Proceedings Paper:

Roberts, D. and Brown, S.F. orcid.org/0000-0001-8229-8004 (2020) Identifying calendar-correlated day-ahead price profile clusters for enhanced energy storage scheduling. In: Cruden, A., (ed.) Energy Reports. 4th Annual CDT Conference in Energy Storage & Its Applications, 09-10 Jul 2019, Southampton, U.K.. Elsevier , pp. 35-42.

https://doi.org/10.1016/j.egyr.2020.02.025

Available online at www.sciencedirect.com
ScienceDirect
Energy Reports 6 (2020) 35–42
www.elsevier.com/locate/egyr

4th Annual CDT Conference in Energy Storage and Its Applications, Professor Andrew Cruden, 2019, 07-19, University of Southampton, U.K

Identifying calendar-correlated day-ahead price profile clusters for enhanced energy storage scheduling

Diarmid Roberts, Solomon F. Brown*

Department of Chemical and Biological Engineering, University of Sheffield, UK

Received 12 February 2020; accepted 25 February 2020

Abstract

Optimising the scheduling of energy storage systems with respect to multiple revenue streams is crucial to the business case for installations in the UK and other countries with high electrical grid penetration. In this work we use hierarchical clustering for the first time to correlate groupings of UK day-ahead electricity price profiles with calendar period. We observe that there are three primary clusters in the 2017–2019 dataset, and hypothesise that these arise from the interplay of winter/summer variations in demand along with longer term variations in the wholesale gas price. Looking at finer detail, we find that in summer 2018 there is a clear split in weekday/weekend price profiles, with the latter showing a significantly delayed price peak, and higher night time prices. These findings demonstrate the usefulness of the approach for revenue stacking, as the optimal bidding for ancillary services to fit around the performance of peak shaving will be influenced by the knowledge of such patterns, especially when the horizon for bidding is beyond the day ahead.

Peer-review under responsibility of the scientific committee of the 4th Annual CDT Conference in Energy Storage and Its Applications, Professor Andrew Cruden, 2019.

Keywords: Hierarchical clustering; Scheduling; Optimisation; Energy storage; Revenue stacking

1. Introduction

Along with increased interconnection and flexible demand, energy storage is one solution to the problem of intermittency in the output of renewable power generation. The rollout of energy storage on a large scale is restricted in many industrialised nations by the difficulty of paying back the initial capital investment within an acceptable time period. For this reason revenue stacking, whereby the energy store performs different tasks in different periods, has recently been a subject of much research. Depending on where the ESS is located, a selection of the following benefits will apply:

- Arbitrage on wholesale energy price (peak shaving)

* Corresponding author.
E-mail address: s.f.brown@sheffield.ac.uk (S.F. Brown).

https://doi.org/10.1016/j.egyr.2020.02.025

Peer-review under responsibility of the scientific committee of the 4th Annual CDT Conference in Energy Storage and Its Applications, Professor Andrew Cruden, 2019.

- Avoidance of use of network charges (peak shaving)
- Provision of ancillary services
- Deferral of electrical infrastructure upgrades
- Local power quality and reliability improvement

Optimising the scheduling of the system in order to maximise revenue is complicated by uncertainty regarding pricing. While the price profile for certain revenue streams may be known well in advance (see DUoS charges in the UK), those of others may not, for example the purchase of hourly electricity beyond the day ahead in the UK. In principle it should be possible to construct a predictive model for the electricity price based on the intersection of predicted demand with the marginal cost curve of generation (which itself changes with the availability of the various types of generation, and their marginal running costs). Such models require detailed knowledge of the generation base, and some residual variability in price remains [1,2].

An alternative approach is to use heuristics to predict the price behaviour, for example, the use of back-casting whereby the observed profile in a number of previous days is used as the predictor for the coming days [3]. One way to improve this approach would be to reduce the variability in the training set by identifying subsets, assuming these correlated to a predictable calendar period. An example would be the division between the working week and the weekend — it the price profile were consistently different at the weekend (due to reduced demand) then the back-casting would be more effective if training set was drawn from the pool of recent days with the same classification.

