# Generation of aggregated plug load profiles in office buildings

Shelly Clemente, Solène Beauchêne, Elyes Nefzaoui

# Generation of aggregated plug load profiles in office buildings

Shelly Clemente[1], Solène Beauchêne [1,2] and Elyes Nefzaoui[1,3,*]
[1] *Efficacity, F-77447 Marne la Vallée Cedex 2, France*
[2] *EDF R&D, EDF Lab Les Renardières, F-77818 Moret sur Loing, France*
[3] *ESYCOM Lab, Univ Gustave Eiffel, CNRS, F-77454 Marne-la-Vallée, France*
**Corresponding author: elyes.nefzaoui@esiee.fr*

## Abstract

Plug loads are one of the most important energy consumption items in non-residential buildings and their weight is continuously increasing in new generation buildings with highly insulated envelopes and high efficiency HVAC systems. An accurate and realistic modeling of plug loads is therefore of paramount importance for building energy modeling and energy efficient operation, as well as for other emerging applications such as renewables integration and demand response.

We present in the present paper a method for generating aggregated electric load profiles for plug loads in office buildings based on more than 6400 field-measured load profiles of individual appliances used in office buildings and a field survey covering more than 1000 office building occupants which provides the distribution of the considered appliances in office buildings. We show that the aggregated load curves scale up following different laws depending on the considered appliances. We also provide aggregated load curve stabilization thresholds for all appliances: 43 for desktop computers, 32 for computer screens, 127 for laptops and 204 for multifunction devices if we mention the main office appliances Finally, by combining the individual appliances' aggregated load curves and the survey results, we propose a model for generating building-level aggregated plug load curves for three main building types defined through the survey data analysis. We show that the aggregated load curves do not scale

up differently when the number of occupants increases depending on the building type and the day type. Obtained results provide unique insights on plug load aggregated load profile shapes and behaviors and can be very useful for researchers and engineers interested in non-residential building energy modeling.

## Keywords

Plug load, load profile, diversity, aggregation, office buildings, energy consumption

## Abbreviations

| | | | |
|---|---|---|---|
| *ALP* | Aggregated Load Profile | $n_{occ}$ | Number of occupants |
| *BALP* | Building Level Load Profile | $n_{th}$ | Stabilization threshold |
| *BEM* | Block Element Modifier | *NALP* | Normalized Aggregated Load Profile |
| *DF* | Diversity Factor | *SAALP* | Single Appliance Aggregated Load Profile |
| *HVAC* | Heating, Ventilation and Air Conditioning | $r_i$ | Appliance per occupant ratio per appliance *i* |
| $n_{app,i}$ | Number of appliances for each appliance *i* | | |

## 1. Introduction

Reducing building energy consumption and related greenhouse gas emissions is one of the major challenges for engineering and research on the built environment. Indeed, buildings account for 29% of worldwide final energy consumption and 49% of the total electricity consumption [1], and this share is continuously increasing [2]. To tackle these issues, recent opportunities have risen in the development of smart infrastructures [3] and the significant role of smart meter deployment plans worldwide, in the United States of America [5] and Europe in particular[6] . The growing availability of data collected from advanced metering infrastructure is therefore a strong asset for

research and development towards a better and realistic understanding and modeling of building energy consumption.

To take the full benefit of the amount of collected data, the current trend in building energy modeling is shifting from traditional physics-based modeling [7] to data-driven methods [8]. To capture the diversity of occupants' behaviors, which is a key parameter for an accurate building energy model [9], stochastic methods are required and can be enhanced by the increasing amount of available energy demand data. Such data has been recently used with different data processing techniques for a variety of purposes related to building energy consumption modeling such as pattern recognition [10], [11], abnormal energy behavior identification [12], building energy demand characterization by load profiles clustering [13] , demand side management for industrial [14] and residential [15] sectors, building energy consumption [16] and peak demand [17] forecasting. A comprehensive review can be found in [8]. These techniques are also used for various practical applications including the identification of priority targets for energy efficiency programs [18], the optimization of equipment sizing, energy storage, electric network operation[19], renewables integration [20] and commercial offers [21]. Studies have mainly covered residential households and then mixed industrial and commercial buildings as highlighted in [17]. Other non-residential buildings such as education, research or office buildings have more seldomly been considered [11], [22].

