

# Clustering Analysis of the Electrical Load in European Countries

Ankit Kumar Tanwar

Department of Electrical Engineering  
Delhi Technological University,  
Delhi, India

Emanuele Crisostomi

Pietro Ferraro  
Marco Raugi  
and Mauro Tucci  
Department of Energy, Systems, Territory  
and Constructions Engineering  
University of Pisa, Italy

Giuseppe Giunta

eni S.p.A  
midstream department  
San Donato Milanese, Italy

**Abstract**—In this paper we used clustering algorithms to compare the typical load profiles of different European countries in different day of the weeks. We find out that better results are obtained if the clustering is not performed directly on the data, but on some features extracted from the data. Clustering results can be exploited by energy providers to tailor more attractive time-varying tariffs for their customers. In particular, despite the relevant differences among the several compared countries, we obtained the interesting result of indentifying a single feature that is able to distinguish weekdays from holidays and pre-holidays in all the examined countries.

## I. INTRODUCTION

### A. Motivation

There are two main motivations to accurately analyze electrical load consumption data: the first reason is to make the power grid operate in a more efficient way, possibly close to real-time. In fact, the recent increase of power generated from renewable sources, mainly PhotoVoltaic (PV) plants and wind farms, has in turn increased the quantity of power that is injected in the grid in a non-dispatchable and fluctuating manner. On the other hand, conventional power plants (e.g., thermo-electric power plants) are continuously maintained switched on at a low-level, as a back-up to match the energy demand, if needed (i.e., if they were switched off, then it would take a too long time before they could effectively be used as a back-up). Such an operation is very expensive, especially as it is usually very rare that they are used in practice. In this framework, the ability to predict the load in an accurate manner would clearly provide a helpful information to the energy suppliers to schedule the operation of power plants in a more efficient, and economically convenient, way.

There is also a second motivation to learn the patterns of electrical load consumption. The electricity market is in fact shifting towards a more dynamic, liberalized and competitive environment; the number of Distributed System Operators (DSOs) and energy retailers is increasing inside single countries, and have also started competing in other national markets. In such a context, energy suppliers have started to offer more diversified energy tariffs, following the trend of diversified bills in other more-established fields (e.g., telecommunications). From this perspective, the ability to know the typical load consumption patterns of

different classes of customers can be used by energy suppliers to offer tailored, and in principle more attractive, energy tariffs.

Accordingly, the objective of this paper is to analyze the consumption data of some European countries of different sizes and of different latitudes. While it is clear that the load is different for a number of reasons, namely, other than size and geographical reasons, there are also different levels of the industrial sector, and use of electrical energy for different purposes (e.g., heating), an interesting finding of this paper is that there are also some important common patterns. In particular, as it will be further elaborated in the paper, some features are equivalently informative in all the considered countries.

This paper is organized as follows: the remainder of the Section is dedicated to explain why we are interested in applying clustering algorithms to the load data, and to review the current state of the art in this specific application. Section II introduces the used data-base, provides some initial insight on the available data, and shortly illustrates the used clustering algorithms and the adopted performance indices. Section III thoroughly compares the hourly load patterns among the considered countries in different cases. Finally, we summarize our findings and outline our current ongoing research in Section IV.

### B. Clustering

Classification and clustering of time series signals is an important area of research in several fields, such as economics, engineering, finance, medicine, biology, physics, geology, and many others. Clustering refers to the ability to aggregate similar objects together, and the basic clustering operation corresponds to take a set of  $N$  objects and group them into  $K$  clusters. There are three main motivations for doing so (from [1]):

(i): First, a good clustering has predictive power; in this case, we perform clustering because we believe the underlying cluster labels are meaningful, will lead to a more efficient description of our data, and will help us choose better actions. This type of clustering is sometimes called “mixture density modelling”, and the objective function that measures

how well the predictive model is working is the information content of the data. In the load consumption example, this property suggests that clusters can be used to predict future energy consumption as well;

(ii): Secondly, clusters allow people to compress the information into a single information, corresponding to the center of the cluster, i.e., centroid. For instance, by classifying load consumption into two categories (working days and holidays), it is possible to identify two clusters and their two corresponding centroids. The centroids can be used to identify the “typical” consumption of that day of the week. Thus, it summarizes in only one 24-hour profile the information content of the load profiles during the, say, whole year. This type of clustering is sometimes called “vector quantization”;

(iii): A third reason to make clusters is to identify the “outliers”, i.e., the cases in which clusters fail to accurately represent particular data. An example of this, in our load consumption case, is represented by working days that for some reasons present load consumptions that are very close to holidays (e.g., working days in the middle of two holidays). Clearly, such anomalous days should be identified and not considered when building the profile of typical working days.

