Disaggregated Electricity Forecasting using Wavelet-Based Clustering of Individual Consumers

Jairo Cugliari
Université de Lyon
Lyon, France
Jairo.Cugliari@univ-lyon2.fr

Yannig Goude
EDF R&D, Clamart, France and
LMO, Univ. Paris-Sud Orsay, France
Yannig.Goude@edf.fr

Jean-Michel Poggi
Univ. Paris Descartes, France and
LMO, Univ. Paris-Sud Orsay, France
Jean-Michel.Poggi@math.u-psud.fr

***Abstract*—Electricity load forecasting is crucial for utilities for production planning as well as marketing offers. Recently, the increasing deployment of smart grids infrastructure requires the development of more flexible data driven forecasting methods adapting quite automatically to new data sets. We propose to build clustering tools useful for forecasting the load consumption. The idea is to disaggregate the global signal in such a way that the sum of disaggregated forecasts significantly improves the prediction of the whole global signal. The strategy is in three steps: first we cluster curves defining super-consumers, then we build a hierarchy of partitions within which the best one is finally selected with respect to a disaggregated forecast criterion. The proposed strategy is applied to a dataset of individual consumers from the French electricity provider EDF. A substantial gain of 16 % in forecast accuracy comparing to the 1 cluster approach is provided by disaggregation while preserving meaningful classes of consumers.**

***Index Terms*—load forecasting, time series, clustering, segmentation, smart meter data.**

I. Introduction

Electric load forecasting is a crucial activity for utility applications, such as unit commitment, market operations and target marketing. Because of the expansion of smart grid infrastructures, a lot of new prospects have come up recently, such as new data driven methods for their flexibility and their ability to automatic fit new datasets. Many applications coming from individual data analysis can be found in literature: load profiling, customer segmentation, dynamic structure of smart meter data and portfolio variations (losses and gains of customers).

One of the key statistical tools involved is the use of clustering methods. The purpose of clustering is to partition a dataset into one or more subsets called clusters, in such a way that each subset is as homogeneous as possible. The elements of a given cluster are then more similar to others in the same cluster as the elements of another one. The recent literature provides several survey papers: from a general one for clustering of time series [13] and a series of papers about clustering methods for electrical load patterns [7], [8], in smart grid environment [11], [20] as well as recent propositions about clustering analysis of residential electricity demand profiles, [15], [17] or data-based method for creating electricity load profiles [16].

There is a need for local electricity load forecasting at different levels of the grid. Bottom-up approaches, based on a two stage process combining clustering and forecasting methods, are a promising perspective. First, it consists in building classes in a population such that each class could be sufficiently well forecast but corresponds to different load shapes or reacts differently to exogenous variables like temperature or prices (see e.g. [12] in the context of demand response). The second stage consists in aggregating forecasts to forecast the total or any subtotal of the population consumption. For example, identify and forecast the consumption of a sub-population reactive to an incentive is an important need to optimize a demand response program.

Few papers consider the problem of clustering individual consumption for forecasting. In [9] short-term forecasting of residential building load is considered and in [1], clustering procedures are compared with respect to their forecasting performances of their corresponding bottom-up forecasts of the total consumption of 6 000 residential customers and SME in Ireland. Good performances are reached and it should be noted that the proposed clustering methods are defined quite independently of the model used for forecasting. Similarly, in [6] clustering algorithms are considered for long-term load forecasting.

In [14], a clustering method is proposed associating more closely hierarchical clustering and multi-linear regression models to improve the forecast of the total consumption of a French industrial subset. Even if a substantial gain is obtained for forecasting, the method requires a sufficiently long dataset (2 or 3 years) and the algorithm is computationally intensive. A recent paper [19] examines forecasting uncertainty in electricity demand using a two-steps GAM modelization and mentions that with the ongoing roll-out of smart metering in many countries worldwide alternative approaches for improving the demand forecast appear. Disaggregation in space is not the only one to consider, for example [5] focused in natural gas consumption and noticed that the consumption data for households and small commercial customers are available in many countries only as long-term sum meter readings, their disaggregation and possibly reaggregation to different time intervals can be useful.

We propose to build clustering tools useful for the two tasks simultaneously: clustering individual customers and forecasting the load consumption. The idea is to disaggregate the global signal in such a way that the sum of disaggregated forecasts significantly improves the prediction of the whole global signal. The general strategy is in three steps: first we cluster individual curves defining super-consumers, then we built a hierarchy of partitions within which a best one is finally selected with respect to a disaggregated forecast criterion.