We focus in the present work on office buildings. We specifically consider the problem of modeling and generating aggregated electric load curves for small power equipment in office buildings. The main loads in an office building are HVAC , lighting and small power equipment or plug loads [23]. HVAC and lighting loads are mainly governed by weather conditions, the day duration and equipment technological features, while plug loads are mainly governed by the

occupants' behavior and show no or very low sensitivity to weather conditions as highlighted in recent commercial buildings use cases in the United States[24], [25] and field monitoring campaigns in residential buildings in Australia. [26]. The non-sensitivity of plug loads to weather conditions is also admitted in several building professional guides such as ASHRAE Handbook of Fundamentals[27] or CIBSE Guide F[28] where office appliances and plug loads are accounted for using constant watt per surface area heat gain without seasonal modulation. The strong dependency of plug loads to occupants makes them harder to model since the occupancy is considered as one of the major factors of a building's energy performance gap [29]. This is particularly critical in energy efficient buildings[30]. In addition, for such buildings, with their very low transmission envelopes, the weight of internal heat sources, such as plug loads, becomes more significant in the building energy budget. They should be carefully taken into account for accurate sizing and operation of such buildings and are therefore attracting an increasing attention from buildings operation[31] and research communities [32], especially for most recent and energy efficient buildings[33]. Plug loads are traditionally accounted for in BEM tools using benchmarked average energy consumption or power load density values [28], weighted by occupancy schedules provided as a limited number of typical occupancy profiles [34] as recommended by standard building energy modeling protocols[35]. A few recent works have highlighted the weaknesses of such an approach, especially for office buildings. They showed the relevance of using field measured data for small office power equipment energy consumption estimation compared to the recommendations of commonly used guidance documents [36], proposed different stochastic strategies to estimate the power demand of such equipment based on random sampling of measured load profiles or a bottom-up construction of load profiles[37] and developed simplified

models to generate plug load profiles combining input occupancy profiles and random sampling over the learned load profiles patterns from measured data[38].

In this paper, we follow this path generate building-level aggregated plug load profiles by summing individual appliance load profiles randomly selected from a large dataset of field-measured load profiles. This method has two main advantages: i) the individual appliance load profiles are not simplified models but real load profiles accounting for load variability while simplified models use occupancy-weighted average load values ii) occupancy information is partially included in individual appliance (such as computers, screens, etc.) measured load profiles. When a large number of different individual load profiles is aggregated, the diversity of the occupancy and the occupants' behavior are implicitly accounted for. The main drawback of this method is the cost of field data collection, the lack of open datasets and the bias that can be induced by limited size datasets. In this work, we use a dataset of several thousand individual appliance load profiles collected in office buildings in Paris metropolitan area (France) from 2017 to 2019. We combine those load profiles with appliance per occupant ratios for each appliance type obtained through a field survey conducted in 2018 and 2019 amongst 1038 office building workers in France. Using the measured appliance load curves, we obtained aggregated load curves for each appliance and a diversity stabilization threshold, i.e. a number of appliances beyond which the aggregated load curve shape is not significantly affected by aggregating additional appliance load curves. The determination of this threshold is useful to optimize load curve generators. By combining the appliance load curves and the survey data, we were also able to generate building-level aggregated load curves for different buildings types and sizes. All developed algorithms have been integrated in a proprietary tool for load profile generation for building energy modeling used in the authors' institutions. In the present paper, we present the main methods, algorithms and results of this work.

The present paper is constructed as follows. We present in section 2 the methods used for data acquisition, pre-processing, structuring and processing through the whole process leading to the computation of aggregated load curves. We focus more particularly on the algorithms used to obtain the stabilization threshold and the aggregated load curves. In section 3, we present and discuss the main result with respect to the used dataset statistics, the stabilization threshold for each appliance and the building-level aggregated load profiles where we specifically compare load profiles of buildings of different sizes, below and above the stabilization threshold.

# 2. Methodology

## 2.1. Data acquisition

Collected data, individual appliance load curves for instance, have been measured in 3 different office buildings in Paris Region, France (additional information on the considered buildings are provided in appendix F), on different appliances commonly found and used in offices, either by individual users or as shared resources, such as desktop computers, laptops, screens, printers, multifunction devices, kettles, coffee machines, etc.

Load curves have been measured on each device using Currentcost EnviR wireless plug load meters [39]. Each ten plug meters communicate with a gateway that has been customized using a Raspberry Pi to store collected data on an SD card. This inhouse solution enables to extend the storage capacity of the commercial device in an unlimited way depending only on the used SD memory card capacity.

In addition to centralizing the measured data, the gateway also provides a time reference and ensures that the ten load curves are synchronized.