### C. State of the Art

The adoption of clustering techniques to analyze load data is not fully novel: for instance, reference [2] analyses the load profiles of a representative sample of Spanish residential users, by means of dynamic clustering (i.e., dynamic in the sense that the load profiles are interpreted as a time-series database). A similar attempt to classify electricity customers has been previously performed, among others, in [4], in the case of 234 non-residential customers in Italy connected to the Medium Voltage (MV) distribution system. One of the objectives of [4] is to compare some clustering algorithms, but the results (i.e., the habits of the customers) are not interpreted. In [3], a clustering analysis is performed for the specific case of a building in a university campus in Greece. In some references, clustering analysis is mainly performed as a basis for load forecast. Among other, we remind references [5]-[8]. In [9], the authors find out that also aggregations of several consumers can be convenient to improve forecasting accuracy. Other interesting very recent related works include [10], where the authors use clustering algorithms to analyse the electrical load profile and for peak load assessment; [11] where load patterns are identified using an initial set of centroids specified according to a user-defined centroid model; [12] where the residential electrical load is modelled using mixture model clustering and Markov models; and finally we mention [13] where neuro-fuzzy classification methods are used to monitor the load in a non-intrusive fashion.

The contribution of this paper with respect to the others previously mentioned is that here clustering algorithms are used to compare the different load profiles in different European countries with the specific objective of identifying some common patterns. In doing so, we identify a specific feature, namely, “morning slope”, that is particularly informative in all the selected countries in distinguishing days

belonging to working days, pre-holidays and festive days. This result is rather surprising since the selected countries present load data that are completely different due to the different sizes of the countries, different latitudes, and different uses of energy (e.g., France, where electrical energy is also used for heating purposes). A second result is the identification of the typical load consumption profile in different days of the week in the selected country, which is an interesting curve that can be used, for instance, by energy retailers or suppliers to tailor tariffs upon typical consumption habits. This work is a natural extension of [14], where clustering techniques were used to analyze only the Italian electrical load.

## II. DATA ANALYSIS AND ALGORITHMS

### A. Data Set

We used the electrical load data freely available from the ENTSO-E database<sup>1</sup>. ENTSO-E is the European Network of Transmission Systems Operators for Electricity. The ENTSO-E statistical database includes a range of historical data sets regarding power systems of ENTSO-E member Transmission Systems Operators (TSOs). Following the merging of former TSOs associations in 2009, ENTSO-E has become the single data competence centre of the European electricity transmission systems; in particular, 41 TSOs from 34 countries are members of ENTSO-E. Among other data, the ENTSO-E database provides hourly and monthly consumption *aggregated* data for each country (i.e., the whole national load value is given). Note that according to the EC Regulation no. 1228/2003, the national TSOs have to obligatorily communicate the data relating the electrical physical flows in transmission systems operators’ networks.

In our analyses, we downloaded daily data of six countries, namely, Italy, France, Germany, Belgium, United Kingdom and Ireland, from year 2010 to 2013<sup>2</sup>. The countries were chosen in order to cover most latitudes in Europe (i.e., from Southern to Northern Europe), different sizes, and different electrical loads, as it will be further illustrated in Section II-B.

### B. Electrical Load Data

Figure 1 illustrates the electrical load data in year 2013 in the six considered countries.

As can be seen from the figure, there are major differences among the different electrical loads, apart from the obvious difference of the average magnitude of the load which depends, among other things, on the size of the countries. In particular, it is possible to note that:

- The electrical load in France is particularly large in winter days. This is due to the fact that electrical energy is also used for heating, as an alternative to gas which is the conventional fuel in (most) of the other European countries;
- The electrical load is particularly high in Italy in summer days due to air conditioning. The same effect

<sup>1</sup><https://www.entsoe.eu/data/data-portal/consumption/Pages/default.aspx>

<sup>2</sup>At the time when the manuscript was prepared, complete data from year 2014 was not available yet for all countries

