Data-driven Modeling of Building Consumption Profile for Optimal Flexibility: Application to Energy Intensive Industry

Camille Pajot¹, Quang Hung Nguyen¹, Benoit Delinchant¹, Yves Marechal¹, Frederic Wurtz¹
Stephane Robin², Benjamin Vincent³, Francois Debray³
¹Univ. Grenoble Alpes, CNRS, Grenoble INP, G2Elab, 38000 Grenoble
²Universite de Lyon, CNRS, GATE UMR 5824, F-69130 Ecully, France
³CNRS, LNCMI, F-38000 Grenoble, France

Abstract

Despite the vast research on the flexibility of buildings consumption, current literature is more about predicting the impacts of energy flexibility than focusing on its modeling. In this paper, a methodology is provided to go from data-driven modeling of the load consumption to an optimization problem with a Mixed-Integer Linear Programming (MILP) formulation. Illustrated on an Energy Intensive Industries (EII) with an economic point of view, the methodology is suitable for any consumption site, allowing optimal energy planning studies at the district scale.

Thus, it facilitates the definition of flexibility strategies to exploit the complementary use of the districts.

Introduction

As global warming and its consequences appear to be the biggest challenge of this century, energy transition policies are becoming crucial. One of the main strategies to reduce our CO₂ emissions is to increase the share of renewable energies in the global energy mix. However, for the electrical power system, higher penetration of variable renewable energies could lead to greater risks of instability (NREL (2018)).

Flexibility on the demand-side addresses this issue by considering that the consumption could be better synchronized with power production. At the market level, new actors appear in order to trade the flexibility on the demand-side: the aggregators. As indicated by their name, these actors gather consumption sites in order to sell a significant amount of consumption increase or diminution. Aggregators mostly target industries for their power levels, with special care for Energy-Intensive Industries (EII), as they represent half of the entire energy consumption of the industrial sector (U.S. Energy Information Administration (2016)).

On the other hand, electricity prices are a key factor for the competitiveness of many EII, so that they may be interested in some schedule changes in order to reduce their energy bills. In Germany, the electricity consumption of the main EII (chemicals, paper, steel, aluminum, copper and textiles) accounts for 27% of the total electricity consumption (Grave et al. (2015)).

For this reason, Paulus and Borggreve (2011) investigated the technical and economic potential of EII to provide ancillary services in Germany. However, using energy flexibility as an answer to a massive integration of renewable energies is largely studied out of Germany too (Lund et al. (2015)).

For instance, Pallonetto et al. (2016) studied the flexibility from an economic point of view, in the context of a European adoption of time-of-use tariffs. De Coninck and Helsen (2016) developed a methodology to quantify the amount of energy that can be shifted with the associated costs. The main methodologies used to quantify building energy flexibility have been reviewed by Reynders et al. (2018), in particular in cases of thermal storage.

Two main approaches were found to quantify energy flexibility. The first one demonstrates the potential gains obtained by using flexibility, while the second approach aims to predict the available energy flexibility itself. This study could be classified within the first approach of quantifying profits earned thanks to energy flexibility, by showing the economic gains obtained by this flexibility.

However, the innovation of this paper is based on the modeling of this flexibility, so that it addresses the prediction of the available energy flexibility too. Whereas this modeling of flexibility on the demand-side is largely studied for thermal needs, we aim to focus on electrical consumption. Indeed, studies about electrical load modulation mostly rely on shifting existing profiles of typical loads while we aim to model the entire load flexibility of a building. The study will particularly target EII in order to provide them a better knowledge of their own operation flexibility and to integrate them into smart cities strategies.

Thus, this paper is proposing a methodology to model and to optimize the energy flexibility of monitored buildings.

In this work, the optimization considers a single-objective: minimizing the electricity consumption costs on the French day-ahead market, during the six first months of 2018. We are particularly demonstrating the potential benefits by considering the process flexibility of an EII into dynamic pricing. To do so, we will assume that the economic value of consump-
tion flexibility for electrical suppliers can be assessed through the electrical day-ahead market. Then, we can compare a non-optimal operation (real data from measures) to the optimal results obtained by considering the flexibility of the EII.

In order to carry out the optimal energy planning, we rely on a Mixed-Integer Linear Programming (MILP) approach which is common in the research filed. For example, Viana and Pedroso (2013) studied the MILP formulation of unit commitment problems in power production planning, which can be transposed to consumption planning. On the demand-side, Silvente and Papageorgiou (2017) investigated the minimization of the operation cost of a microgrid, by applying a penalization for deviations from the previous energy consumption. Moreover, Tenfen and Finardi (2015) implemented new optimization models for curtailable loads. Among others, the MILP formulation has the advantage to allow a quick optimization with a large number of variables.

