi^2$  and Fisher’s exact tests and examined the effect of 4 factors: holidays, weekends, seasons and days-of-the-week. Our dependency analysis of real-world data suggests that customers are not similarly affected by aforementioned factors: modeling flexibility merely based on device characteristics and assuming that all customers will utilize them in the same manner does not reflect observed behavior. This further indicates the need to model customers individually in order to incorporate the uncertainties influencing their flexibility behavior due to their appliance usage habits.

#### 4.1. Future Work

The configuration time in the flexibility profile has a cyclic nature, hence, it is plausible that configurations shortly after midnight are the tails of distributions of the ones in the late evening. In order account for this, while still keeping the modeling in linear space, we adopted a transformation of the  $x$ -axis as shown in the figures throughout the paper, and changed the reference point to the middle of the largest gap in the configuration times of the customer for both Model I and Model II. However, it could be statistically meaningful (though more complex) to consider distributions of flexibility on a cylinder, with configuration times on a circle, while deadlines are on the vertical axis. Such mathematically more complex analysis is left for future work.

Finally, the comparison of test customers suggests similarities among some customers in terms of their flexibility behavior (e.g., similar deadline clusters). In future research, we will use clustering to group customers with similar flexibility profiles for each appliance. Clustering of similar customers may improve the exploitation of their flexibility by DR algorithms (e.g., by tailoring a specific DR algorithm for a group of similar customers).

## 5. Appendix

In this appendix, we summarize the estimated parameter values of Model II for the customer flexibility profiles for all three appliance types. We also give guidelines for generating synthetic data based on the summarized parameters. The notations used in the table headers are explained first:

- $[X_{min}, X_{max}]$ : the minimum and maximum values of the configuration time,
- $[Y_{min}, Y_{max}]$ : the minimum and maximum values of the deadline,
- Ref.: the largest gap in the configuration times of the customer (The data is shifted around this reference to ensure the continuity of the configuration times.)
- $\eta_k$ : the weight of  $k^{th}$  component of the mixture
- $\mu_k$ : the normalized mean of the  $k^{th}$  component of the mixture

- $\Sigma^k$ : the covariance matrix of the  $k^{th}$  component of the mixture. To compactly represent the covariance matrix in the tables, we note it as  $(\Sigma_{11}, \Sigma_{12}; \Sigma_{21}, \Sigma_{22})$  which corresponds to  $\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$ .

The following steps explain how to re-generate a customer flexibility profile:

- *Step 1*: sample from the  $K$ -component Gaussian mixture model based on the given parameters and weights for each customer.
- *Step 2*: the data was normalized prior to the modeling (i.e., using  $X_{normalized} = \frac{X - X_{min}}{X_{max} - X_{min}}$  for normalizing configuration times and  $Y_{normalized} = \frac{Y - Y_{min}}{Y_{max} - Y_{min}}$  for normalizing deadlines). Hence, de-normalize the generated samples accordingly.
- *Step 3*: shift the configuration times around the “Ref.” offset.

The aforementioned steps will regenerate a customer flexibility profile with deadlines ranging from [0 h, 24 h] and configuration times between [0, 96] (which represent indices of the 15 min timeslots). Note that since we had 15 min measurements to obtain the flexibility information, the configuration times are between [0, 96] in the re-generated sample and correspond to time-slot of the day. To obtain the configuration times in terms of hour of the day, multiply the values by 1/4.

Table 6: Summary of estimated parameters of Model II based on customer flexibility profiles for tumble dryers