The proposed strategies are applied to a dataset of individual consumers from the French electricity provider EDF. A substantial gain is provided by disaggregation while preserving meaningful classes of consumers. A key remark to assess the value of the methods is the fact that no prior information is necessary to obtain very interesting results ensuring the inherent flexibility for adapting quite automatically to new datasets.

The outline of the paper is as follows. In Section II, we develop the statistical setup on how to use wavelets for clustering and forecasting curves. Section III deals with the French national application. Section IV proposes a short discussion before closing the paper with the conclusion in Section V.

II. Methods

Among the promising new adaptive data driven methods, functional data analysis received a particular attention in the field of time series analysis where they are particularly appropriate to model physical quantities that are by essence continuous in time, like electricity demand. In this section we present three aspects intimately connected: the way to represent and compare functional data, the methods to cluster individuals characterized by functional variables and finally the nonparametric forecasting. The key mathematical ingredient used here for clustering and forecasting curves is wavelets.

A. *Curves, wavelets and related similarities*

Wavelets allow to cope with functional data (or curves) by hierarchically decomposing finite energy signals in a broad trend (the smooth part giving information about the mean level of the curve) plus a set of localized changes kept in the details parts giving information about the shape of the curve. Then wavelets are well suited for identifying highly discriminant local time and scale features. Two similarity measures are defined in [3]. For each curve, through its empirical orthogonal wavelet transform, the first one is simply the Relative Contributions (denoted by RC) of scales to the total energy based, in other words the distribution of energy across scales to compactly represent data. This can be sufficient to make the signals well distinguishable, at least at a first glance. The second similarity measure, through the index WER (for Wavelet Extended R) is more accurate and can be useful at the second stage. It is based on the whole time-scale representations and uses wavelet-coherence tools, a wavelet generalization of the Fourier sprectrum coherence to non stationary signals. Similarity measures combined with an efficient feature selection technique in the wavelet domain is then used within more or less classical clustering algorithms to effectively differentiate among high-dimensional populations.

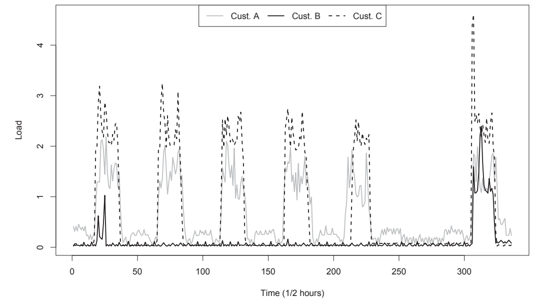

Fig. 1. Load curves of 3 customers during one week.

Figure 1 displays three load curves during one week coming from different individual customers[1]. As an example, we represent the multidimensional scaling (a classical visualization method) of both the RC based distance and the WER distance on Figure 2. If we take as a reference customer A, the RC distance leads to a similar dissimilarity to the other customers while the WER distance considers that customer C is much closer to customer A than he is to customer B.

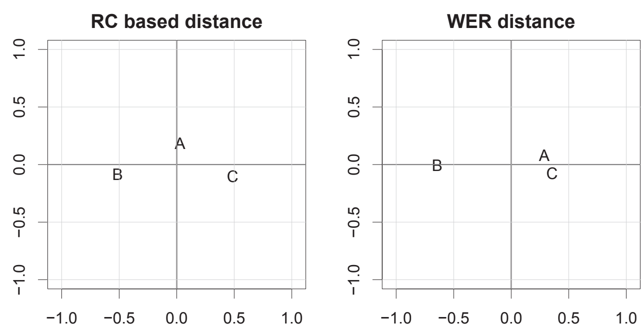

Fig. 2. Multidimensional scaling of 3 individual curves using RC based distance and WER distance.

B. *Clustering methods: PAM, CLARA and AHC*

Two types of clustering methods are usually distinguished: those based on iterative partitioning versus the methods building hierarchy of partitions. In both cases, a clustering method involves the choice of a measure of dissimilarity, a criterion of homogeneity, an algorithm, and the selection of the number of clusters. In our strategy, we consider two different clustering methods the first to get super-consumers and the second providing a hierarchy of super-consumers clusterings.