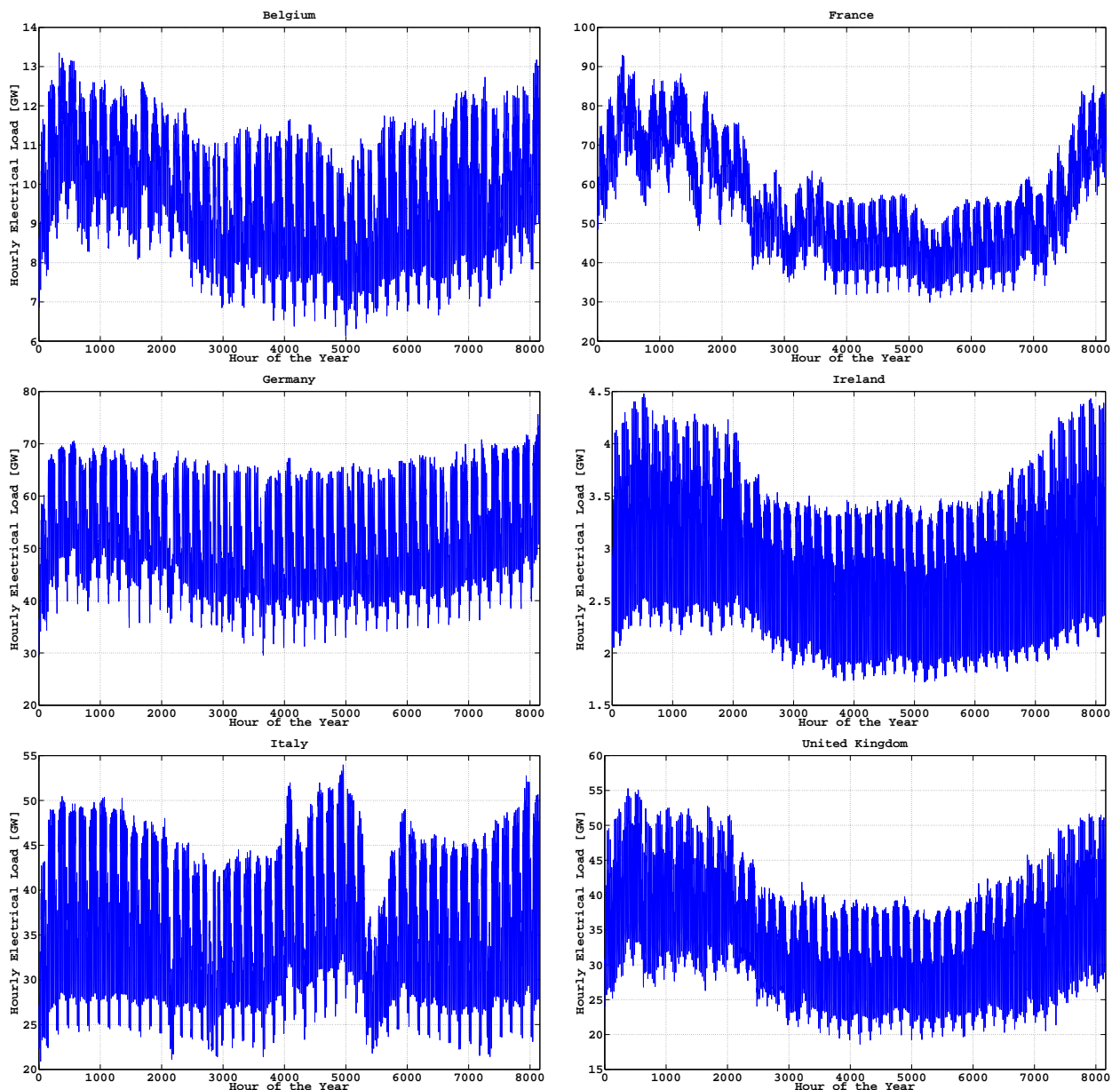


Fig. 1. Hourly electrical load in the different considered European countries in the whole year 2013.

is not as clear in other considered countries, due to the fact that the weather is not as hot as in Italy; also, note that the load is dramatically low for a couple of weeks around August 15, this is due to the fact that almost all companies and offices close for holidays;

- The electrical load is practically constant throughout the year in Germany. This is due to the fact that a large component of the load is the industrial load which is not affected too much by seasonal patterns;

### C. Fuzzy c-Means

We used a clustering algorithm to partition  $N$  data points in a 24-dimensional space (i.e., a daily load) into  $K$  classes, where  $N$  and  $K$  depend to the application of interest. There are many ways to do so, and we decided to adopt one of

the most popular clustering algorithms: Fuzzy c-Means. The fuzzy c-Means (FCM) algorithm is a soft version of the so-called K-means algorithm. In the K-means algorithm,  $K$  initial points are randomly chosen as the center of the  $K$  classes, denoted as centroids. Then an *assignment step*, where each point is assigned to the nearest centroid, and an *update step*, where the centroids are adjusted to match the means of their data points, are performed iteratively. The K-means algorithm always converges to a fixed point [1]. A drawback of the K-means algorithm, is that points are assigned to exactly one cluster, and all points within a cluster equally belong to that cluster. One way to circumvent such a criticism, is to make the algorithm softer, and associate each points with a degree to quantify the membership to each cluster. The FCM is an example of such an algorithm, and it was initially proposed by Dunn [15] and then refined by Bezdek [16]. It can be shown

that basic fuzzy C-means algorithms correspond to maximum-likelihood algorithms for fitting a mixture of Gaussians to data [1].

#### D. Performance Index

It is well known, and also rather obvious, that the electrical consumption load has a daily pattern that is repeated every 24 hours. However, it is also known that this is not the only pattern, and the electrical load does also have a weekly pattern, which strongly depends on the specific day of the week, namely, working days vs. holidays. As an illustrative example, Figure 2 compares the electrical load in two weeks of year 2013, from April 15 to April 28. The two weeks were arbitrarily chosen within the year, just for the sake of comparison, in such a way that no holidays occur within the week in any of the selected countries for the period under exam. In Figure 2, weekdays are shown in black, while Saturdays are shown in blue and Sundays are in red. As