Customer (Data Size)	$[X_{min}, X_{max}]$	$[Y_{min}, Y_{max}]$	Ref.	$\eta_k$	$\mu_k$	$\Sigma^k = (\Sigma_{11}, \Sigma_{12}; \Sigma_{21}, \Sigma_{22})$
1 (126)	[13.193, 92.846]	[0.134, 22.852]	0	2.97E-01	(5.87E-01, 1.78E-01)	(3.03E-03, -3.01E-04; -3.01E-04, 1.12E-03)
				5.64E-01	(6.83E-01, 1.77E-01)	(8.26E-03, -3.24E-05; -3.24E-05, 1.30E-04)
				1.38E-01	(4.16E-01, 5.38E-01)	(6.34E-02, -4.41E-02; -4.41E-02, 9.87E-02)
2 (163)	[17.166, 92.075]	[3.307, 22]	0	2.01E-01	(2.19E-01, 6.21E-01)	(8.00E-03, 4.91E-03; 4.91E-03, 2.92E-02)
				1.64E-01	(6.08E-01, 1.82E-01)	(3.91E-02, 5.37E-03; 5.37E-03, 1.21E-02)
				3.72E-01	(5.27E-01, 8.13E-01)	(3.02E-02, 4.77E-03; -4.77E-03, 7.28E-03)
3 (104)	[4.169, 91.279]	[0.291, 23.997]	21	2.63E-01	(8.82E-01, 1.91E-01)	(4.97E-03, 1.17E-04; 1.17E-04, 5.37E-03)
				1.93E-01	(4.33E-01, 4.45E-01)	(5.38E-02, -2.12E-02; -2.12E-02, 8.58E-02)
				3.14E-01	(3.31E-01, 1.83E-01)	(5.36E-03, -3.08E-05; -3.08E-05, 2.83E-05)
4 (118)	[10.361, 84.917]	[0.173, 22.587]	19	3.71E-01	(7.01E-01, 1.80E-01)	(4.14E-03, -7.09E-05; -7.09E-05, 1.57E-05)
				1.22E-01	(5.22E-01, 2.31E-01)	(5.03E-02, -2.40E-04; -2.40E-04, 1.88E-04)
				1.67E-01	(2.39E-01, 6.28E-01)	(2.09E-02, 1.09E-02; 1.09E-02, 2.39E-02)
5 (103)	[28.775, 92.834]	[5.157, 22.305]	0	2.78E-01	(6.98E-01, 4.23E-01)	(1.19E-02, 8.69E-03; 8.69E-03, 4.21E-02)
				1.69E-01	(3.66E-01, 1.95E-01)	(2.12E-02, 2.11E-05; 2.11E-05, 1.60E-02)
				3.86E-01	(7.55E-01, 1.16E-01)	(5.64E-03, -4.51E-04; -4.51E-04, 6.89E-03)
6 (166)	[8.908, 87.13]	[2.005, 22.813]	17	5.52E-01	(8.47E-01, 1.66E-01)	(6.04E-03, 2.45E-04; 2.45E-04, 3.76E-04)
				3.13E-01	(3.42E-01, 6.77E-01)	(1.94E-02, 1.34E-02; 1.34E-02, 1.47E-02)
				1.35E-01	(6.48E-01, 1.90E-01)	(5.26E-02, 5.57E-03; 5.57E-03, 5.12E-03)
7 (163)	[28.38, 93.954]	[0.174, 23.301]	0	3.91E-01	(7.87E-01, 2.45E-01)	(1.49E-02, 5.41E-04; 5.41E-04, 1.70E-03)
				1.71E-01	(2.25E-01, 6.31E-01)	(1.58E-02, 5.23E-03; 5.23E-03, 4.51E-02)
				2.37E-01	(6.26E-01, 8.33E-01)	(1.31E-02, 8.17E-03; 8.17E-03, 8.76E-03)
8 (116)	[11.333, 83.927]	[0.006, 23.921]	13	2.02E-01	(8.03E-01, 3.93E-01)	(1.13E-02, 1.75E-03; 1.75E-03, 4.64E-02)
				2.72E-01	(1.32E-01, 6.51E-01)	(1.04E-02, 2.33E-03; 2.33E-03, 1.36E-02)
				3.03E-01	(7.48E-01, 4.82E-01)	(1.09E-02, -3.67E-03; -3.67E-03, 4.33E-02)
9 (103)	[11.333, 83.927]	[0.006, 23.921]	13	4.24E-01	(6.30E-01, 1.95E-01)	(2.92E-02, 2.00E-04; 2.00E-04, 8.27E-04)
				1.96E-01	(9.58E-01, 1.65E-01)	(1.72E-03, 5.36E-04; 5.36E-04, 3.56E-03)
				2.25E-01	(7.33E-01, 1.53E-01)	(3.21E-03, 3.08E-04; 3.08E-04, 4.84E-04)
10 (103)	[11.333, 83.927]	[0.006, 23.921]	13	4.89E-01	(8.94E-01, 1.43E-01)	(4.13E-03, 2.36E-04; 2.36E-04, 1.26E-04)
				8.99E-02	(1.83E-01, 7.60E-01)	(2.07E-02, 8.54E-03; 8.54E-03, 1.04E-02)
				1.96E-01	(9.58E-01, 1.65E-01)	(1.72E-03, 5.36E-04; 5.36E-04, 3.56E-03)

Table 7: Summary of estimated parameters of Model II based on customer flexibility profiles for washing machine