In this work, the optimization problem is formulated as MILP with the help of the open source python library OMEGAlpes\(^1\) (Optimization ModEls Generation As Linear Programming for Energy Systems). To do so, we relied on a model of flexible load provided by OMEGAlpes and focused on the quick definition of the flexibility parameters required to instantiate the model. Thus, the MILP formulation of flexible consumption is not the subject of this paper and is described by Anonymous (2019).

**Study case**

The LNCMI (French National High Magnetic Field Laboratory) is a host laboratory for experiments in high magnetic fields. At its Grenoble site, electricity consumption required in order to provide magnetic static fields up to 37T reaches about 15 GWh per year (up to 24 MW electrical power). This high-consumption level put the LNCMI as an EII with an operation strongly affected by energy bill. Indeed, with the main purpose of providing scientists access to the magnets for experiments, the LNCMI still have to respect its annual allocated budget for electricity consumption. As the current operation is based on a schedule of experiments handled by hand, the LNCMI can very easily re-schedule it. Despite the very high importance of flexibility for the electrical system, this modularity has still not been exploited by energy suppliers. In order to negotiate with them, a first step is to characterize the LNCMI flexibility, i.e. the degrees of freedom of its operation. To do so, a better knowledge of the particularities of the process is required.

**Paper structure**

After describing the used approach for flexibility potential estimation, this paper will focus on the model-

---

\(^{1}\)https://pypi.org/project/omegalpes/

---

Figure 1: Electrical consumption profile of the LNCMI during 5 days of the month of June 2018.

Figure 2: Illustration of a re-schedule of each hourly consumption slots from a real consumption profile.

**Methods**

**Approach for the estimation of the optimal flexibility potential**

Frank et al. (2017) draw a distinction between energy flexibility (power variation) and time flexibility (starting time variation). In order to re-schedule the experiments while keeping them unchanged, this study case only focuses on the load shifting (time flexibility). The flexibility potential will be quantified by the financial savings obtained by the new planning of experiments. As it reflects the hourly price variation of the electricity costs on the flexibility market, the day-ahead price of the spot market was chosen to estimate the economic value generated by the time flexibility.

Then, a reference scenario is needed to estimate the savings. In this case, we will consider as a reference, the cost of the real LNCMI consumption on the spot market. As the LNCMI does not buy its electricity on this market, this reference is a fictitious value, based on real-data for both spot prices and consumption values. Our aim is to provide a methodology to quickly compare this reference with an optimized consumption profile.

A first approach (further referred to as "Opt. (a)") is to re-schedule each hourly consumption slots according to the market prices. An example of re-schedule can be found Figure 2.
This method presents the advantage of being easy and fast but could be characterized as naive from an industrial point of view. Indeed, this new consumption profile does not take into account the complexity of the building operation. However, it allows a first estimation of the maximal gains that can be obtained.

For further studies, it becomes crucial to quantify the flexibility of the building, i.e. what are the degrees of freedom for the load shifting. In our case, the LNCMI has great flexibility on the experiments schedule. Nevertheless, the experiments could be characterized as motifs of the consumption profile, that cannot be shortened through smaller motifs. Therefore, power values belonging to the same experiment should not be separated, contrary to Opt. (a).

This second approach (further referred to as "Opt. (b)") will be detailed in the next section.

Finally, the optimization problem is built with OMEGAAlpes. The MILP problem studied here is defined by the following characteristics:

- **Time period:**
  - From January to June 2018
  - Hourly time slots

- **Objective:**
  - Minimizing the consumption cost of LNCMI on the day ahead electricity spot market.

- **Hypothesis:**
  - Opt. (a): The hourly consumption are disconnected from one another.
  - Opt. (b): Consumption profiles are approximated to basic shapes.

- **Constraints:**
  - Opt. (a): Each hourly power value from the real consumption profile has to be scheduled during the time period.
  - Opt. (b): The number of experiments per month is constant, i.e. the scientific activity is equally distributed during the year.

Thus, the estimation of the optimal flexibility potential stands on three main stages (Figure 3): the calculation of the reference to be compared, the modeling of the flexibility (by two approaches) and the MILP formulation in order to solve the optimization problem.

In this paper, we particularly focus on accurate modeling of the time flexibility thanks to a method based on a classification approach (Opt. (b)).