One approach to reducing the variability in a dataset is to use clustering techniques. Clustering algorithms are machine-learning methods that can identify statistical subclasses within datasets [4]. However, clustering techniques may also be applied to higher dimensionality data-sets. Electricity price or demand profiles are perfect examples of these, for example, a day-ahead price profile at hourly resolution is a 24D vector.

The number of articles where clustering is applied to energy system optimisation is small. Most commonly, clustering has been applied to simplify the exogenous variable dataset. For example, Teichgraeber and Brandt [5] used a variety of clustering techniques to find clusters in 365 daily electricity price profiles. An optimal energy storage schedule was calculated for each cluster, and a weighted average (based on cluster size) was used to estimate the total revenue. Therefore the optimisation was only run once for each cluster, as opposed to 365 times. The authors judge the performance of the clustering technique by comparison of the optimal revenue based on the clusters to that based on the individual days.

Zhang et al. [6] used the k-means approach to identify clusters in both electricity price profile and load profile for an industrial site in China. For each combination of load and price profile, an optimal schedule was determined and the revenue calculated by the weighted average. The approach relies on an assumption of independence between load profile and price profile. In reality, the particular load profiles are likely to correlate with particular price profiles (e.g. seasonal variation) and the approach would benefit from analysing these relationships.

Green et al. [7] used k-means clustering to identify clusters in UK demand profiles again with the objective of simplifying the exogenous variable dataset. However, the authors also correlate the clusters to particular calendar periods; when the number of clusters, k, is set to six there is strong correlation of cluster membership with the six calendar periods resulting from splitting by weekend/weekday and summer/winter/shoulder season.

Álvarez et al. [8] also found that k-means and fuzzy c-means approaches were able to identify 6 clusters in Spanish electricity price data with good partitioning of cluster membership by weekend/holiday and weekday.

In the following work we use hierarchical clustering for the first time to identify profile clusters in UK day-ahead electricity price data and correlate these to calendar periods. We favour this approach, as it is deterministic – i.e. is not subject to a selected starting point – and the output provides a wealth of information on cluster hierarchy. We then discuss the benefit of knowledge of these clusters when scheduling an energy storage system to perform other services alongside peak shaving.

2. Method

Day-ahead price data for the N2EX trading region (UK) at hourly resolution were downloaded from the NordPool portal [9]. The dataset consists of price profiles from 1st January 2017 to 2nd May 2019.

The hierarchical clustering analysis was performed in Python using the scipy.cluster library at https://docs.scipy.org/doc/scipy/reference/cluster.html, which also provides functions for k-means clustering [10].

The hierarchical clustering method makes no presumption about the number of clusters in the data set. It starts from individual points and agglomerates then starting from the closest pairs. The agglomeration process continues until all of the points have been merged in one cluster. The primary output is a dendrogram showing the distances travelled in order to make each merge, so it is possible to track the process (see Fig. 1). The process is deterministic, as it makes no assumption regarding the location of cluster centres, or the number of clusters. Hence the same dendrogram will always be returned for a given dataset.

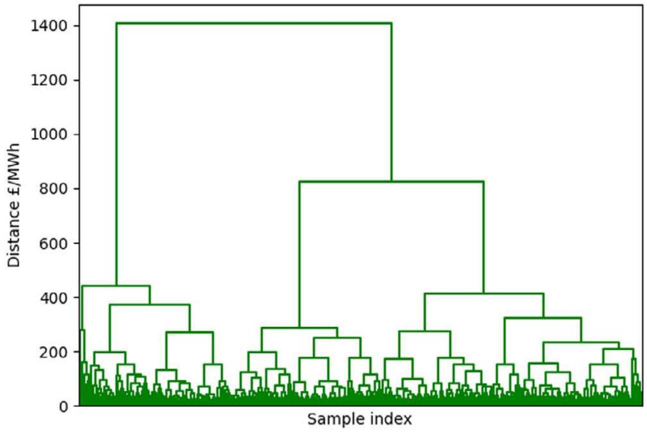

Fig. 1. Dendrogram showing clustering steps.

When two points are clustered, there are a number of methods for describing the combined centre of the cluster (e.g. centroid, medioid) and the distance between clusters. In the following analysis, we apply the Ward method for determination of inter-cluster distance [11]. In this method, the "closest" two clusters are the ones that result in the minimum increase to intra-cluster variance upon merging.