Customer (Data Size)	$[X_{min}, X_{max}]$	$[Y_{min}, Y_{max}]$	Ref.	$\eta_k$	$\mu_k$	$\Sigma^k = (\Sigma_{11}, \Sigma_{12}; \Sigma_{21}, \Sigma_{22})$
1 (105)	[12.185, 82.902]	[4.184, 22.734]	16	1.33E-01	(5.74E-01, 2.39E-01)	(5.40E-02, -1.50E-02; -1.50E-02, 2.02E-02)
				3.55E-01	(5.75E-01, 1.78E-01)	(1.49E-03, -3.92E-05; -3.92E-05, 4.24E-04)
				6.53E-02	(2.68E-01, 8.82E-01)	(1.71E-02, 1.99E-03; 1.99E-03, 2.33E-03)
				4.47E-01	(7.08E-01, 1.77E-01)	(5.80E-03, -1.07E-05; -1.07E-05, 1.45E-04)
2 (253)	[4.828, 91.584]	[2.455, 23.077]	21	1.45E-01	(5.90E-01, 1.61E-01)	(3.57E-02, 1.61E-03; 1.61E-03, 7.50E-03)
				2.53E-01	(8.81E-01, 1.92E-01)	(4.96E-03, 5.29E-05; 5.29E-05, 4.71E-03)
				1.38E-01	(2.16E-01, 5.52E-01)	(1.06E-02, 7.61E-03; 7.61E-03, 2.31E-02)
				1.34E-01	(2.60E-01, 8.20E-01)	(1.11E-02, 9.08E-04; 9.08E-04, 1.12E-02)
3 (122)	[8.626, 94.296]	[4.369, 22.527]	0	8.11E-02	(7.28E-01, 6.34E-01)	(2.88E-02, -4.19E-03; -4.19E-03, 1.48E-02)
				2.49E-01	(5.48E-01, 8.29E-01)	(1.72E-02, 7.79E-04; 7.79E-04, 2.43E-03)
				3.82E-01	(7.01E-01, 1.80E-01)	(3.49E-03, -6.93E-05; -6.93E-05, 1.56E-05)
				2.96E-01	(4.87E-01, 3.41E-01)	(5.67E-02, -1.74E-02; -1.74E-02, 6.43E-02)
4 (133)	[11.301, 84.537]	[5.742, 21.778]	15	3.21E-01	(3.31E-01, 1.83E-01)	(5.00E-03, -3.14E-05; -3.14E-05, 2.79E-05)
				1.67E-01	(2.39E-01, 6.28E-01)	(2.09E-02, 1.09E-02; 1.09E-02, 2.39E-02)
				2.78E-01	(6.98E-01, 4.23E-01)	(1.19E-02, 8.69E-03; 8.69E-03, 4.21E-02)
				1.69E-01	(3.66E-01, 1.95E-01)	(2.12E-02, 2.11E-05; 2.11E-05, 1.60E-02)
5 (100)	[2.206, 92.976]	[0.605, 23.737]	63	3.86E-01	(7.55E-01, 1.16E-01)	(5.64E-03, -4.51E-04; -4.51E-04, 6.89E-03)
				5.39E-01	(8.47E-01, 1.66E-01)	(5.87E-03, 2.47E-04; 2.47E-04, 3.74E-04)
				1.20E-01	(2.69E-01, 6.62E-01)	(4.57E-02, 3.67E-02; 3.67E-02, 3.96E-02)
				1.34E-01	(6.53E-01, 1.89E-01)	(5.13E-02, 5.50E-03; 5.50E-03, 5.06E-03)
6 (109)	[14.609, 81.239]	[2.258, 23.972]	19	2.07E-01	(3.66E-01, 6.74E-01)	(3.45E-03, 8.79E-04; 8.79E-04, 1.04E-03)
				2.13E-01	(6.32E-01, 9.23E-01)	(2.69E-02, 2.16E-03; 2.16E-03, 4.19E-03)
				3.31E-01	(2.26E-01, 5.48E-01)	(1.40E-02, 6.31E-03; 6.31E-03, 1.13E-02)
				1.41E-01	(5.47E-01, 1.33E-01)	(3.21E-02, -5.03E-03; -5.03E-03, 7.68E-03)
7 (103)	[5.258, 95.513]	[2.025, 21.795]	0	3.15E-01	(8.50E-01, 2.94E-01)	(6.80E-03, 3.78E-04; 3.78E-04, 1.71E-02)
				2.61E-01	(2.59E-01, 2.12E-01)	(2.07E-02, -1.93E-03; -1.93E-03, 7.93E-03)
				4.43E-01	(8.34E-01, 2.41E-01)	(1.09E-02, 1.88E-03; 1.88E-03, 4.99E-03)
				1.07E-01	(5.88E-01, 7.62E-01)	(1.26E-02, 6.44E-05; 6.44E-05, 1.68E-02)
				1.89E-01	(3.10E-01, 8.72E-01)	(4.00E-03, 3.76E-04; 3.76E-04, 7.59E-03)

(Continued on next page)

Table 7: Summary of estimated parameters of Model II based on customer flexibility profiles for washing machine (*continued*)