Due to the potentially large number of individuals to be classified and in order to be scalable, we propose first to use the partitioning algorithm CLARA (Clustering Large Applications, see [10]). It performs sequential clustering considering several randomly drawn sub-samples. For the first draw, m individuals are processed by a simple robust version of well-known k-means (called PAM standing for Partitioning Around

[1]Due to confidentiality issues we display on this example data from http://www.ucd.ie/issda/data/commissionforenergyregulationcer/.

Medoids, see [10]) to obtain k centers. Then $m-k$ individuals (different from the previous ones) are randomly drawn and merged with the k previously obtained centers and processed by a new PAM clustering. Thus, the algorithm is repeated preserving at each stage the previously obtained centers. On the latest iteration the obtained centers are the final ones. Finally to get an effective classification, each observation of the initial data set is assigned to its nearest center. The individuals are classified using the energy distribution across scales of whole curve and it is sufficient for this stage.

At the second stage, we aim at finding the super-consumers classes and we use an agglomerative hierarchical clustering (AHC) with the Ward method, providing a more flexible framework and a hierarchy of partitions. The algorithm starts from a distance (or dissimilarity) matrix between the observations which are in their own clusters at the beginning. Then, iteratively the nearest pairs of clusters are merged as one goes up the hierarchy. We will use two different notions of distance: the Euclidean distance calculated on the wavelet energy distribution across scales descriptors and the distance WER based on wavelet coherence.

C. Nonparametric forecasting using wavelets

In the context of economic seasonal univariate continuous time series, it is often natural to segment it in time, into consecutive curves, for example days, which are then treated as a discrete time series of functions. In particular, in the electrical context, the shape of the curves exhibits rich information about the calendar day type, the meteorological conditions or the existence of special electricity tariffs. Using the information contained in the shape of the load curves leads to very elegant formulation of functional forecasting.

The basic idea of nonparametric forecasting is that similar cases in the past have similar future consequences. For example the electricity consumption is divided into blocks of one day size. Then, using a dissimilarity measure, the blocks similar to the last observed block are searched in the past and a vector of weights is built. Finally, the forecast of the next day is obtained by a weighted average of the most similar future days using previous vector of weights. From the statistical point of view, the model is an estimate of the regression function using the kernel method, of the last block against all the blocks in the past. In [2] this basic model is extended to the case of stationary functional random variables. But in the context of electrical power demand, the hypothesis of stationarity generally fails: an evolving mean level and the existence of groups that may be seen as classes of stationarity are to be considered. Corrections to take into account these two main nonstationary features are considered in [4] defining a flexible nonparametric function-valued forecast model called KWF (*Kernel + Wavelet + Functional*) well suited to handle nonstationary series. The predictor can be seen as a weighted average of futures of past situations, where the weights increase with the similarity between the past situations and the actual one. Again the similarity is defined thanks to the wavelet decompositions of the two segments.

D. From individual curves to a hierarchy of partitions for forecasting: big picture

We present in this section the whole procedure to obtain a hierarchy of customers. Indeed we construct from N individual electrical demand records a hierarchical time series. On the top of the hierarchy we find the global demand and on the bottom the individual ones. Figure 3 represents this hierarchy where at some medium level we calculate the demand for K aggregates.

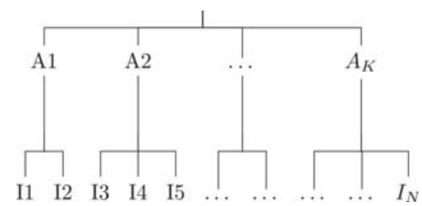

Fig. 3. Schematic representation of a hierarchy of customers.

For each customer we dispose with P records evenly sampled at a relatively high frequency (e.g. 1/4, 1/2 or hourly records) and maybe with noise. We adopt the strategy of describing this data as curves, i.e. functions of time.

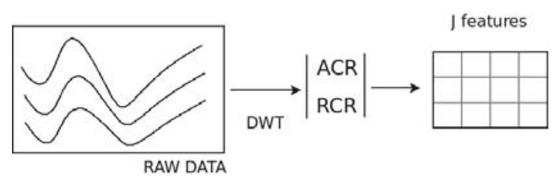

Fig. 4. From curves to matrices.

For this we project each curve into a wavelet basis using the Discrete Wavelet Transform (DWT) and we keep the projection coefficients. Then, we compute a handy number of features (in total $J = [\log_2(P)]$) known as the relative energetic contributions. By this we fall into the classical framework of data analysis where data is tabulated into a matrix with lines containing observations and columns containing variables (see Figure 4).