Plug loads have been measured with a 6 second time step which is the shortest time step enabled by the used meters. Consequently, each daily load curve contains 14400 data points. Different measurement campaigns have been performed on different sites, buildings and appliances. Each

campaign lasted at least three weeks to ensure having both business day and closing day load curves. Many campaigns lasted more than two months.

## 2.2. Data pre-processing: cleaning, meta-data, indexing, data-frames structures, etc.

Collected data was pre-processed before being used in the final program for aggregated load curves generation. First, raw data was cleaned to remove outliers, for example values which are too large compared to the nominal load of each equipment (user defined thresholds for each equipment category are used to determine whether a value is an outlier or not. The threshold value for each appliance type is given in appendix F, section A6.2). In addition, load curves with many contiguous missing points were not considered. When missing data represents less than 10 contiguous measurement points, one minute in our case given the time step of 6 seconds, we use mean imputation to replace them.

## 2.3. Data structure

For each monitored equipment, data is first stored in a date-indexed DataFrame with one daily load curve per line. Each line therefore contains 14400 measurement points.

After a first data pre-processing, not presented here, we noticed that the appliance load curves can be classified into two clusters: business day clusters (mainly constituted of non-nil load curves suggesting that the appliance has been used) and closing day clusters (mainly constituted of nil load curves suggesting no use of the appliance during that day). Nevertheless, it may happen to have a nil load curve for a business day if one user is absent from office or a non-nil load curve for a closing day if, exceptionally, a user is present on the premises during one closing day. We therefore tag the different load curves either as a business day load curve, vacation load curve, or a clo sing day load curve. The tag is only based on official calendars: business days (from Monday to Friday), closing days (Saturday, Sunday, and national holydays), vacation days

correspond to school vacation based on an official calendar provided by the Ministry of Education. Indeed, based on a previous work[40] of our group and national statistics[41], we note that school holidays affect in a non-negligible manner the presence of occupants in workplaces and therefore the aggregated load curves. In the following, the date index is not used anymore. Then, load curves of all appliances of the same type (desktop computers, laptops, etc.) are gathered in the same DataFrame with the line number (load curve number) used as an index. A day type tag is also added to each load curve based on the above mentioned rule.

## 2.4.Data processing

The goal of data processing is to generate aggregated load profiles. Two types of aggregated load profiles are considered: Single Appliance type Aggregated Load Profiles (SAALP), an aggregation of load profiles for n computers for example, and Building level Aggregated Load Profiles (BALP) including different kinds of appliances. For SAALP, the goal is to determine whether the aggregated load profile shape stabilizes for a large number of appliances. For this purpose, we calculate the diversity factor defined and commented in section 2.4.1. For BALP, the goal is to generate building-level load profiles that can be used for different purposes such as building energy modeling, renewable generation device sizing, demand response, etc. For this purpose, generated load profiles need to be realistic. We present in section 2.4.2 the model used for BALP generation which is strongly based on field collected data.

To generate the aggregated load profiles (ALP) presented in this paper, a house made code was developed. We show in Figure 1 the workflow of the used program.