The clustering process may be followed from the bottom of the plot, where a single point represents each day-ahead price profile in the studied period. The dendrogram is a record of every cluster merge that is performed (where two branches join at a horizontal line), and the distance that must be travelled to unite the two clusters.

Determining which jumps on the dendrogram represent merging of genuine clusters is a qualitative process. An elbow plot may be used to inform this choice, as it shows the progressively larger distances that must be travelled to joining clusters. Where there are jumps on the plot, there is likely to be a valid clustering. If unsupervised identification is required, the acceleration on the elbow plot may be used as an indicator. This compares the jump size to those of the previous jumps in order to discern a step change.

3. Results and discussion

The dendrogram output for the Ward clustering on the day-ahead electricity price dataset is shown in Fig. 1.

In addition to the dendrogram in Fig. 1, the cluster merge record for the dataset (last 70 merges) is presented as an elbow plot in Fig. 2.

The elbow plot and the corresponding acceleration plot indicate that there are three primary clusters in the data, as the jumps in distance required to make the last two merges are much greater than any preceding jumps. The arithmetic means of the profiles in each cluster (centroids) are shown in Fig. 3.

These clusters exhibit similar profile shape, with an offset in the absolute price. The occurrence of these clusters is shown in calendar format in Fig. A.1. The low price cluster corresponds to summer 2017 and the beginning of the summer 2019 dataset at the time of writing. The high price cluster corresponds to winter 2018/2019, and the intermediate cluster summer 2018 and winter 2017/2018. There appear to be two factors at play. Firstly, according to Ofgem [12], the natural gas price rose through the summer of 2017, peaking in autumn 2018 before falling back to where it started by the end of Q1 2019. As the natural gas price is the primary determinant of the short run marginal cost of the combined gas cycle power plants that provide the bulk of the UK dispatchable power generation, this trend is a likely explanation for the inter-year variation in cluster occurrence. Secondly, there is a

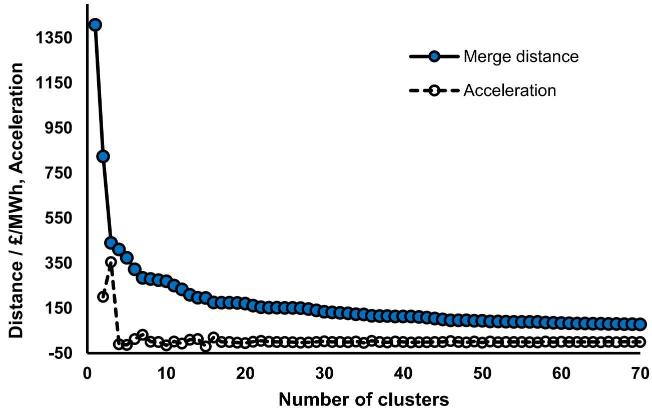

Fig. 2. Elbow plot showing the distances required to make the last 70 cluster merges for the dataset, and the corresponding acceleration in merge distance.

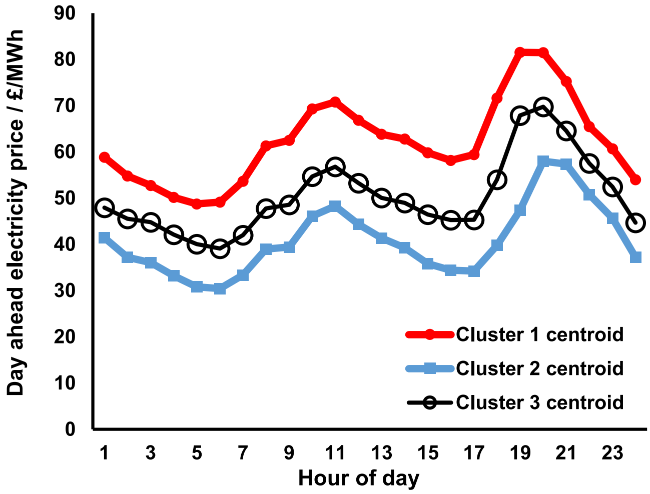

Fig. 3. Centroids of the three clusters identified in the dataset using the acceleration method. Note that on the N2EX market, the first hour of the day ahead is 2300 to 0000.

seasonal variation in national electrical demand, with the summer demand being lower and hence incurring lower marginal costs of generation. This is the most likely explanation for the intra-year variation. It also appears that during the summer of 2018, there is a split between weekday and weekend with the lower price cluster 3 more likely to occur on the latter.