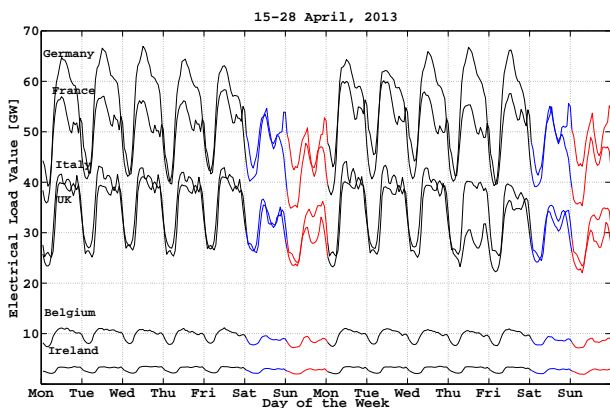


Fig. 2. Daily load in the six countries. Daily patterns (i.e., weekdays vs. holidays) are clearly visible.

can be seen from the Figure, the load is typically lower on Sundays (i.e., when many commercial activities and some industries are closed), and is higher during the weekdays. Some intermediate behaviour is sometimes obtained on Saturdays.

Thus, the objective of this paper is to test the ability of a clustering algorithm, namely, the fuzzy c-Means explained in the previous section, to automatically classify daily profiles as belonging to a weekday or to a holiday. In some cases, we shall also be interested in further split them into three classes, i.e., weekdays, holidays and pre-holidays (i.e., non-festive days before holidays). In particular, we are interested in comparing the results that are obtained in each of the six selected countries. Then, the performance index corresponds to the number of days belonging to a given cluster that have been automatically classified in the correct cluster. The “correct” classification is thus taken from the calendar. At this regard, note that different countries have a different set of festive days, for instance, Figure 3 shows the Italian calendar of year 2013, where the days have been classified as festive, holidays and pre-holidays.

For the reasons listed above, being interested in finding two or three hard partitions (namely, weekdays, holidays and pre-holidays) we use a basic defuzzification step after

running the FCM algorithm, not exploiting the membership degree; this information will be further explored in a future work.

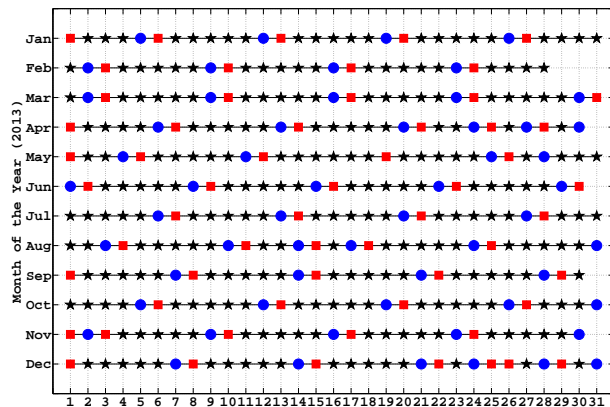


Fig. 3. Working (black stars) vs. pre-holidays (blue circles) vs. holidays (red squares) in Italy in 2013.

### III. CLUSTERING RESULTS

#### A. Direct Load Clustering

Table I shows what happens if one runs the FCM algorithm directly on the raw data, or on the normalized data<sup>3</sup>. The objective is to obtain two clusters of data, one corresponding to holidays and pre-holidays, and another one corresponding to working weekdays. As can be seen from Table I, pre-

TABLE I. ABILITY OF THE CLUSTERING ALGORITHMS TO ASSOCIATE A GIVEN LOAD PROFILE WITH THE CORRESPONDING DAY IN YEARS 2011-2013 (AS A PERCENTAGE OF THE OVERALL NUMBER OF DAYS)

	2011-13	Raw Data	Normalized Data
Belgium		73.03%	71.53%
France		52.84%	52.84%
Germany		93.70%	92.95%
Ireland		57.84%	57.97%
Italy		92.81%	89.60%
UK		57.15%	57.22%

processing of the data does not seem to influence the results of the clustering. Also, it can be seen that the direct clustering is already quite effective on both Italy and Germany, while the clustering approach classifies data in a different manner in the other countries. In particular, from visual inspection of the results, it can be seen that the data are not automatically clustered upon the day of the week, but rather upon the average value of the load, which is heavily affected by some seasonality effects.

#### B. Feature-based Clustering

The results of Table I can be improved if the clustering procedure is performed on some features extracted from the data, rather than from the data itself. In this work, we consider the features listed in Table II. Some of them are conventional choices in data analysis, others are known to be interesting for the specific application of interest. Taking one feature at a time into consideration, the cluster analysis

<sup>3</sup>here, by normalized we intend that the data are pre-processed to have a zero mean and a unitary standard deviation

TABLE III. 2-CLUSTERS CLUSTERING UPON SELECTED FEATURES RATHER THAN ON THE LOAD DATA IN YEARS 2011-2013 (AS A PERCENTAGE OF THE OVERALL NUMBER OF DAYS)