Customer (Data Size)	$[X_{min}, X_{max}]$	$[Y_{min}, Y_{max}]$	Ref.	$\eta_k$	$\mu_k$	$\Sigma^k=(\Sigma_{11}, \Sigma_{12}; \Sigma_{21}, \Sigma_{22})$
8 (139)	[4.010, 91.115]	[0.377, 23.523]	32	5.56E-01	(2.85E-01, 6.90E-01)	(2.87E-02, 2.63E-02; 2.63E-02, 2.59E-02)
				2.83E-01	(6.70E-01, 3.72E-01)	(1.07E-02, 1.29E-03; 1.29E-03, 4.98E-03)
				1.60E-01	(4.82E-01, 5.76E-01)	(7.16E-02, -1.96E-02; -1.96E-02, 5.59E-02)
9 (110)	[3.688, 91.611]	[4.907, 22.561]	13	3.84E-01	(4.30E-01, 7.52E-01)	(5.01E-03, 9.85E-04; 9.85E-04, 1.42E-03)
				3.11E-01	(8.90E-01, 1.45E-01)	(2.65E-03, 3.09E-04; 3.09E-04, 2.12E-03)
				3.06E-01	(1.81E-01, 5.21E-01)	(2.75E-02, 5.51E-03; 5.51E-03, 4.18E-02)
10 (104)	[4.023, 92.217]	[0.323, 23.781]	26	9.18E-02	(2.76E-01, 1.30E-01)	(2.77E-02, -9.61E-04; -9.61E-04, 2.60E-04)
				2.18E-01	(7.93E-01, 1.10E-01)	(1.01E-02, -1.11E-03; -1.11E-03, 2.16E-03)
				2.98E-01	(8.71E-01, 1.21E-01)	(4.04E-03, -6.91E-05; -6.91E-05, 8.53E-05)
11 (218)	[3.635, 91.896]	[3.700, 21.281]	18	3.44E-01	(8.59E-01, 1.52E-01)	(5.19E-03, 7.90E-05; 7.90E-05, 6.73E-05)
				4.82E-02	(2.42E-01, 9.92E-01)	(1.99E-02, 6.14E-04; 6.14E-04, 2.02E-04)
				1.08E-01	(7.34E-01, 2.77E-01)	(6.87E-03, -2.04E-03; -2.04E-03, 4.41E-03)
12 (173)	[3.442, 92.674]	[0.143, 23.429]	20	5.36E-01	(8.26E-01, 2.59E-01)	(2.98E-03, -2.82E-05; -2.82E-05, 2.15E-04)
				6.22E-02	(9.23E-01, 3.76E-01)	(3.99E-03, 1.05E-03; 1.05E-03, 3.06E-03)
				2.31E-01	(8.60E-01, 3.10E-01)	(1.10E-03, -3.84E-06; -3.84E-06, 3.80E-04)
12 (173)	[3.442, 92.674]	[0.143, 23.429]	20	6.19E-02	(2.48E-01, 8.72E-01)	(1.66E-02, 6.77E-03; 6.77E-03, 4.63E-03)
				5.94E-01	(2.03E-01, 6.02E-02)	(1.76E-03, -1.86E-05; -1.86E-05, 2.84E-05)
				2.84E-01	(2.64E-01, 6.15E-02)	(7.07E-03, 5.14E-05; 5.14E-05, 1.99E-05)
1 (199)	[16.814, 79.651]	[2.350, 21.473]	55	2.19E-02	(2.42E-02, 9.99E-01)	(8.98E-03, 1.54E-04; 1.54E-04, 9.65E-05)
				5.64E-02	(4.61E-01, 5.88E-02)	(1.17E-02, 2.02E-04; 2.02E-04, 3.64E-05)
				4.42E-02	(7.01E-01, 5.34E-02)	(6.89E-02, 6.36E-03; 6.36E-03, 2.47E-03)

Table 8: Summary of estimated parameters of Model II based on customer flexibility profiles for dishwasher

Customer (Data Size)	$[X_{min}, X_{max}]$	$[Y_{min}, Y_{max}]$	Ref.	$\eta_k$	$\mu_k$	$\Sigma^k=(\Sigma_{11}, \Sigma_{12}; \Sigma_{21}, \Sigma_{22})$
1 (199)	[16.814, 79.651]	[2.350, 21.473]	55	6.65E-02	(5.62E-01, 5.17E-02)	(5.66E-02, 4.10E-03; 4.10E-03, 1.66E-03)
				2.36E-02	(9.71E-03, 1.00E+00)	(1.61E-02, 1.14E-04; 1.14E-04, 1.17E-04)
				9.10E-01	(2.23E-01, 6.05E-02)	(4.28E-03, 2.07E-05; 2.07E-05, 2.59E-05)

(Continued on next page)

Table 8: Summary of estimated parameters of Model II based on customer flexibility profiles for dishwasher (*continued*)