Although the 3 to 2 cluster merge stands out in the elbow plot, there may be further meaningful clusters in the dataset. Hees [13] gives an example where this approach (and other unsupervised cutoff identifiers) may fail to identify valid clusterings.

In Fig. A.2 we display the calendar occurrence of the seven clusters that correspond to the local maximum in the acceleration plot at $n = 7$ (Fig. 2). Clusters 4 and 6 and 7 are sub-clusters of 1, whereas clusters 5, 8 and 9 are sub clusters of 3 (with the numbers denoting the order of the splitting). The low price cluster 2 remains un-split at this point. At this level in the dendrogram, the weekday/weekend split in Q2 and the first half of Q3 2018 becomes more obvious. The centroids of the subclusters represented in this period are shown in Fig. 4.

Membership of cluster 9 is almost exclusively associated with weekends and public holidays in this sub-period. The profile of this cluster is interesting: it does not display a lower price in general than the weekday clusters 5 and 8 as might be expected given the lower demand for electricity at the weekend. In fact, it exhibits the highest overnight prices, and the peak shifts 3 h later in the evening.

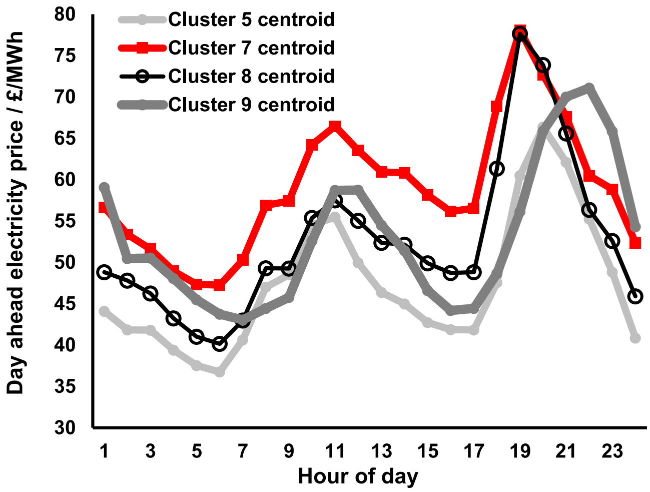

Fig. 4. Centroids of the four clusters identified in Q2 and the first half of Q3 2018, where a marked weekday/weekend split in cluster occurrence is observed (see Fig. A.2). Note that on the N2EX market, the first hour of the day ahead is 2300 to 0000.

One area where the information revealed by cluster analysis may prove useful is in decision making regarding multiple revenue streams, especially when decisions are required further in advance. For example, the firm frequency response (FFR) service must be bid for monthly in 4h blocks [14]. Based on the cluster occurrence identified in summer 2018, during the week it would make sense to bid for FFR in the 1900 to 2300 block during the week, as this is past the peak hours where it may be most lucrative to discharge the battery at maximum power to reduce import costs. However, at the weekend, where the day-ahead price peaks in the hour 2000 to 2100, it may make more sense to bid for FFR in the 1500 to 1900 block.

Clustering may also help to inform back-casting methods. For example, although the existence of a weekday/weekend split appears to be dependent on season and gas price over the longer term, cluster analysis on recent weeks would allow the ESS owner to determine whether these conditions are prevalent.

4. Conclusions

We have shown that hierarchical clustering is a useful method for identifying clusters of similar profiles in UK day-ahead electricity price data, and that these clusters correlate to both calendar periods and longer-term energy price trends. Knowledge of such patterns is likely to be useful for scheduling ESS operation, especially when considering co-scheduling for the provision of services and the performance of multi-hour energy management. This is demonstrated by the identification of a weekend specific profile cluster in summer 2018, where the peak price occurs several hours later than in the rest of the dataset. Knowledge of such a cluster could lead to improved scheduling of FFR to avoid conflict with peak shaving.