**Disaggregation of the consumption curve**

As explained previously, an efficient forecast of the LNCMI consumption profile should keep the experiments consumption profiles intact, by only shifting them. This assumption allows us to guarantee the respect of scientific needs reflected by the load curves.

The first step of our modeling will be to separate the experiments from the entire consumption profile. To be selected, an experiment on the load curve must fulfill two criteria:

- it should be longer than 30 minutes
- its energy consumption ($E_c$) should be greater than 100 kWh

These constraints prevent the algorithm from considering the industrial process testing or maintaining periods as real reschedulable consumptions. Furthermore, the model must take into account that some consumption profile could be turned off for a short duration, before starting again, and should be selected as a unique consumption motif. To do so, the selection principle is based on the observation of the mean power consumption during a sliding observation window. Now that consumption curve is split by experiment, the next step consists in forecasting a standard load consumption.

**Modeling of consumption patterns**

From one year to another, the experiments will not be exactly the same, but will still behave similarly according to the experiment purposes. After observation of the load profiles, three main patterns seemed to appear:

- rectangle (Figure 4)
- triangle (Figure 5)
- saw-tooth (Figure 6)

After a validation of these motifs by the expert on experiments planning, models have been converted into parametric equations (1 to 6).

The rectangle profile can be described by its duration: $D$ and its power value: $P$, according to (1).

$$\hat{p}_t(D, P) = \begin{cases} P & \text{if } t \in [0; D] \\ 0 & \text{otherwise} \end{cases} (1)$$

For the triangle profile, one must add the time of the peak, so that it can be described by (2).
Figure 4: Rectangle motif and parameters.

Figure 5: Triangle motif and parameters.

\[
\hat{p}_t(D, P, \alpha) = \begin{cases} 
  \frac{P}{D} t & \text{if } t \in [0; \alpha D] \\
  \frac{P}{D} - t (1 - \alpha) D & \text{if } t \in [\alpha D; D] \\
  0 & \text{otherwise}
\end{cases} \tag{2}
\]

Finally, the saw-tooth shape requires two additional parameters: the duration between peaks: \(d\) and the number of peaks: \(n\), as presented by (3).

Figure 6: Saw-tooth motif and parameters.

\[
\hat{p}_t(D, P, \alpha, d, n) = \begin{cases} 
  \frac{P}{D} t & \text{if } t \in [\tau^r_n; \tau^p_n] \\
  \frac{P}{D} - t (1 - \alpha) D & \text{if } t \in [\tau^p_n; \tau^e_n] \\
  0 & \text{otherwise}
\end{cases} \tag{3}
\]

Where:

\[
\tau^r_n = (n - 1)(d + D) \tag{4}
\]

\[
\tau^p_n = \alpha D + (n - 1)(d + D) \tag{5}
\]

\[
\tau^e_n = (n - 1)(d + D) + D \tag{6}
\]

Classification of consumption patterns

Once the patterns defined, each experiment needs to be classified into one of the three categories. To do so, a first step consists in the extraction of the main features of each experiment (duration, mean power and maximal power). As the duration of an experiment should not be changed by the load shifting, it was assumed that the duration of an experiment should remain intact after the classification process.

Then, each experiment would be classified into one of the three patterns presented before. The selection of the corresponding pattern is based on a distance calculation defined as (7).

\[
dist(exp, pat) = \sqrt{\sum_{i=1}^{D} (p_t(exp) - \hat{p}_t(pat))^2} / D \tag{7}
\]

Where \(p_t(exp)\) is the power of the experiment at \(t\), while \(\hat{p}_t(pat)\) is its estimation defined by the evaluated pattern.

Indeed, the algorithm proceeds by iterating steps for each experiment:

- Adjusting the parameters of each pattern in order to find the nearest profile to this pattern
- Definition of the corresponding pattern by selecting the one with the smaller distance

Finally, a confidence indicator will be attributed for the clustering. Here, a classification will be called reliable when \(\text{dist} < 0.4\). Then, the proportion of certain classification (\(R_{\text{rate}}\)) could be calculated for each pattern (8), in order to identify if the patterns are really suited to model the experiments.

\[
R_{\text{rate}, pat} = \sum_{\text{if } \text{dist}(exp, pat) < 0.4, \text{exp}} \text{N}_{\exp, pat} / \text{N}_{\exp, pat} \tag{8}
\]

Where \(\text{N}_{\exp, pat}\) is the number of experiments classified in the pattern.