Now we begin with the proper clustering step. Since the time complexity of this step depends on the number of observation N and of variables P we decide to screen out irrelevant features using a feature selection algorithm for unsupervised learning described in [18]. Besides we produce first a coarse clustering by means of the CLARA algorithm. This allow us to construct a sufficiently large number, K' of prototypical customers that we call super-customers (SC). For each SC we can compute the synchrone demand, that is the direct sum of the individual demand of customers that belongs to the group at each time step of the sampling grid. Notice the parallelism with the initial situation, we have now K' coarsely aggregated demands over P records that can be seen as a discretized noisy sampling of a curve.

The aim of the second stage of the clustering is to aggregate the SC into a small number K of aggregates and to

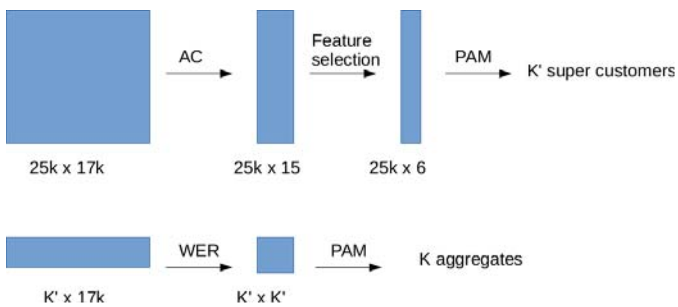

Fig. 5. Two step clustering.

construct a hierarchy. Instead of projecting curves and work with coefficients, we choose to use a notion of dissimilarity between objects of functional nature. Since the number of SC, K', is small enough, we can construct a dissimilarity matrix between the SC which is the input of the classical Agglomerative Hierarchical Clustering (AHC) which we use with the Ward link. The output of the AHC is the desired hierarchy of (super-)customers.

The scheme of this two step clustering procedure is summarized in Figure 5.

III. French national application

To optimize its production units each days, the French electricity provider EDF needs to forecast the electricity consumption of its portfolio (the sum of all its customers consumption) at a high accuracy. One of the different strategies to do that is to partition the consumers into classes, forecast the sum of consumption in each class and aggregate them to obtain the forecast of the total portfolio. This is called the bottom-up forecast. The complexity of this process stands in the fact that information concerning customers could be quite heterogeneous: from the real time consumption signal (customers equipped with smart meters) to a load profile and half-year consumption index, or exogenous information describing customers electricity usages.

In this section, we focus on a subset composed of big customers equipped with smart meters and consider the specific issue of forecasting the total consumption of this subgroup (the baseline) using only past electricity demands as entry of our forecasting model. We are interested here in how a given forecasting algorithm classically applied on the overall consumption of such a portfolio could be improved using a clever partitioning of customers and a bottom-up approach. We apply the proposed method to a dataset provided by EDF consisting in approximately 25000 half-hourly load consumption series over two years (2009-2010). The first year is used for partitioning and the calibration of our forecasting algorithm, then the second year is used as a test set to simulate a real forecasting use-case. Note that we represent normalized data (divided by the maximum) due to confidentiality reasons.

A. Numerical experiments

All our experiments are conducted on a standard laptop (64-bits Linux, 4 x 2.7Ghz processors, 8Gb RAM) using the statistical language R v.3.2.2.

The available dataset is composed of the records of 25011 clients at 15720 time points per year. A very few (less than 100) data points are missing for individual customers. Since we work with functional data techniques this is not a problem. After inspecting a summary of descriptive statistics not reported here, we decide to discard 1% of the curves exhibiting the most extreme variance values. We screen out those of these curves which are too flat (usually associated to null records of demand) or curves that are extremely fluctuating (often associated with measurement errors).

Then, each curve is transformed using the DWT with the *Symmlet 6* wavelet from the family of the least asymmetric wavelets with periodic boundary condition. After dropping the scale coefficients we compute the relative energetic contributions. Notice that up to here all the treatment, except the read of the file, can be easily parallelized.

At the end of this step we have the dataset structured as a matrix of about 24000 rows and 15 columns which represents a space economy of about 99% with respect to the original dataset. The new features are then used as input of a feature selection algorithm to screen out the least informative variables. As a result we keep only the 8 most energetic contributions associated to the medium range of the wavelet spectrum. In other words, the scales related to too low or too high frequencies are irrelevant to discover cluster structure.