Customer (Data Size)	$[X_{min}, X_{max}]$	$[Y_{min}, Y_{max}]$	Ref.	$\eta_k$	$\mu_k$	$\Sigma^k=(\Sigma_{11}, \Sigma_{12}; \Sigma_{21}, \Sigma_{22})$
2 (105)	[13.308, 82.571]	[1.634, 17.467]	15	1.87E-01	(7.92E-01, 9.49E-02)	(1.14E-02, -4.53E-04; -4.53E-04, 1.75E-03)
				5.03E-02	(2.42E-01, 9.92E-01)	(2.39E-02, 3.19E-04; 3.19E-04, 8.45E-04)
				9.68E-02	(2.86E-01, 1.29E-01)	(3.16E-02, -1.27E-03; -1.27E-03, 5.81E-04)
				6.66E-01	(8.60E-01, 1.38E-01)	(4.94E-03, -2.35E-04; -2.35E-04, 4.16E-04)
3 (202)	[9.193, 87.692]	[3.805, 15.130]	69	8.04E-02	(8.46E-02, 1.28E-01)	(3.41E-03, 1.23E-04; 1.23E-04, 1.08E-03)
				5.63E-01	(2.83E-01, 1.06E-01)	(2.29E-03, 2.85E-05; 2.85E-05, 1.01E-04)
				6.56E-02	(7.38E-01, 7.35E-01)	(1.23E-02, -9.87E-03; -9.87E-03, 5.89E-02)
				1.48E-01	(2.20E-01, 1.30E-01)	(4.39E-03, -1.02E-03; -1.02E-03, 1.14E-02)
4 (246)	[12.292, 82.805]	[0.175, 16.604]	17	1.43E-01	(3.12E-01, 1.90E-01)	(1.81E-03, -1.30E-04; -1.30E-04, 2.46E-04)
				5.23E-01	(8.24E-01, 2.58E-01)	(3.07E-03, -3.33E-05; -3.33E-05, 2.04E-04)
				1.71E-01	(7.95E-01, 3.02E-01)	(1.14E-02, 1.56E-03; 1.56E-03, 5.47E-03)
				6.11E-02	(2.49E-01, 8.73E-01)	(1.62E-02, 6.63E-03; 6.63E-03, 4.57E-03)
5 (129)	[15.868, 79.402]	[0.281, 18.923]	17	2.46E-01	(8.59E-01, 3.08E-01)	(1.25E-03, 2.08E-04; 2.08E-04, 5.58E-04)
				1.96E-01	(9.58E-01, 1.65E-01)	(1.72E-03, 5.36E-04; 5.36E-04, 3.56E-03)
				2.25E-01	(7.33E-01, 1.53E-01)	(3.21E-03, 3.08E-04; 3.08E-04, 4.84E-04)
				4.89E-01	(8.94E-01, 1.43E-01)	(4.13E-03, 2.36E-04; 2.36E-04, 1.26E-04)
6 (113)	[9.579, 85.951]	[0.121, 21.758]	18	8.99E-02	(1.83E-01, 7.60E-01)	(2.07E-02, 8.54E-03; 8.54E-03, 1.04E-02)
				3.05E-01	(3.31E-01, 1.83E-01)	(5.58E-03, -2.40E-05; -2.40E-05, 2.81E-05)
				9.94E-02	(4.83E-01, 1.89E-01)	(4.25E-02, -1.22E-05; -1.22E-05, 4.79E-05)
				1.02E-01	(5.75E-01, 1.92E-01)	(4.20E-02, 4.29E-05; 4.29E-05, 6.22E-05)
7 (285)	[7.202, 88.372]	[2.383, 18.716]	22	3.65E-01	(7.03E-01, 1.81E-01)	(5.10E-03, -6.01E-05; -6.01E-05, 1.58E-05)
				1.29E-01	(4.00E-01, 5.57E-01)	(5.20E-02, -1.28E-02; -1.28E-02, 7.99E-02)
				1.33E-01	(5.74E-01, 2.39E-01)	(5.40E-02, -1.50E-02; -1.50E-02, 2.02E-02)
				3.55E-01	(5.75E-01, 1.78E-01)	(1.49E-03, -3.92E-05; -3.92E-05, 4.24E-04)
8 (168)	[11.963, 84.778]	[0.345, 22.742]	16	6.53E-02	(2.68E-01, 8.82E-01)	(1.71E-02, 1.99E-03; 1.99E-03, 2.33E-03)
				4.47E-01	(7.08E-01, 1.77E-01)	(5.80E-03, -1.07E-05; -1.07E-05, 1.45E-04)
				2.01E-01	(2.62E-01, 6.53E-01)	(3.37E-02, 1.20E-02; 1.20E-02, 2.47E-02)
				1.20E-01	(7.41E-01, 2.67E-01)	(1.64E-02, -2.96E-04; -2.96E-04, 8.23E-03)
				5.50E-01	(7.81E-01, 2.20E-01)	(1.02E-02, 2.82E-04; 2.82E-04, 1.41E-04)
				1.29E-01	(3.00E-01, 2.26E-01)	(3.60E-02, 2.01E-04; 2.01E-04, 3.28E-04)

(Continued on next page)

Table 8: Summary of estimated parameters of Model II based on customer flexibility profiles for dishwasher (*continued*)

