Modeling Real-World Flexibility of Residential Power Consumption: Exploring the Cylindrical WeiSSVM Distribution

Nasrin Sadeghianpourhamami
IDLab Ghent University-imec, Belgium
nasrin.sadeghianpourhamami@ugent.be

Dries F. Benoit
Faculty of Economics and Business Administration, Ghent University, Belgium
dries.benoit@ugent.be

Dirk Deschrijver
IDLab Ghent University-imec, Belgium
dirk.deschrijver@ugent.be

Chris Develder
IDLab Ghent University-imec, Belgium
chris.develder@ugent.be

ABSTRACT

A user's power consumption flexibility is defined in terms of amount, time and duration of availability. The timing of flexibility is circular in nature. Therefore, it is natural to adopt circular distributions to model this data. This paper investigates the key research question whether that leads to better generative models than using conventional linear distributions. In particular, it fits Gaussian mixture models and a very flexible recent cylindrical (WeiSSVM) distribution mixture to real-world field trial data. Using a predictive accuracy performance measure, it is found that the latter does not provide substantially better fits. Shortcomings of both models are pointed out and it is concluded that research for appropriate statistical models for the observed data is still open.

KEYWORDS

Residential Flexibility Modeling, Cylindrical Distributions

ACM Reference Format:
Nasrin Sadeghianpourhamami, Dries F. Benoit, Dirk Deschrijver, and Chris Develder. 2018. Modeling Real-World Flexibility of Residential Power Consumption: Exploring the Cylindrical WeiSSVM Distribution. In *e-Energy '18: The Ninth International Conference on Future Energy Systems, June 12–15, 2018, Karlsruhe, Germany.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3208903.3212041

1 INTRODUCTION

User energy consumption flexibility is typically characterized by the amount and the duration of the deferrable energy at various times of the days. The timing aspect of the flexibility is greatly influenced by user lifestyle and energy consumption habits. Hence, to derive generative models of user flexibility, its timing aspect is quantified using configuration time and deadline [5]. The *configuration time* indicates when the users configure their smart appliances flexibly and the *deadline* is the latest possible start time of the appliance. The configuration time is of cyclic nature while the deadline is a linear quantity. Hence, flexibility measurements can naturally be regarded as bivariate cylindrical data.

Initial studies modeling the energy consumption flexibility avoid the cylindrical representation by defining a heuristic algorithm that identifies the middle of the largest gap on the circular axis to wrap the data around and proceed to modeling using probabilistic models defined on linear scales (e.g., [5] [2]). However, such heuristic algorithms might fail in situations where such a reference point is challenging or impossible to find. On the other hand, most of the existing cylindrical distributions are limited in terms of flexibility in modeling cross-correlation between the cylindrical and linear variables as well as modeling skewness and heterogeneity in the data (which requires mixture models). Recently, Abe and Ley proposed a cylindrical distribution (named WeiSSVM) which is tractable, flexible, has well-known conditional and marginal distributions and models the skewness and cross-correlation between the cylindrical and linear variables [1].

This paper investigates how well WeiSSVM mixtures can model the energy consumption flexibility compared to using distributions defined on the linear scale. The analysis is based on data from year-long measurements in the LINEAR pilot project [3], where [5] previously modeled the user behavior towards smart wet-appliances with Gaussian mixture models (GMM).

2 MODELING USER ENERGY CONSUMPTION FLEXIBILITY WITH WEISSVM MIXTURES

2.0.1 PDF of WeiSSVM Mixtures. The WeiSSVM distribution is a combination of a Weibull distribution and the sine-skewed Von-Mises distribution. Its probability density is defined as [1]:

$$f(\theta, x|\boldsymbol{\zeta}) \mapsto \frac{\alpha\beta^{\alpha}}{2\pi\cosh(\kappa)} \cdot (1 + \lambda\sin(\theta - \mu)) \cdot x^{\alpha-1} \cdot \exp[-(\beta x)^{\alpha}(1 - \tanh(\kappa)\cos(\theta - \mu)],$$

with random variables $(\theta, x) \in [0, 2\pi) \times [0, \infty)$, and distribution parameters $\boldsymbol{\zeta} = (\alpha, \beta, \mu, \kappa, \lambda)$. The parameter vector of the WeiSSVM distribution comprises $\alpha, \beta > 0$, which are linear shape and scale parameters respectively, $0 \leq \mu < 2\pi$ is a circular location parameter, $\kappa \geq 0$ controls the circular concentration, and $-1 \leq \lambda \leq 1$ controls the circular skewness. The mixture of a K-component WeiSSVM distribution has the following density function:

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

e-Energy '18, June 12–15, 2018, Karlsruhe, Germany

ACM ISBN 978-1-4503-5767-8/18/06…$15.00
https://doi.org/10.1145/3208903.3212041

Table 1: *elppd* of GMM and WMM fits for dishwashers.

User	WMM	GMM	User	WMM	GMM
1	-73.0(3)	**-45.1(3)**	9	-496.9(5)	**-469.8(5)**
2	-173.0(2)	**-158.0(2)**	10	-579.0(4)	**-556.1(4)**
3	-247.8(5)	**-229.3(4)**	11	-486.7(4)	**-464.7(4)**
4	-341.3(5)	**-314.7(4)**	12	-807.6(3)	**-786.3(5)**
5	-198.0(5)	**-157.5(5)**	13	-1054.3(6)	**-1011.7(5)**
6	-245.6(4)	**-219.3(4)**	14	-450.8 (4)	**-413.7(4)**
7	**-401.5(5)**	-411.5(4)	15	-1031.9(5)	**-1015.6(5)**
8	-488.6(4)	**-456.6(4)**			

where $f_k(\theta, x|\boldsymbol{\zeta_k})$ is the probability density of the k^{th} component indexed by parameter set $\boldsymbol{\zeta_k}$; η_k is the weight of the k^{th} component, thus $\boldsymbol{\eta} = (\eta_1, \eta_2, ..., \eta_K)$ is the weight distribution constrained by:

$$\eta_k \geq 0, \quad \eta_1 + \eta_2 + ... + \eta_K = 1.$$

Hence, $\boldsymbol{\vartheta} = (\zeta_1, ..., \zeta_K, \boldsymbol{\eta})$ is the parameter vector of the mixture model.

2.0.2 Model Parameter Estimation. Note that the likelihood of the WMMs is a high dimensional function. Hence, Bayesian methods are more reliable than point-estimates (e.g., expectation maximization) in estimating the parameters of the WMMs since they output the entire posterior distribution. The Metropolis-Hastings algorithm [4] is used to estimate the parameters of the WMMs for the user flexibility.

2.0.3 Measures for Model Comparison. One of the use cases of generative models of the energy usage flexibility is to generate data samples (e.g., to simulate scenarios for assessing DR impact). Hence, it is natural to compare the generative models in terms of their out-of-sample predictive accuracy. For comparing the models, we use log point-wise prediction density (*elppd*), calculated from the posterior samples, as a popular method of quantizing the out-of-sample predictive accuracy of a model [6]. When modeling data using mixture models, one also needs to identify the optimal number of mixtures. This value is determined by finding a knee point in the plot of the *elppd* measure vs. the number of mixtures.

3 ANALYSIS AND MODEL COMPARISON

This section presents the results of fitting WMMs as a generative model of the energy usage flexibility for the households that participated in the aforementioned LINEAR pilot project [3]. This data (comprising 15 dishwashers, 12 washing machines and 8 tumble dryers) was previously modeled using GMMs [5]. The models are compared next. In terms of the resulting clusters: GMMs typically identify either clusters in parallel with the x-axis or clusters along a diagonal as seen from the example users depicted in the bottom row of Fig. 1. The former indicates configurations with similar deadline while the latter maps to configurations with similar flexibility duration. WMMs also identify clusters in parallel with the x-axis. As opposed to GMMs, WMMs find clusters in parallel with the y-axis, indicating similar configuration time but varying deadlines and hence different flexibility duration. Also, due to the inherent nature of the distributions, each GMM component is symmetrical, but