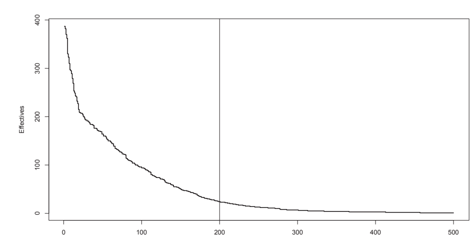

Fig. 6. Distribution of the number of observations per cluster after the CLARA procedure for 500 clusters. The vertical line at $K' = 200$ shows the final choice for the number of SC.

For the CLARA procedure we choose 5 sequential passes of PAM over 4000 observations each for K' classes. We first tried $K' = 500$ but it resulted in a too large number of unitary classes. Figure 6 depicts the distribution of the number of observations per cluster, which guides us to pick $K' = 200$ which is our final choice. The aggregates are then computed to obtain 200 load curves for the second stage clustering.

Now the curves are treated with the Continuous Wavelet Transform using the *Morlet* wavelet. We retain only scales related to frequencies below six months and then compute the Wavelet Extended R^2 between each pair of curves. The resulting matrix of dissimilarity is used on the ACH with the Ward link.

B. Focus on forecasting

To evaluate our approach we compare the day-ahead forecasts performance obtained by the KWF algorithm on the baseline of the 24000 customers to the ones obtained with a bottom-up approach on the different partitions. To measure this performance we use the (daily) MAPE (Mean Absolute Percentage Error), defined as

$$\text{MAPE}(y, \widehat{y}) = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{y_i - \widehat{y}_i}{y_i} \right|$$

where $y = (y_1, ..., y_n)$ is the target signal, $\widehat{y} = (\widehat{y}_1, ..., \widehat{y}_n)$ is the forecast, n the number of observations (here $n = 48$ corresponding to one day). We present the forecasting results in Figure 7. We clearly see that the forecasts obtained on the baseline (the horizontal line) are significantly improved by the complete partitioning (full gray line), reaching an optimum at 15 clusters. To situate our results, let us recall that the French national load daily forecasts have an accuracy of 1.5 % MAPE and that [1] achieve a 2.6 % MAPE on an Irish residential and SME dataset. They obtained an overall improvement with 10 clusters of 20 % over the 1 cluster approach. We observe a similar gain of 16 % corresponding to 15 clusters.

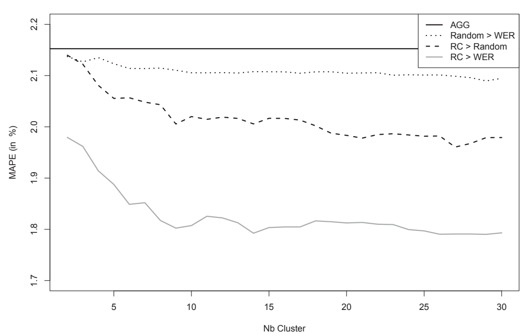

Fig. 7. Mean of the daily MAPE over the test set as a function of the number of clusters for the baseline forecast, i.e. without clustering (black solid line) and the cluster-based alternative (gray solid line). Dotted and dashed lines represent a tentative evaluation of the minimal contribution of disaggregation for each one of the two stages on the clustering.

In addition, the dashed curve is a tentative evaluation of the minimal contribution of disaggregation by starting from the super-consumers and aggregate them at random. Conversely, the dotted curve is the contribution of disaggregation by obtaining randomly the super-consumers and then aggregate them with the WER distance. The difference between the solid grey curve and the dashed one important and illustrates the additional reduction provided by the second step of clustering with respect to the first step leading to the definition of super-consumers which is already uniformly better than the baseline forecast.

C. *Focus on clustering*

We focus here on the 15 clusters obtained by our method and representing an optimum for forecasting the baseline. We exhibit with a few graphs the properties of the different clusters.

First, we note that our clustering approach produce a number of clusters which is close to the one commonly used for commercial purposes and based on exogenous information on the customers. Additionally, the mean number of observations per cluster is here around 1500, which is of the same order of magnitude than the commercial clusters.