Customer (Data Size)	$[X_{min}, X_{max}]$	$[Y_{min}, Y_{max}]$	Ref.	$\eta_k$	$\mu_k$	$\Sigma^k=(\Sigma_{11}, \Sigma_{12}; \Sigma_{21}, \Sigma_{22})$
9 (186)	[13.429, 82.243]	[0.390, 23.384]	16	1.02E-01	(8.77E-01, 2.24E-01)	(5.10E-03, 9.73E-04; 9.73E-04, 2.10E-03)
				5.46E-02	(5.07E-01, 9.15E-01)	(1.25E-02, 1.19E-04; 1.19E-04, 3.05E-03)
				2.32E-01	(3.63E-01, 6.74E-01)	(3.87E-03, 1.05E-03; 1.05E-03, 1.31E-03)
				1.00E-01	(2.57E-01, 2.81E-01)	(4.21E-02, -2.69E-02; -2.69E-02, 2.47E-02)
				5.11E-01	(8.40E-01, 1.65E-01)	(6.31E-03, 1.85E-04; 1.85E-04, 3.67E-04)
10 (141)	[11.126, 84.580]	[4.143, 21.225]	18	1.83E-01	(6.95E-01, 1.39E-01)	(1.73E-02, 4.82E-03; 4.82E-03, 9.42E-03)
				2.47E-01	(2.23E-01, 4.66E-01)	(1.13E-02, 1.72E-02; 1.72E-02, 5.01E-02)
				2.25E-01	(6.67E-01, 6.23E-01)	(4.66E-03, 1.61E-03; 1.61E-03, 2.78E-02)
				3.45E-01	(7.46E-02, 5.91E-01)	(2.71E-03, 8.89E-05; 8.89E-05, 1.54E-02)
				3.91E-01	(7.87E-01, 2.45E-01)	(1.49E-02, 5.41E-04; 5.41E-04, 1.70E-03)
11 (116)	[12.960, 83.511]	[0.028, 23.573]	20	1.71E-01	(2.25E-01, 6.31E-01)	(1.58E-02, 5.23E-03; 5.23E-03, 4.51E-02)
				2.37E-01	(6.26E-01, 8.33E-01)	(1.31E-02, 8.17E-03; 8.17E-03, 8.76E-03)
				2.02E-01	(8.03E-01, 3.93E-01)	(1.13E-02, 1.75E-03; 1.75E-03, 4.64E-02)
				3.61E-01	(6.10E-01, 1.93E-01)	(2.90E-02, -5.22E-05; -5.22E-05, 6.16E-04)
				1.61E-01	(9.97E-02, 6.86E-01)	(6.50E-03, -1.27E-04; -1.27E-04, 1.14E-03)
12 (199)	[4.797, 91.170]	[0.381, 23.127]	23	2.28E-01	(7.59E-01, 3.02E-01)	(9.53E-03, -1.69E-03; -1.69E-03, 1.21E-02)
				1.01E-01	(7.70E-01, 6.61E-01)	(8.24E-03, -2.98E-04; -2.98E-04, 1.22E-03)
				1.49E-01	(2.83E-01, 6.45E-01)	(4.79E-02, 2.51E-02; 2.51E-02, 3.13E-02)
				1.65E-01	(8.25E-01, 7.16E-01)	(4.81E-03, 8.75E-04; 8.75E-04, 2.68E-03)
				1.96E-01	(3.90E-01, 7.38E-01)	(2.26E-02, 1.99E-02; 1.99E-02, 2.31E-02)
13 (270)	[6.595, 90.142]	[2.639, 19.398]	8	2.09E-01	(8.19E-01, 1.09E-01)	(5.31E-03, -1.33E-04; -1.33E-04, 2.21E-03)
				1.83E-01	(1.52E-01, 7.11E-01)	(4.90E-03, 2.83E-03; 2.83E-03, 1.06E-02)
				2.47E-01	(8.29E-01, 3.68E-01)	(6.68E-03, 7.99E-04; 7.99E-04, 1.33E-02)
				6.94E-02	(2.72E-01, 1.76E-01)	(2.38E-02, -2.95E-03; -2.95E-03, 1.70E-03)
				1.65E-01	(2.96E-01, 7.17E-01)	(4.37E-02, 1.11E-02; 1.11E-02, 1.91E-02)
14 (139)	[30.056, 91.676]	[0.164, 22.142]	0	2.94E-01	(8.90E-01, 1.46E-01)	(2.58E-03, 2.36E-04; 2.36E-04, 1.96E-03)
				3.35E-01	(4.25E-01, 7.50E-01)	(3.43E-03, 1.39E-04; 1.39E-04, 7.54E-04)
				1.36E-01	(8.13E-02, 4.80E-01)	(3.46E-03, 7.27E-04; 7.27E-04, 1.76E-03)

(Continued on next page)



Table 8: Summary of estimated parameters of Model II based on customer flexibility profiles for dishwasher (*continued*)

Customer (Data Size)	$[X_{min}, X_{max}]$	$[Y_{min}, Y_{max}]$	Ref.	$\eta_k$	$\mu_k$	$\Sigma^k = (\Sigma_{11}, \Sigma_{12}; \Sigma_{21}, \Sigma_{22})$
15 (253)	[9.304, 85.805]	[3.928, 23.743]	18	1.67E-01	(2.39E-01, 6.28E-01)	(2.09E-02, 1.09E-02; 1.09E-02, 2.39E-02)
				2.78E-01	(6.98E-01, 4.23E-01)	(1.19E-02, 8.69E-03; 8.69E-03, 4.21E-02)
				1.69E-01	(3.66E-01, 1.95E-01)	(2.12E-02, 2.11E-05; 2.11E-05, 1.60E-02)
				3.86E-01	(7.55E-01, 1.16E-01)	(5.64E-03, -4.51E-04; -4.51E-04, 6.89E-03)