Hierarchical clustering may also be useful for identifying calendar trends in other markets, such as the markets for ancillary services. More work would be required if an unsupervised approach to cluster allocation was desired, as the elbow method tested here only identified longer term trends in the profile shape. It would also be interesting to compare the results of the hierarchical clustering analysis with those obtained by alternative methods as assessed by Teichgraeber and Brandt [5]. Lastly, further work is required to understand better the underlying drivers of the profiles shapes for the identified clusters. One approach would be to correlate these with demand profile clusters.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors wish to acknowledge the financial support provided by EPSRC grant EP/L016818/1 to the Energy Storage and its Applications CDT. They would also like to thank Drax Power Ltd for financial support.

Appendix. Cluster occurrences by calendar view

A.1. $n = 3$ Cluster membership

See Fig. A.1.

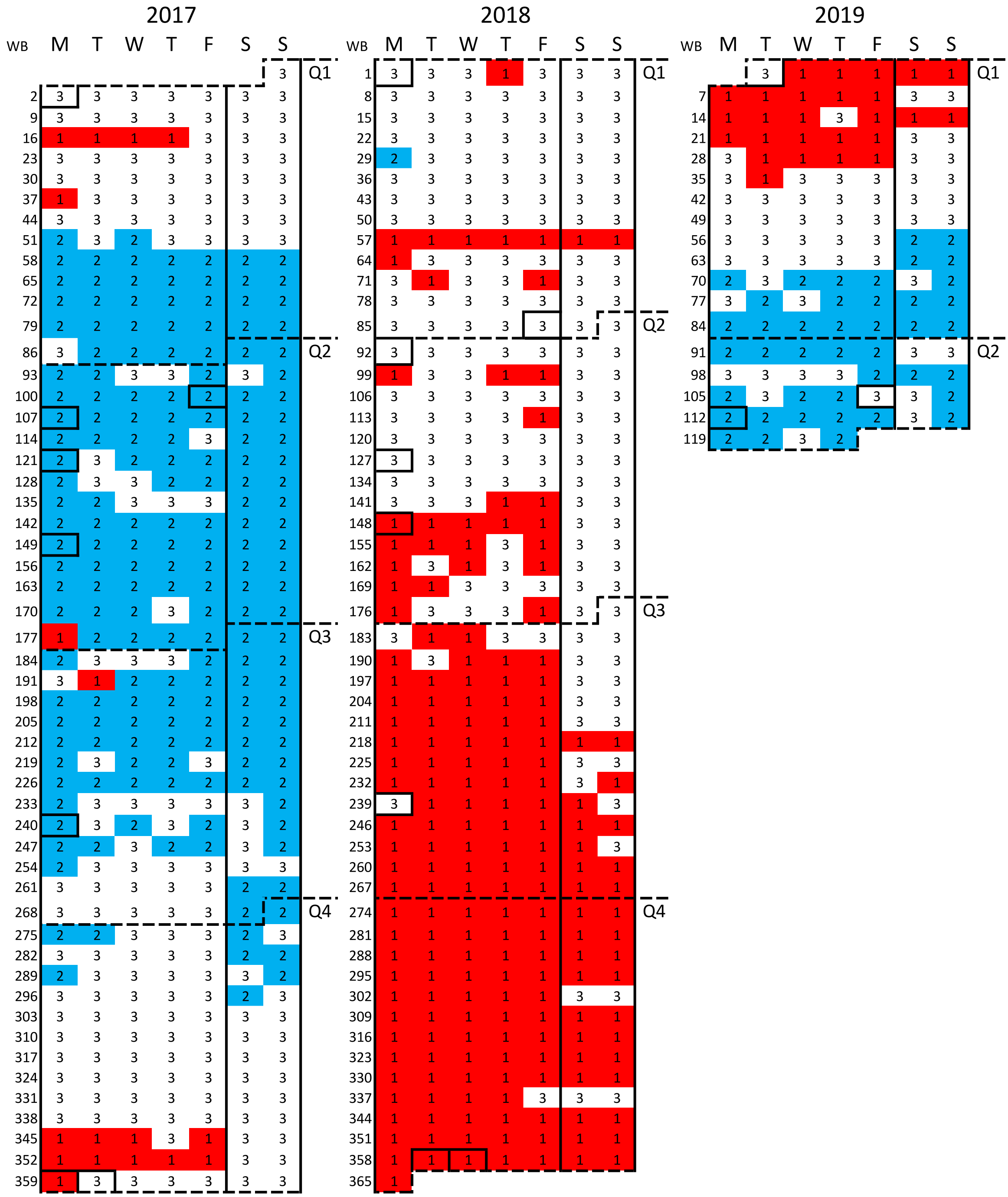

Fig. A.1. Calendar view of $n = 3$ cluster occurrence in day-ahead electricity price-profiles from 1st January 2017 to 2nd May 2019. A heavy border indicates public holidays and weekday/weekend split.

A.2. $n = 7$ Cluster membershp

See Fig. A.2.