One important feature of electrical data is their seasonal variation along the year, often due to meteorological variation but also here, as we deal with professional customers, with the business activities or industrial processes along the year. To clearly distinguish this effect within the 15 clusters, we smooth the total consumption in each cluster along time and represent it on Figure 8. We observe classical winter/summer patterns corresponding to electrical heating in the clusters (see e.g. 1 and 9) or cooling as in clusters (see e.g. 5 and 11). Some yearly patterns are heavily impacted by holidays like in cluster 3 for which we observe a large decrease of the load in august and during the winter holidays. On the other hand, clusters like 7 looks pretty flat along the year. Those different yearly shapes could be explained by customers activities, if they are industrial, business offices, located in a touristic area, etc.

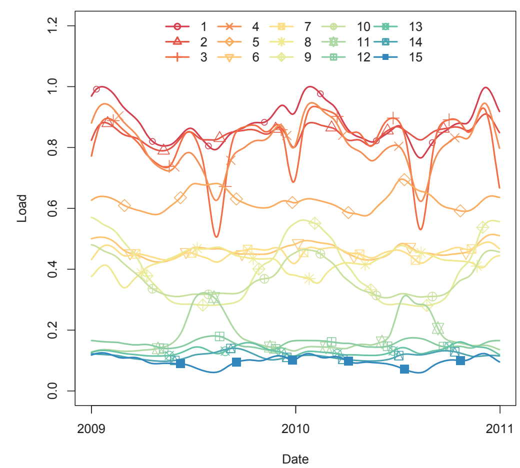

Fig. 8. Smoothed total consumption of the 15 clusters of the optimal partition in function of time.

Another interesting feature to look at is the weekly shapes of each cluster. We calculate the mean load of each cluster by hour of day and day of week and plot it on Figure 9. Here again some very different profiles are obtained, depending of the weekly activities of customers per cluster. We observe some classical day/night and week-days/week-end shapes (see e.g. clusters 1, 2, 3 and 4) where the consumption is higher during the days than at night and higher during week days than week ends, probably corresponding to office buildings. In contrast, we also observe profiles like 9 and 10 where the consumption is higher during the nights probably due to industrial activities and profile like 5 where the consumption is low only on Sunday meaning that Saturday is an active day.

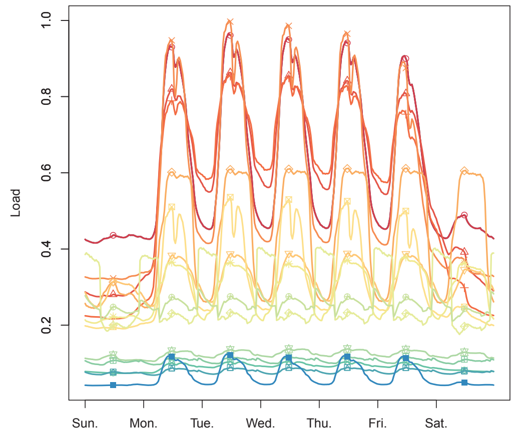

Fig. 9. Weekly profiles of the 15 clusters.

IV. Discussion

Concerning the clustering algorithm, CLARA is one choice among many others. The current research on data analysis for high volumes of data provides with a large span of variants that are able to cope with very large data sets. However, merging of the first and second stage is not straightforward since the computational burden of calculating the dissimilarity matrix. Note that for this we should arbitrate the trade-off between a large value for K' which gives more SC but hardens the calculation of the matrix.

To end this short discussion let us emphasize that the idea is to define a flexible method allowing sensitivity and adaptation to customer losses and changes in perimeter of the wavelet unsupervised classification of individual consumption by a hierarchy of partitions. This seems to be preferable to the inclusion, for the optimization purpose, of the forecasting in the cluster definition. Here forecasting refines disaggregation, since it operates at a different scale and the forecasting objective is introduced only for the selection of a partition in the hierarchy.

V. Conclusion

Load data if often seen as curves by electrical engineers but rarely by forecasters. The use of wavelets allows to efficiently compress curves while keeping discriminative information as it can be seen on the first stage clustering. There, the relative energetic contributions are shown to be effective in using this information while the size of the problem decreases dramatically from P ($\sim$ 17000 for the French dataset) to $[\log_2(P)]$ (= 15).

Another aspect of wavelets is used on the second stage of clustering. Instead of compressing the curves, the Continuous Wavelet Transform is used to produce a redundant description of the curve. The redundancy is translated into a mathematical representation much larger that is also computationally heavier to obtain (in particular due to the use of complex numbers). However, this dissimilarity seems to effectively detect structures of electrical demand which are both interpretable and able to enhance the forecast of a given prediction model.