Results

Classification results

To forecast a standard consumption profile of the LNCMI, the first step is to select the consumption profiles of each experiment. The results presented Figure 7 were obtained with a Gaussian sliding window of 2 hours, with a standard deviation of 1.5. During one year, 595 experiments have been identified.

The second step of the consumption forecast is to attribute each experiment to a cluster, defined by a specific shape of the power curve (rectangle, triangle or saw-tooth). Each experiment belongs then to its closer motif (Figure 8), according to the distance defined by (7).

Then, a confidence indicator was introduced in order to express the level of certitude obtained by the classification. The number of experiments classified in each pattern (\(\text{N}_{\exp}\)), as well as the mean distance
obtained (dist\textsubscript{mean}) and this reliability rate (R\textsubscript{rate}) are presented in Table 1.

The low certainty rates obtained for the experiments defined as a saw-tooth consumption profile shows the difficulty to classify some experiments with very atypical profiles.

The modeling of the experiments consumption profiles offers a better understanding of how the EII is operating. Then, the clustering algorithm enabled the classification of these experiments into three classes: rectangle, triangle, and saw-tooth, according to the consumption profiles’ shape. The results from the rectangle pattern strongly suggest that this shape is well fitting for about 19% of the experiments while considering all experiments modeled by rectangle motifs lead to 43% of reliability.

Even if a simple classification of such a complex consumption leads to few highly-reliable profiles, the main parameter for the optimization problem stays the duration of the consumption profiles to shift. In our case, regardless of the assigned motif, the duration of the consumption profile is set to the experiment duration. For this reason, we found the use of these results suitable for the optimization problem.

**Optimization results**

<table>
<thead>
<tr>
<th></th>
<th>Rectangle</th>
<th>Triangle</th>
<th>Saw-tooth</th>
</tr>
</thead>
<tbody>
<tr>
<td>N\textsubscript{exp}</td>
<td>111</td>
<td>220</td>
<td>264</td>
</tr>
<tr>
<td>dist\textsubscript{mean}</td>
<td>0,51</td>
<td>0,49</td>
<td>0,71</td>
</tr>
<tr>
<td>R\textsubscript{rate}</td>
<td>43%</td>
<td>37%</td>
<td>13%</td>
</tr>
</tbody>
</table>

Table 1: Classification results.

With nearly no constraint, the results from Opt. (a) follows as much as possible the variations of electrical prices, while the profile obtained from Opt. (b) has to respect the experiments shapes (Figure 9). The first approach (Opt. (a)) seems to provide a good support for the planning of experiments. Indeed, Figure 9 shows that the consumption profile obtained by Opt. (b) fits quite well the results from Opt. (a). In order to validate the confidence that can be given to the first approach for the estimation of the economical gain, the economical results from both Opt. (a) and Opt. (b) can be found in Table 2 under the label "Cost".

From an electricity cost of reference of 37,8\(€\)/MWh to 18,0\(€\)/MWh, the Opt. (a) provides a reduction of 52%. As explained previously, this value allows us to estimate the maximal reduction that can be realized. More constrained, the second optimization (Opt. (b)) reduces the cost of 43% with a mean cost of 21,5\(€\)/MWh. The results confirm the interest in using the first approach (Opt. (a)) for a quick estimation of the potential financial gains.

However, further studies would rather stand on the methodology provided in Opt. (b), which provides a better understanding of the constraints of the process. Although this analysis does not justify a true advantage of this method, this result can be explained by the wide freedom on the schedule, with 595 diversified experiments. Some series of experiments could yet need to be considered together as they form a single scientific study. This conclusion was only possible thanks to a better understanding of the operation given by the modeling of the experiments. Thus, it is important to guarantee that the forecast provides realistic consumption profiles. This point was already discussed in the classification results, but can be confirmed by having the same amount of energy consumption (E\textsubscript{c}) for Opt. (b) as for the real consumption profile (see Table 2). Indeed, Opt. (a) shifts real consumption slots, while Opt. (b) schedules the pattern models of the experiment consumption

<table>
<thead>
<tr>
<th></th>
<th>Ref.</th>
<th>Opt. (a)</th>
<th>Opt. (b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost [€/MWh]</td>
<td>37,8</td>
<td>18,0</td>
<td>21,5</td>
</tr>
<tr>
<td>E\textsubscript{c} [GWh]</td>
<td>8,76</td>
<td>8,76</td>
<td>8,85</td>
</tr>
<tr>
<td>Work by night [%]</td>
<td>35%</td>
<td>54%</td>
<td>56%</td>
</tr>
</tbody>
</table>

Table 2: Optimization results.