## 6. References

- [1] N. Sadeghianpourhamami, M. Strobbe, C. Develder, Real-world user flexibility of energy consumption: Two-stage generative model construction, in: Proc. ACM Symp. Appl. Comput. (SAC2016), 2016, pp. 1–6.
- [2] P. Siano, Demand response and smart grids survey, *Renewable and Sustainable Energy Reviews* 30 (2014) 461 – 478. doi:10.1016/j.rser.2013.10.022.
- [3] J. Vardakas, N. Zorba, C. Verikoukis, A survey on demand response programs in smart grids: Pricing methods and optimization algorithms, *IEEE Commun. Surveys Tuts* 17 (1) (2015) 152–178. doi:10.1109/COMST.2014.2341586.
- [4] S. Nolan, M. O'Malley, Challenges and barriers to demand response deployment and evaluation, *Appl. Energy* 152 (2015) 1–10. doi:10.1016/j.apenergy.2015.04.083.
- [5] M. Alczar-Ortega, C. Calpe, T. Theisen, J. F. Carbonell-Carretero, Methodology for the identification, evaluation and prioritization of market handicaps which prevent the implementation of demand response: Application to european electricity markets, *Energy Policy* 86 (C) (2015) 529–543.
- [6] C. Goldman, N. Hopper, R. Bharvirkar, B. Neenan, P. Cappers, Estimating demand response market potential among large commercial and industrial customers: a scoping study, Lawrence Berkeley National Laboratory.
- [7] Demand response from day-ahead hourly pricing for large customers, *The Electricity Journal* 19 (3) (2006) 52 – 63. doi:10.1016/j.tej.2006.02.002.
- [8] D. Jang, J. Eom, M. G. Kim, J. J. Rho, Demand responses of korean commercial and industrial businesses to critical peak pricing of electricity, *Journal of Cleaner Production* 90 (2015) 275 – 290. doi:http://dx.doi.org/10.1016/j.jclepro.2014.11.052.  
URL <http://www.sciencedirect.com/science/article/pii/S0959652614012396>
- [9] R. Deshmukh, G. Ghatikar, R. Yin, G. G. Das, S. K. Saha, Estimation of potential and value of demand response for industrial and commercial consumers in delhi, Presented at the India Smart Grid Week (ISGW).
- [10] Z. Liu, I. Liu, S. Low, A. Wierman, Pricing data center demand response, in: *ACM SIGMETRICS Performance Evaluation Review*, Vol. 42, ACM, 2014, pp. 111–123.
- [11] V. Gomez, M. Chertkov, S. Backhaus, H. Kappen, Learning price-elasticity of smart consumers in power distribution systems, in: Proc. 3rd IEEE Int. Conf. Smart Grid Commun. (SmartGridComm2012), 2012, pp. 647–652. doi:10.1109/SmartGridComm.2012.6486059.
- [12] A. van Stiphout, J. Engels, D. Guldentops, G. Deconinck, Quantifying the flexibility of residential electricity demand in 2050: a bottom-up approach, in: *PowerTech*, IEEE Eindhoven, 2015, pp. 1–6. doi:10.1109/PTC.2015.7232411.
- [13] P. Thimmapuram, J. Kim, Consumers' price elasticity of demand modeling with economic effects on electricity markets using an agent-based model, *IEEE Trans. Smart Grid* 4 (1) (2013) 390–397. doi:10.1109/TSG.2012.2234487.
- [14] A. Moshari, G. Yousefi, A. Ebrahimi, S. Haghbin, Demand-side behavior in the smart grid environment, in: Proc. 1st IEEE PES Innovative Smart Grid Technol. Conf. Europe (ISGT Europe), 2010, pp. 1–7. doi:10.1109/ISGTEUROPE.2010.5638956.
- [15] N. Paterakis, J. Catalao, A. Tascikaraoglu, A. Bakirtzis, O. Erdinc, Demand response driven load pattern elasticity analysis for smart households, in: Proc. 5th IEEE Int. Conf. Power Engineering, Energy and Electrical Drives (POWERENG), 2015, pp. 399–404. doi:10.1109/PowerEng.2015.7266350.
- [16] P. Siano, D. Sarno, Assessing the benefits of residential demand response in a real time distribution energy market, *Appl. Energy* 161 (2016) 533–551. doi:10.1016/j.apenergy.2015.10.017.
- [17] Q. Hu, X. Fang, F. Li, X. Xu, C. fei Chen, H. Hu, An approach to assess the responsive residential demand to financial incentives, in: *IEEE PES General Meeting*, 2015, pp. 1–5. doi:10.1109/PESGM.2015.7286623.
- [18] M. Pipattanasomporn, M. Kuzlu, S. Rahman, Y. Teklu, Load profiles of selected major household appliances and their demand response opportunities, *IEEE Trans. Smart Grid* 5 (2) (2014) 742–750. doi:10.1109/TSG.2013.2268664.
- [19] F. Wattjes, S. Janssen, J. Slootweg, Framework for estimating flexibility of commercial and industrial customers in smart grids, in: Proc. 4th IEEE PES Innovative Smart Grid Technol. Conf. Europe (ISGT Europe), 2013, pp. 1–5. doi:10.1109/ISGTEurope.2013.6695406.
- [20] I. Laicane, D. Blumberga, A. Blumberga, M. Rosa, Reducing household electricity consumption through demand side management: The role of home appliance scheduling and peak load reduction, *Energy Procedia* 72 (2015) 222 – 229. doi:10.1016/j.egypro.2015.06.032.
- [21] A. Safdarian, M. Fotuhi-Firuzabad, M. Lehtonen, Benefits of demand response on operation of distribution networks: A case study, *IEEE Syst. J. PP* (99) (2014) 1–9. doi:10.1109/JSYST.2013.2297792.
- [22] K. Kouzelis, Z. Tan, B. Bak-Jensen, J. Pillai, E. Ritchie, Estimation of residential heat pump consumption for flexibility market applications, *IEEE Trans. Smart Grid* 6 (4) (2015) 1852–1864. doi:10.1109/TSG.2015.2414490.
- [23] W. Labeeuw, J. Stragier, G. Deconinck, Potential of active demand reduction with residential wet appliances: A case study for Belgium, *IEEE Trans. Smart Grid* 6 (1) (2015) 315–323. doi:10.1109/TSG.2014.2357343.
- [24] C. B. Kobus, E. A. Klaassen, R. Mugge, J. P. Schoormans, A real-life assessment on the effect of smart appliances for shifting households electricity demand, *Appl. Energy* 147 (2015) 335–343. doi:10.1016/j.apenergy.2015.01.073.
- [25] R. D'Hulst, W. Labeeuw, B. Beusen, S. Claessens, G. Deconinck, K. Vanthournout, Demand response flexibility and flexibility potential of residential smart appliances: Experiences from large pilot test in Belgium, *Appl. Energy* 155 (2015) 79–90. doi:10.1016/j.apenergy.2015.05.101.
- [26] B. Dupont, P. Vingerhoets, P. Tant, K. Vanthournout, W. Cardinaels, T. De Rybel, E. Peeters, R. Belmans, Linear breakthrough project: Large-scale implementation of smart grid technologies in distribution grids, in: Proc. 3rd IEEE PES Innovative Smart Grid Technol. Conf. Europe (ISGT Europe), 2012, pp. 1–8. doi:10.1109/ISGTEurope.2012.6465708.
- [27] I. Sajjad, G. Chicco, M. Aziz, A. Rasool, Potential of residential demand flexibility – Italian scenario, in: Proc. 11th IEEE Int. Multi-Conf. Systems, Signals Devices (SSD), 2014, pp. 1–6. doi:10.1109/SSD.2014.6808849.