Feature/Country	Belgium	France	Germany	Ireland	Italy	UK
<b>Daily Load</b>	69.95%	52.57%	91.17%	57.02%	90.42%	56.82%
<b>Daily Mean</b>	69.95%	52.57%	91.17%	57.02%	90.42%	56.81%
<b>Daily Variance</b>	68.79%	79.26%	84.12%	85.42%	82.89%	66.32%
<b>Min-max</b>	80.08%	79.26%	87.54%	61.26%	87.82%	55.85%
<b>Max Peak</b>	75.56%	52.70%	94.59%	54.14%	91.92%	55.85%
<b>Peak Hour</b>	60.57%	64.75%	62.35%	62.56%	59.14%	59.48%
<b>Morning Slope</b>	96.37%	96.10%	96.71%	93.98%	94.69%	95.07%
<b>Night Slope</b>	57.70%	59.27%	78.71%	55.99%	60.64%	51.88%
<b>FFT</b>	81.38%	53.87%	95.21%	64.20%	94.66%	60.03%
<b>Kurtosis</b>	59.89%	54.07%	51.27%	57.70%	75.70%	68.31%
<b>Partial Daily Mean</b>	83.23%	53.46%	95.00%	79.19%	58.38%	58.38%
<b>Partial Daily Variance</b>	65.98%	59.55%	61.05%	72.96%	66.46%	69.20%
<b>Partial Min-max</b>	62.97%	57.36%	59.14%	68.38%	60.03%	66.67%
<b>Partial Min-diff</b>	60.78%	55.03%	50.92%	84.33%	62.15%	82.61%
<b>Partial Max-diff</b>	62.97%	63.66%	68.24%	59.62%	68.31%	59.00%
<b>Late Morning</b>	71.25%	56.61%	51.75%	84.12%	69.20%	83.23%
<b>Late Afternoon</b>	63.38%	58.04%	69.40%	53.53%	75.02%	53.73%
<b>FFT Peak</b>	69.95%	52.57%	91.17%	57.02%	90.42%	56.81%

TABLE II. LIST OF INTERESTING FEATURES

Feature	Definition
<b>Daily Load</b>	Sum of daily load values
<b>Daily Mean</b>	Mean of daily load values
<b>Daily Variance</b>	Variance of daily load values
<b>Min-max</b>	Difference between the maximum and the minimum value of the daily load
<b>Max Peak</b>	Maximum value of the daily load
<b>Peak Hour</b>	Hour of the day at which we have the maximum value of the daily load
<b>Morning Slope</b>	Difference between the load value at 10.00 am and at 06.00 am
<b>Night Slope</b>	Difference between the load value at 11.00 pm and at 09.00 pm
<b>FFT</b>	Sum of the absolute values of the Fast Fourier Transform of the daily load values
<b>Kurtosis</b>	Kurtosis of daily load values
<b>Partial Daily Mean</b>	Mean of daily load values between 11.00 am and 08.00 pm
<b>Partial Daily Variance</b>	Variance of the daily load values between 11.00 am and 08.00 pm
<b>Partial Min-max</b>	Difference between the maximum and the minimum value of the load between 11.00 am and 08.00 pm
<b>Partial Min-diff</b>	Difference between the average load and the minimum load between 11.00 am and 08.00 pm
<b>Partial Max-diff</b>	Difference between the maximum load and the average load between 11.00 am and 08.00 pm
<b>Early Afternoon</b>	Difference between the load value at 11.00 am and at 03.00 pm
<b>Late Afternoon</b>	Difference between the load value at 09.00 pm and at 06.00 pm
<b>FFT Peak</b>	Maximum of the absolute values of the Fast Fourier Transform of the daily load values

is iterated to obtain a significance scoring of each feature. Accordingly, Table III shows the performance of clustering upon the features previously listed. As can be seen from Table III, feature ‘‘Morning Slope’’ is the most informative feature in each of the six selected country, and in each case it manages to infer the correct day of the week around 95-96% of the times.

**Comment :** Note that it might be pointless to try to improve the previous results up to, say, 100 %. The reason is that some days of the year, though belonging to a certain class just from the point of view of the calendar, still, in every day life, they correspond to another class. As an example, in Italy practically every single office and most industries are closed for two weeks around August 15; thus, all the weekdays appear as ‘‘working days’’ from the calendar, but

it is well known that in truth they are holidays. Obviously, the load follows the real life trend, rather than the calendar. A similar pattern can be noticed during Christmas holidays in most countries. Therefore, the only way to increase the matching probability closer to 100 %, is to ‘‘redefine’’ the calendar according to people’s habits. Since any choice in this direction can be interpreted as arbitrary, and as in any case such operations require some knowledge of the habits of people in different countries, in this paper we simply used the calendar as a term of comparison. However, one should be aware that reaching higher matching probabilities would be hard, and under some points of view, even wrong.