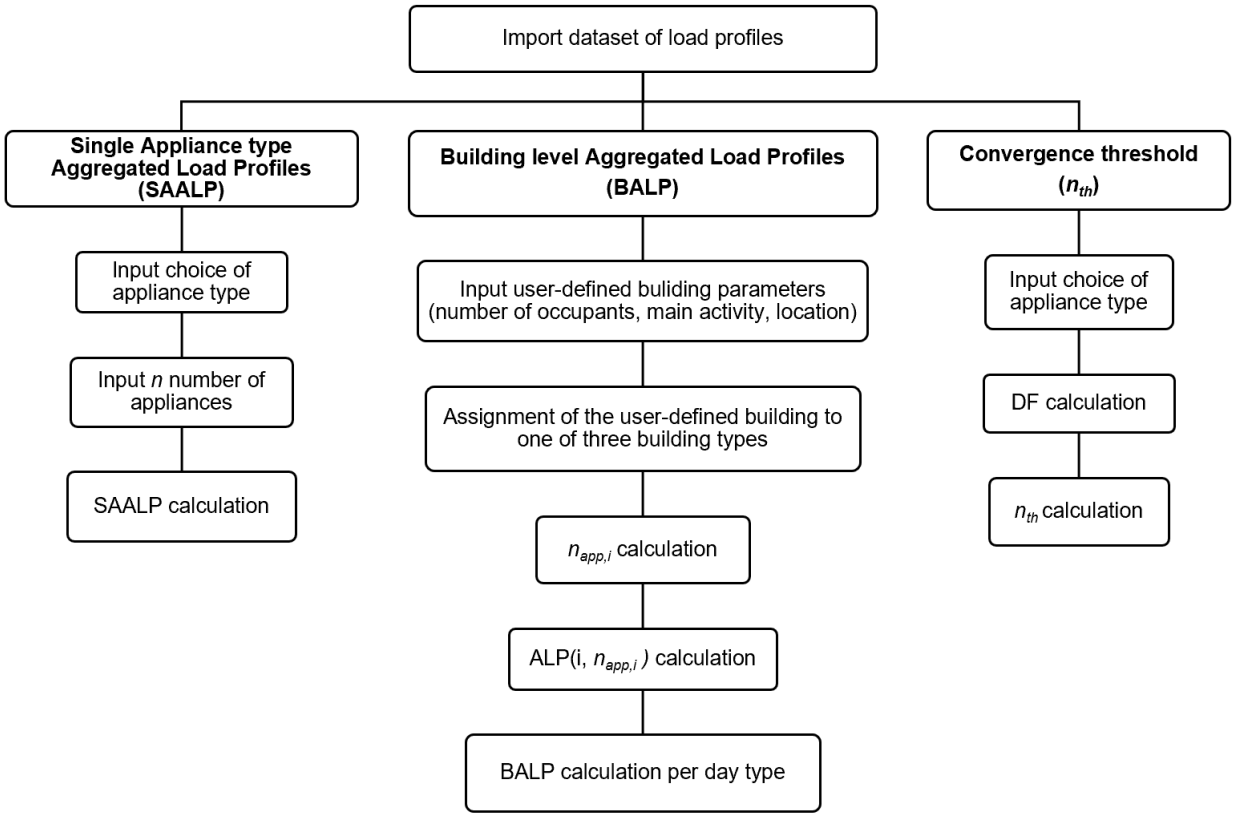



*Figure 1: Workflow of the aggregated load curve generator.*

First the user imports a dataset containing measured and pre-processed appliance load profiles. Then he can choose between three main functions: generating a SAALP, a BALP or determining the stabilization threshold, defined in the following paragraph, for a given appliance. The main steps of the three available functions are described below.

## 2.5. Diversity factor

The diversity factor (DF) is defined by:

$$DF\ (appliance\ type, n) = \frac{\sum_{i=1}^{n} \max(IALP_i)}{\max(\sum_{i=1}^{n} IALP_i)} \quad (1)$$

where $IALP_i$ is an individual appliance load curve and $n$ the number of appliances considered for aggregation. The sum in the denominator is a point wise sum of $IALP_i$.

To compare the aggregated load curve shapes we also defined a normalized aggregated load curve as follows:

$$NALP = \frac{SAALP}{max(SAALP)} \quad (2)$$

where $SAALP$ is the aggregated load profile defined as:

$$SAALP(appliance\ type, n) = \sum_{i=1}^{n} IALP_i \quad (3)$$

one should note here that $NALP$, $SAALP$ and $IALP_i$ are time-series of 14400 points corresponding to daily load profiles with a time step of 6 seconds.

## 2.6.Stabilization threshold

For a given appliance, when a large number of load curves $n$ are summed to obtain a SAALP, the shape of the NALP is expected to stabilize for $n$ large enough. The number $n$ at which the NALP stabilizes is called the stabilization threshold and noted $n_{th}$.

Empirically, we systematically observe that DF tends to and stabilizes around a constant value for a number of appliances $n$ larger than $n_{th}$.

To determine the stabilization threshold, we can also calculate the Euclidean distance between two consecutive NALP, *i.e.* between $NALP(n+1)$ and $NALP(n)$. This difference is expected to tend to zero for $n$ large enough.

In the following, the stabilization threshold is determined using the DF as follows:

The DF is calculated for $N$ appliances where $N$is the number of load curves available in the input dataset for a given appliance. If the DF stabilizes, i.e if $DF(n+1) - DF(n)$ goes towards zero for $n$ large enough, we determine $n_{th}$ as the number of appliances at which

$$|DF(n) - DF(N)| < \epsilon \times DF(N)\ \forall n > n_{th} \quad (4)$$

In the following, we use $\epsilon = 0.1$.