- [28] K. Vanthournout, B. Dupont, W. Foubert, C. Stuckens, S. Claessens, An automated residential demand response pilot experiment, based on day-ahead dynamic pricing, *Appl. Energy* 155 (2015) 195 – 203. doi:10.1016/j.apenergy.2015.05.100.
- [29] G. Hamerly, C. Elkan, Learning the k in k-means, in: *Proc. 17th Ann. Conf. Neural Information Processing Systems*, 2003.
- [30] K. Mets, F. Depuydt, C. Develder, Two-stage load pattern clustering using fast wavelet transformation, *IEEE Trans. Smart Grid* PP (99). doi:10.1109/TSG.2015.2446935.
- [31] G. Chicco, Overview and performance assessment of the clustering methods for electrical load pattern grouping, *Energy* 42 (1) (2012) 68 – 80. doi:10.1016/j.energy.2011.12.031.
- [32] Fr Finite mixture and Markov switching models.
- [33] H. Akaike, A new look at the statistical model identification, *IEEE Trans. Autom. Control* 19 (6) (1974) 716–723. doi:10.1109/TAC.1974.1100705.
- [34] G. Schwarz, Estimating the Dimension of a Model, *The Ann. Statist.* 6 (2) (1978) 461–464. doi:10.2307/2958889.
- [35] W. Xie, P. O. Lewis, Y. Fan, L. Kuo, M.-H. Chen, Improving marginal likelihood estimation for bayesian phylogenetic model selection, *Systematic Biology* 60 (2) (2011) 150–160. doi:10.1093/sysbio/syq085.
- [36] M. Stephens, Bayesian analysis of mixture models with an unknown number of componentsan alternative to reversible jump methods, *Ann. Statist.* 28 (1) (2000) 40–74. doi:10.1214/aos/1016120364.
- [37] J. A. Peacock, Two-dimensional goodness-of-fit testing in astronomy, *Monthly Notices of the Royal Astronomical Society* 202 (3) (1983) 615–627. doi:10.1093/mnras/202.3.615.
- [38] G. H. Freeman, J. H. Halton, Note on an exact treatment of contingency, goodness of fit and other problems of significance, *Biometrika* 38 (1/2) (1951) 141–149.