### C. Three-cluster Clustering

In this section, we are interested in investigating the ability of the clustering algorithms to directly classify data into weekdays, pre-holidays and holidays. This task is considerably more complicated. As previously mentioned, some working days should be in truth treated as festive days; thus, the difference between pre-holidays/holidays and weekdays becomes even more subtle. In this section we compare four possible strategies: (i) a direct clustering procedure that separates the raw data in three classes; (ii) a direct clustering procedure that separates the data in three classes based on some feature; (iii) a hierarchical clustering procedure that first separates the raw data into two classes as before, working days and non-working days, and then further separates the non-working days into holidays and pre-holidays; and (iv) the same hierarchical procedure of before performed on the features rather than on the raw data. Table IV illustrates the obtained results, but due to space limits we only report results for a single (most informative) feature. As can be noted in Table IV, hierarchical clustering provides better results than direct clustering, when taking features into considerations, while it performs quite poorly for what regards raw data. Consistently with the 2-clusters clustering ‘‘Morning Slope’’ appears to be the most informative feature, performing slightly worse than the previous case. On the other hand, other, and generally different, features are the most informative in case of direct clustering. Obviously, results could be improved if the clustering was based on a set of features, rather than on a single feature, and preliminary analysis show that combinations of two or three features perform dramatically better in the case of direct

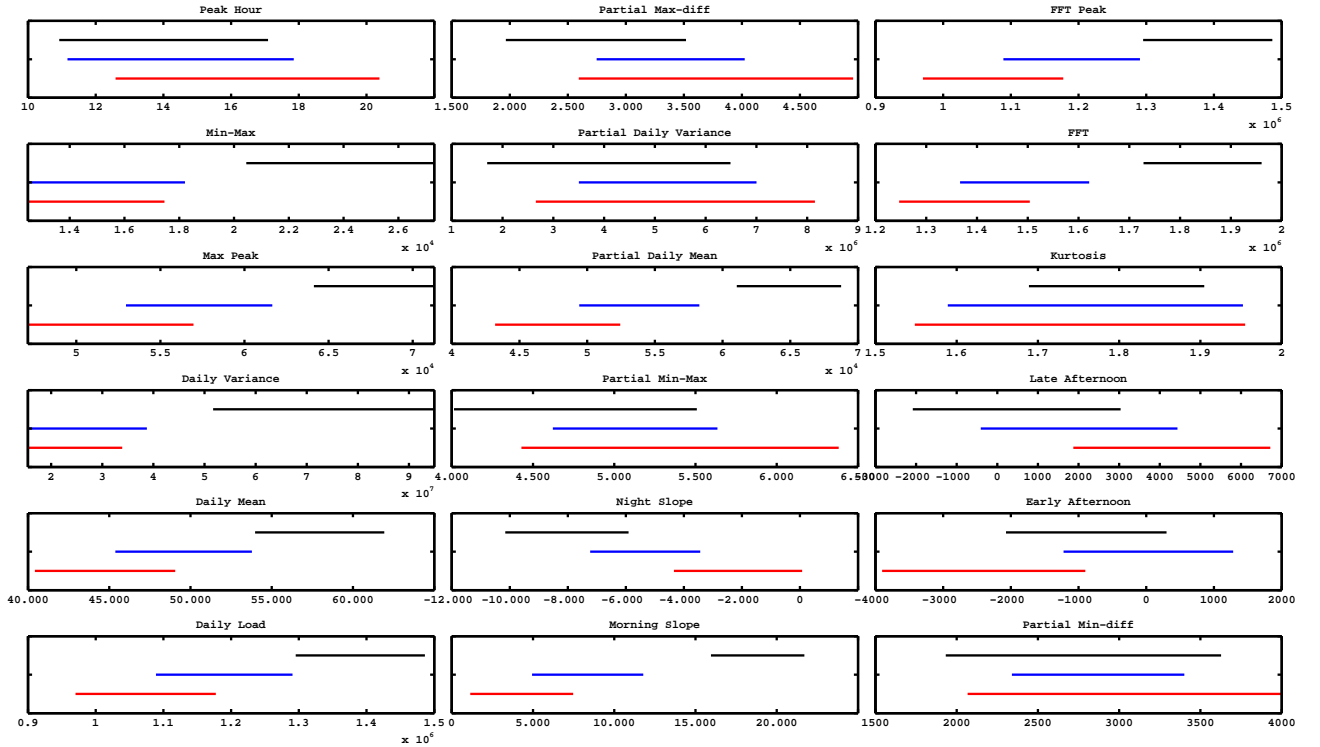


Fig. 4. Average and standard deviation of the features in Germany in the different days of the calendar. Top (black) lines refer to weekdays, middle (blue) line refers to pre-holidays, and the bottom red line refers to holidays.

TABLE IV. ABILITY OF THE CLUSTERING ALGORITHMS TO ASSOCIATE A GIVEN LOAD PROFILE WITH THE CORRESPONDING DAY IN YEARS 2011-2013 (AS A PERCENTAGE OF THE OVERALL NUMBER OF DAYS)

Country / Method	Direct Raw Data	Direct Feature	Hierarchical Raw Data	Hierarchical Feature
Belgium	52.57%	53.18% (Peak Hour)	42.44%	90.62% (Morning Slope)
France	39.08%	58.45% (Morning Slope)	41.34%	94.79% (Morning Slope)
Germany	54.28%	59.14% (Max Peak)	63.72%	90.07% (Morning Slope)
Ireland	53.25%	63.18% (Part. Daily Var.)	38.53%	90.82% (Morning Slope)
Italy	54.96%	76.52% (Kurtosis)	63.24%	93.50% (Morning Slope)
UK	51.75%	73.58% (Kurtosis)	38.33%	89.32% (Morning Slope)

clustering, with percentages up to 90% and slightly better in the case of hierarchical clustering with percentages up to 95% for each country; this aspect will be further investigated in future works.