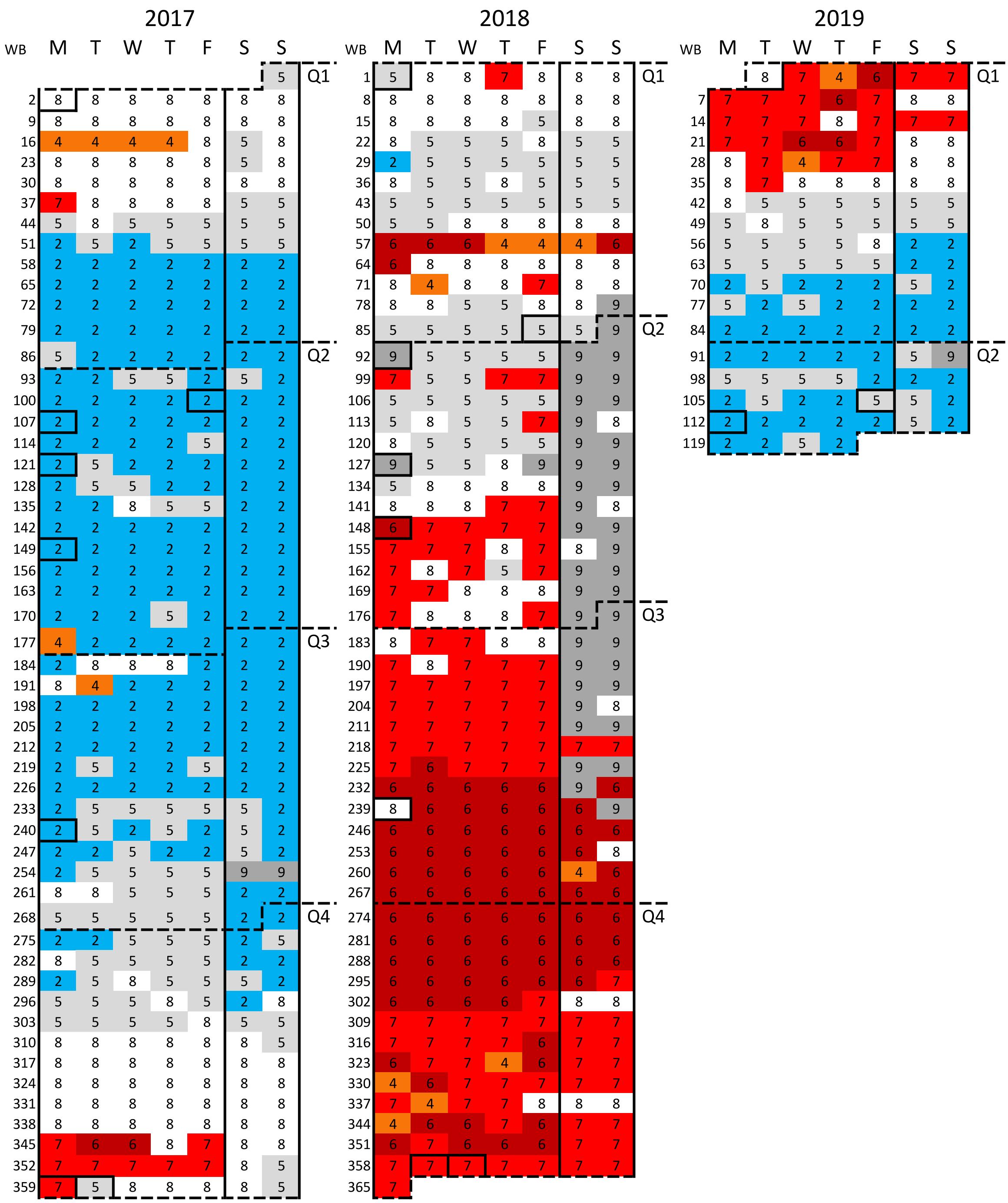

Fig. A.2. Calendar view of $n = 7$ cluster occurrence in day-ahead electricity price-profiles from 1st January 2017 to 2nd May 2019. A heavy border indicates public holidays and weekday/weekend split.

References

[1] Pape Christian, Hagemann Simon, Weber Christoph. Are fundamentals enough? Explaining price variations in the day-ahead and intraday power market. Energy Econ 2016;54(2016):376–87.

[2] Staffell Iain, Green Richard. Is there still merit in the Merit order stack? The impact of dynamic constraints on optimal plant mix. IEEE Trans Power Syst 2014;31(2016):43–53.

[3] Siohansi Ramteen, Denholm Paul, Jenkin Thomas, Weiss Jurgen. Estimating the value of energy storage in PJM: Arbitrage and some welfare effects. Energy Econ 2009;31(2009):269–77.

[4] scikit-learn. Clustering [Internet]. 2019, [Cited 2019 May 8] Available from: https://scikit-learn.org/stable/modules/clustering.html.

[5] Teichgraeber Holger, Brandt Adam R. Clustering methods to find representative periods for the optimization of energy systems: An initial framework and comparison. Appl Energy 2019;239(2019):1283–93.

[6] Zhang Delong, Li Jianlin, Liu Xueqin, Guo Jianbo, Xu Shaohua. A stochastic optimization method for energy storage sizing based on an expected value model. Energies 2019;12(2019):702–15.

[7] Green Richard, Staffell Iain, Vasilakos Nicholas. Divide and conquer? k-means clustering of demand data allows rapid and accurate simulations of the British Electricity System. IEEE Trans Eng Manage 2014;61(2014):251–60.