As a perspective, we suggest to apply our method to bigger datasets from smart meters deployment to study both the scalability and the generalization of this approach to residential data.

References

[1] C. Alzate and M. Sinn, Improved electricity load forecasting via kernel spectral clustering of smartmeter, *International Conference on Data Mining*, vol. 948, pp. 943 – 948, 2013

[2] A. Antoniadis, E. Paparoditis and T. Sapatinas, A functional wavelet-kernel approach for time series prediction, *Journal of the Royal Statistical Society, Series B*, vol. 68(5), pp. 837 – 857, 2006

[3] A. Antoniadis, X. Brossat, J. Cugliari, and J.-M. Poggi, Clustering functional data using wavelets, *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 11(1), 2013

[4] A. Antoniadis, X. Brossat, J. Cugliari, J.-M. Poggi, Prévision d'un processus à valeurs fonctionnelles en présence de non stationnarités. Application à la consommation d'électricité Journal de la Société Française de Statistique, Vol. 153, No. 2, 52–78, 2012

[5] Brabec, M. and Konár, O. and Malỳ, M. and Kasanickỳ, I and Pelikán, E., Statistical models for disaggregation and reaggregation of natural gas consumption data, *Journal of Applied Statistics*, vol. 42(5), pp. 921–937, 2015

[6] Carević, S. and Capuder, T. and Delimar, M. Applications of clustering algorithms in long-term load forecasting *Proceedings Energy Conference and Exhibition (EnergyCon), 2010 IEEE International* 688–693, 2010

[7] G. Chicco Overview and performance assessment of the clustering methods for electrical load pattern grouping, Energy , 42, 68 – 80, 2012.

[8] Figueiredo, V., Rodrigues, F., Vale, Z., Gouveia, J. B. An electric energy consumer characterization framework based on data mining techniques. Power Systems, IEEE Transactions on, 20(2), 596–602, 2005

[9] Iwafune, Y., Yagita, Y., Ikegami, T., Ogimoto K. Short-term forecasting of residential building load for distributed energy management *Proceedings Energy Conference (ENERGYCON), 2014 IEEE International* 1197–1204, 2014

[10] Kaufman, L. and Rousseeuw, P Finding groups in data: An introduction to cluster analysis, Hoboken NJ John Wiley & Sons Inc, 1990

[11] J. Kwac, Flora, J., Rajagopal, R. Household Energy Consumption Segmentation Using Hourly Data Smart Grid, IEEE Transactions on, 5, 420–430, 2014

[12] Labeeuw, W., Stragier, J., and Deconinck, G. Potential of active demand reduction with residential wet appliances: A case study for Belgium. Smart Grid, IEEE Transactions on, 6(1), 315–323, 2015

[13] Warren Liao, T. Clustering of time series data–a survey Pattern recognition, 38(11), 1857–1874, 2005

[14] M. Misiti, Y. Misiti, G. Oppenheim, and J.-M. Poggi, Optimized Clusters for Disaggregated Electricity Load Forecasting, *REVSTAT – Statistical Journal*, vol. 8(2), pp. 105 – 124, 2010

[15] Mutanen, A., Ruska, M., Repo, S., Jarventausta, P. Customer classification and load profiling method for distribution systems. Power Delivery, IEEE Transactions on, 26(3), 1755–1763, 2011

[16] T., Räsänen, D., Voukantsis, H., Niska, K., Karatzas, M., Kolehmainen Data-based method for creating electricity use load profiles using large amount of customer-specific hourly measured electricity use data Applied Energy, 87(11), 3538–3545, 2010

[17] J.D. Rhodes, W.J. Cole, C.R. Upshaw, T.F. Edgar, M.E. Webber Clustering analysis of residential electricity demand profiles Preprint submitted to Applied Energy, March 18, 2014

[18] D. Steinley and M. Brusco, A new variable weighting and selection procedure for k-means cluster analysis. *Multivariate Behavioral Research*, 43:32, 2008.

[19] Wijaya, T. K., Sinn, M., and Chen, B., Forecasting Uncertainty in Electricity Demand, *AAAI-15 Workshop on Computational Sustainability, EPFL-CONF-203769*, 2015

[20] K. Zhou, S. Yang, C. Shen A review of electric load classification in smart grid environment, Renewable and Sustainable Energy Reviews, 24, 103 – 110, 2013.