For the sake of completeness, in Figure 4, we show how informative every single feature is, in the case of a particular country, here Germany. Figure 4 shows the mean value and the standard deviation of the features of interest in the German case, in the period 2010-2013. In each subgraph, a line is centered on the mean value of that given feature

in the corresponding set of days (top, black, working days; middle, blue, pre-holidays; bottom, red, holidays). Clearly, if the lines are horizontally spatially separated, this implies that the corresponding feature is expected to be useful in identifying the different days of the week. As anticipated in Table IV, Figure 4 confirms that the “Max Peak” and the “Morning Slope” are two of the most valuable feature in the case of Germany.

#### IV. CONCLUSION

As we had initially anticipated, one of main advantages of clustering lies in summarizing information in a single data-vector (i.e., the centroid of the cluster) which corresponds to some average consumption profile which is typical of a specific country. Such an information can be useful, for instance, for electrical energy providers to tailor their tariffs accordingly. Figure 5 shows the centroids of the two clusters (i.e., weekdays vs. holidays) for the six selected countries. We believe that the information shown in Figure 5 can be very useful for electrical energy suppliers and retails as an indication of some average profiles. This information can be useful, for instance, to plan optimal scheduling of dispatchable power plants to support power plants that exploit renewable energy to provide the required energy.



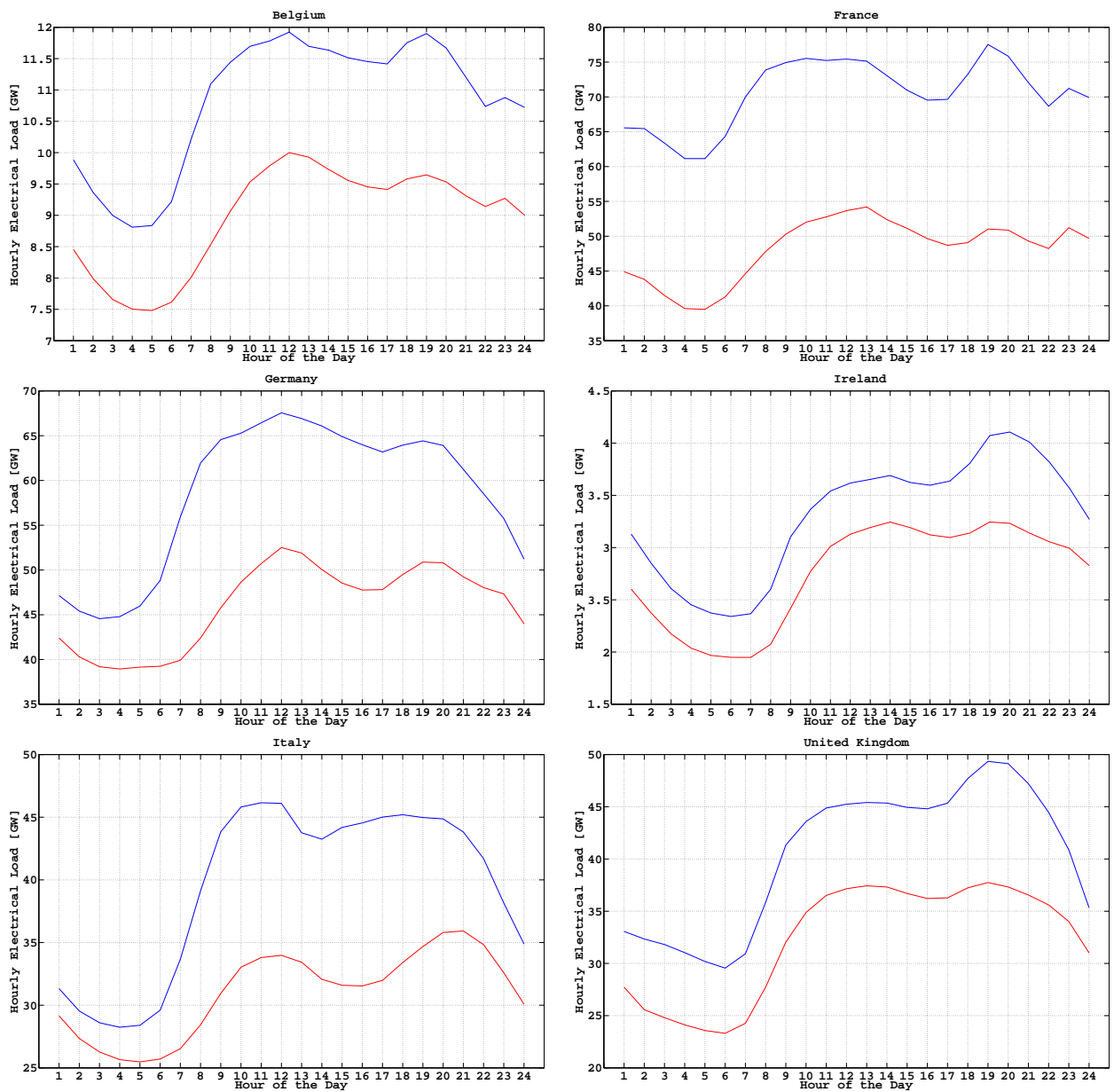


Fig. 5. Centroids of the clusters of working days and holidays in the six selected countries. In all figures, the top curve, in blue, corresponds to working days, and the curve below, in red, to holidays.