[8] Martinez Álvarez F, Troncoso A, Riquelme JC, Riquelme JM. Discovering patterns in electricity price using clustering techniques. In: Proceedings of the international conference on renewable energy and power quality. 2007, pp. 245-252.

[9] NordPool. Historical Market Data [Internet]. Available from: https://www.nordpoolgroup.com/historical-market-data/.

[10] SciPyorg. Clustering package (scipy.cluster) [Internet]. 2019, [Updated 2019 Feb 10; cited 2019 May 1] Available from: https://docs.scipy.org/doc/scipy/reference/cluster.html.

[11] Scipy.org. scipy.cluster.hierarchy.linkage [Internet]. 2019, [Updated 2019 Feb 10; cited 2019 May 9] Available from: https://docs.scipy.org/doc/scipy/reference/generated/scipy.cluster.hierarchy.linkage.html#scipy.cluster.hierarchy.linkage.

[12] Ofgem. Gas prices: day-ahead contracts –monthly average (GB). [Internet]. 2019, [Cited 2019 July 4] https://www.ofgem.gov.uk/data-portal/gas-prices-day-ahead-contracts-monthly-average-gb.

[13] Blog Jorn's. SciPy Hierarchical clustering and Dendrogram Tutorial [Internet]. 2019, [Updated 2015 Aug 26; cited 2019 May 2] Available from: https://joernhees.de/blog/2015/08/26/scipy-hierarchical-clustering-and-dendrogram-tutorial/.

[14] National Grid ESO. Firm frequency response (FFR) [Internet]. 2019, [Cited 2019 May 8] Available from: https://www.nationalgrideso.com/balancing-services/frequency-response-services/firm-frequency-response-ffr?overview.