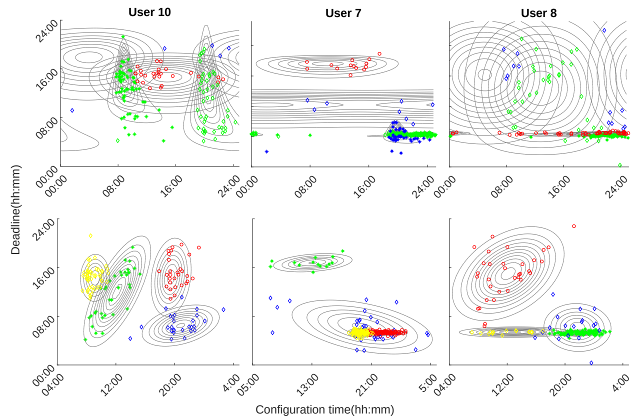

Fig. 1: Comparison of distribution shapes of the mixture components and identified clusters with WMMs (top row) and GMMs (bottom row) for selected users of dishwashers. Note that data on bottom row is wrapped around a new x-axis reference, while for the data on top row, the x-axis is circular (i.e., 00:00 and 24:00 are the same points)

WMMs have skewed distributions with increasing concentration along the linear axis.

To compare the predictive accuracy, we calculated the *elppd* values for the users of all three appliances and summarized the values for dishwashers in Table 1. The numbers between parentheses are the optimum number of mixtures. The bold values in Table 1 indicate a better generative model. As seen from Table 1, WMMs perform comparable or worse than the GMMs for the majority of users of dishwashers. Similar results are obtained for users of washing machines and tumble dryers. This is due to inherent characteristics of the WMMs and GMMs. WMMs are more suitable in modeling datasets in which the circular concentration increases along the linear axis, while the measurements in LINEAR show that users do not exhibit such trait. Instead, the areas of high densities indicating a consistent daily behavior is seen in configuration patterns of the users for all three appliances. This characteristic is well modeled with GMMs.

4 CONCLUSION

An analysis of the LINEAR dataset shows that GMMs are preferred over WMMs, because they compute better generative models. Despite having better predictive accuracy than WMMs, GMMs still suffer from the following limitations: (1) GMMs are defined for linear scale and are not suitable for modeling scenarios where an adequate reference on the circular axis is impossible or challenging to find, (2) GMMs are defined on both negative and positive values, but the flexibility characteristics are positive quantities. Hence, GMMs are prone to generating meaningless samples.

Hence, defining a suitable distribution for modeling user energy consumption behavior cylindrically is still an open research issue. Given that (compared to real-life trials) simulating scenarios for testing DR algorithms' impacts is highly cost effective, it is also a highly relevant issue.

REFERENCES

[1] Toshihiro Abe and Christophe Ley. 2017. A tractable, parsimonious and flexible model for cylindrical data, with applications. *Econometrics and Statistics* 4, Supplement C (2017), 91 – 104. https://doi.org/10.1016/j.ecosta.2016.04.001

[2] Chris Develder, Nasrin Sadeghianpourhamami, Matthias Strobbe, and Nazir Refa. 2016. Quantifying flexibility in EV charging as DR potential: Analysis of two real-world data sets. In *Proc. 7th IEEE Int. Conf. Smart Grid Communications (SmartGridComm 2016).* Sydney, Australia, 600–605. https://doi.org/10.1109/SmartGridComm.2016.7778827

[3] B. Dupont, P. Vingerhoets, P. Tant, K. Vanthournout, W. Cardinaels, T. De Rybel, E. Peeters, and R. Belmans. 2012. LINEAR breakthrough project: Large-scale implementation of smart grid technologies in distribution grids. In *2012 3rd IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe).* 1–8. https://doi.org/10.1109/ISGTEurope.2012.6465708

[4] W. K. Hastings. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* (1970), 97–109. https://doi.org/10.1093/biomet/57.1.97

[5] N. Sadeghianpourhamami, T. Demeester, D.F. Benoit, M. Strobbe, and C. Develder. 2016. Modeling and analysis of residential flexibility: Timing of white good usage. *Applied Energy* 179, Supplement C (2016), 790 – 805. https://doi.org/10.1016/j.apenergy.2016.07.012

[6] Aki Vehtari, Andrew Gelman, and Jonah Gabry. 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing* 27, 5 (01 Sep 2017), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4