The results presented in this paper are focused on the aggregated national electrical energy consumption, and do not distinguish between, for instance, residential load and industrial load. An interesting line of research is to repeat the same approach taking only residential load into account, to design optimal tariffs. In this case, the most challenging task appears to be the ability to collect a significant amount of household consumption data, since they are usually not revealed for privacy reasons and, most important, for business and competitiveness reasons. Another interesting line of research focuses on developing automatic methods to identifying the most informative sets of features that would allow to better clusterise the data into three clusters, and this improve the results outlined in Table IV.

## V. ACKNOWLEDGEMENT

This paper is related to the University of Pisa project PRA 2015 "Metodi e tecniche innovative per l'Integrazione di Sistemi per l'energia Elettrica e Termica (MISSET)"

## REFERENCES

- [1] D.J.C. MacKay, *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, VI Edition, 2007.
- [2] I. Benítez, A. Quijano, J.-L. Díez and I. Delgado, *Dynamic clustering segmentation applied to load profiles of energy consumption from Spanish customers*, *Electrical Power and Energy Systems*, vol. 55, pp. 437-448, 2014.
- [3] I.P. Panapakidis, T.A. Papadopoulos, G.C. Christoforidis and G.K. Papa-  
giannis, *Analysis of the Electricity Demand Patterns of a Building in a*

University Campus, IEEE 12th International Conference on Environment and Electrical Engineering, 2013.

- [4] G. Chicco, R. Napoli and F. Piglion, *Comparisons among clustering techniques for electricity customer classification*, IEEE Transactions on Power Systems, vol. 21, no. 2, pp. 933-940, 2006.
- [5] G. Grigoraç, F. Scarlatache and G. Cârțină, *Load estimation for distribution systems using clustering techniques*, IEEE International Conference on Optimization of Electrical and Electronic Equipment, 2012.
- [6] Y.J. Xia, Y.H. Yang, F. Ge, J. Su and H. Yu, *Pattern Analysis for Load Forecasting*, IEEE 8th International Conference on Computing Technology and Information Management (ICCM), 2012.
- [7] V.S. Kodogiannis and I. Petrounias, *Power Load Forecasting Using Adaptive Fuzzy Inference Neural Networks*, 6th IEEE International Conference on Intelligent Systems, 2012.
- [8] I.P. Panapakidis, M.C. Alexiadis, and G.K. Papagiannis, *Load Profiling in the Deregulated Electricity Markets: A Review of the Applications*, 9th IEEE International Conference on the European Energy Market, 2012.
- [9] S. Humeau, T.K. Wijaya, M. Vasirani and K. Aberer, *Electricity Load Forecasting for Residential Customers: Exploiting Aggregation and Correlation between Households*, 3rd IFIP Conference on Sustainable Internet and ICT for Sustainability, Palermo, Italy, 2013.
- [10] D.D. Sharma and S.N. Singh, *Electrical Load Profile Analysis and Peak Load Assessment using Clustering Technique*, IEEE PES General Meeting, National Harbor, MD, 2014.
- [11] G. Chicco, O.-M. Ionel and R. Porumb, *Electrical Load Pattern Grouping Based on Centroid Model With Ant Colony Clustering*, IEEE Transactions on Power Systems, vol. 28, no. 2, pp. 1706-1715, 2013.
- [12] W. Labeeuw and G. Deconinck, *Residential Electrical Load Model Based on Mixture Model Clustering and Markov Models*, IEEE Transactions on Industrial Informatics, vol. 9, no. 3, pp. 1561-1569, 2013.
- [13] Y.-H. Lin and M.-S. Tsai, *Non-Intrusive Load Monitoring by Novel Neuro-Fuzzy Classification Considering Uncertainties*, IEEE Transactions on Smart Grid, vol. 5, no. 5, pp. 2376-2384, 2014.
- [14] L. Semeraro, E. Crisostomi, G. Giunta, A. Landi, M. Raugi and M. Tucci, *Electrical Load Clustering: the Italian case*, IEEE PES Conference on Innovative Smart Grid Technologies (ISGT) Europe, Istanbul, Turkey, 2014.
- [15] J.C. Dunn, *A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters*, Journal of Cybernetics, vol. 3, no. 3, 1973.
- [16] J.C. Bezdek, *Pattern recognition with fuzzy objective function algorithms*, Plenum Press, New York, 1981.