Finally, Table 2 shows the results obtained by the two optimizations, as well as the reference values, for the rate of hours of work by night during the 6 months. As could be expected, more experiments have been scheduled by night with the optimization based on
electricity prices with an evolution of the time of work by night that could increase from 35% to 56% with an optimal schedule.

This aspect was not the subject of this paper, but since it can guide the EII to get a better knowledge of their way of consuming, this will be discussed in the next section.

Discussion

Industrial constraints and operating hours

For some EII, it could happen that the industrial constraints are not all reflected by the consumption profiles. Once identified, these constraints can be added to the optimization problem while keeping the same methodology. For instance, the operating hours could be very problematic in such load shifting problems. As the night hours are often associated with lower prices, the optimal solution may result in more work during the night.

For this reason, two solutions could be investigated. At first, a constraint can be easily added in order to avoid any working hour. If the purpose is only to limit the work by night, another option is to integrate financial penalties for the operating hours happening during the night. In this particular case, lots of experiments are already scheduled by night. However, including the improvement of the working quality is the topic of further studies.

Potential of the approach for multi-carriers energy planning at the district scale

Although this paper focuses on EII, it can be pointed out that this methodology can also be used for load shifting purposes in residential buildings. With lower power values, the residential consumption was not the first center of interest of aggregators. With a large amount of residential building to gather, the residential area becomes yet targeted for diffuse load shedding/load shifting strategies.

Indeed, households can also benefit from variable tariffs, while having shiftable consumption, such as the washing machine, the dishwasher, the dryer, the boilers for hot water or the heat pumps for building heating. The development of smart electric meters will not only lead to a better understanding of the consumption profile but allows the use of methodologies such as this one.

Then, optimization studies can be realized at the district scale by gathering models of several types of buildings (EII, residential and tertiary). This allows to go from local points of view, driven by only one actor, to look at the global interest of multiple stakeholders. An example of an application is to use consumption profiles diversity for peak-shaving purposes. By reducing the cost of energy production and distribution, this benefits the entire district and can be achieved by sending signals to the consumers in order to motivate them to shift non-necessary loads.

Conclusion

This work focuses on the data-driven modeling of a consumption profile for optimal flexibility from an economic perspective. As the profitability of EII partly relies on the electricity bill, flexibility on the demand side appears as a way to reduce costs and thus increase their competitiveness. More significantly, the modeling of its consumption profile and its flexibility can improve their knowledge of their own operation. This aspect can be crucial for a better understanding of the constraints of the EII to reduce their energy bill.

In order to quickly quantify the impact of flexibility on the demand-side, the method proposed here relies on three main steps:

1. Disaggregation of the load curve: to easily identify the changes of use.

2. Modeling of the main consumption patterns:
   - Definition of the main shapes of consumption profiles.
   - Classification of the uses by similar shapes.

3. Quick formulation and resolution of the optimization problem thanks to a MILP-models generation tool (OMEGAlpes).

These steps can mostly be automatized and will be more and more used, thanks to the massive deployment of load monitoring through smart grids and smart cities, but will also be improved along with machine learning techniques.

This methodology was illustrated in the study case in order to quantify how the load shifting could benefit an EII. In the case of LNCMI, showing the potential gains realized by using its flexibility could lead to a better negotiation with the electricity supplier. Indeed, knowing the degrees of freedom allowed by their operation is a first step of a discussion on variable tariffs. Moreover, the methodology shows what could be the impacts of this flexibility on the LNCMI operation, by proving a new schedule for the experiments. This opens the discussion and highlights which constraints can or cannot be considered.

Although applied to an EII, the presented approach can be used for several kinds of energy uses. Thus, studies can be realized at the district scale by gathering different types of buildings (EII, residential and tertiary), through a single modeling technique that can be mostly automatized.

Acknowledgments

This work has been partially supported by the ANR project ANR-15-IDEX-02.
Nomenclature

$\alpha$: Ratio between time of peak and duration

d: Duration between two triangles motifs

D: Duration of the motif

n: Number of triangles in the saw-tooth

$p_t$: Power value at the instant $t$

P: Maximal power value of the motif

$\tau^n_e$: End of the $n^{th}$ triangle motif

$\tau^n_p$: Time of peak of the $n^{th}$ triangle motif

$\tau^n_s$: Start of the $n^{th}$ triangle motif

References