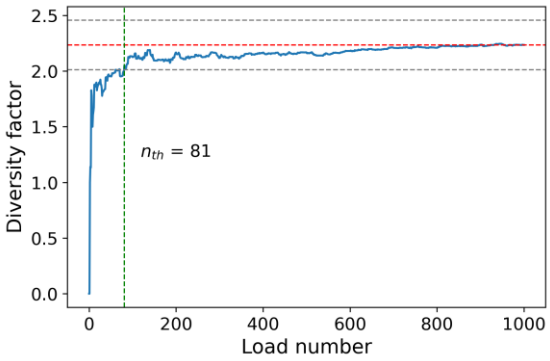



*Figure 2: Example of DF stabilization*

To determine $n_{th}$ for each appliance we make 50 simulations in which we compute DF(i) for $i \in [1, N]$ where N is the number of load curves for the considered appliances in our dataset, i.e. the size if the dataset and consequently the maximum number of such appliance load curves that we can sum. For each simulation we determine the number of appliances $n_{th}$ which satisfies inequality 4. Then we calculate the average of the 50 values of $n_{th}$ obtained which is then considered as the stabilization threshold of appliance under consideration.

## 2.7. Building level aggregated load profile model

In this paragraph, we present the method used to generate building-level aggregated load profiles. For a given building, several input parameters are user-defined such as: the number of occupants, the main activity and the location which enable the determination of appliance per user ratios for all appliances present in office buildings. Such ratios are determined based on the results of an inquiry conducted by authors in 2018 and 2019 which covers 1038 occupants of office buildings in France. The analysis and clustering of the inquiry answers lead to three main building categories whose appliance per user ratios are provided in Appendix A.

According to the user-defined input parameters, the considered building is assigned to one of the three categories and the ratios of the resulting category, as shown in Appendix A, are used for simulation. For each equipment $i$, the appliance / occupant ratio $r_i$ is multiplied by the user

defined number of occupants $n_{occ}$ to obtain the number of appliances for each appliance $n_{app,i} = r_i \times n_{occ}$. Then, for each appliance $i$, $n_{app,i}$ load curves are randomly sampled from the load profile datasets and summed. We obtain an aggregated load profile for each appliance type $i$ :

$$ALP(appliance\ i, n_{app,i}) = \sum_{j=1}^{n_{app,i}} IALP_j \quad (5)$$

where $IALP_j$ are randomly chosen from the load profiles dataset.

The sum in equation 5 is a point wise sum of $IALP_j$ vectors. We remind that with a 6-second timestep, each IALP contains 14400 data points corresponding to a 24h load profile.

Considering that a number $m$ of different appliances can be encountered and used in an office building (those different appliances were obtained from the above-mentioned inquiry), the aggregated load profile is then obtained by summing the aggregated load profiles of each appliance type as follows:

$$BALP = \sum_{k=1}^{m} ALP(appliance\ k, n_{app,k}) \quad (6)$$

where the sum is also a point wise sum.

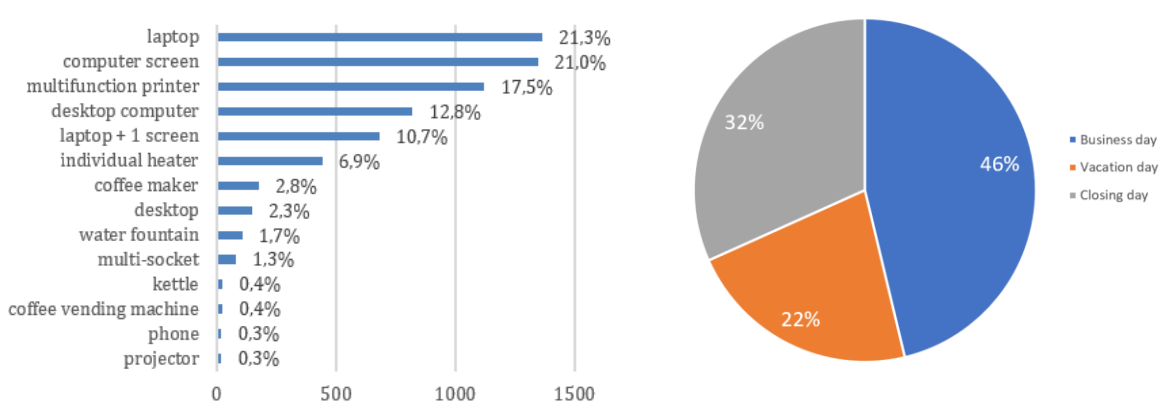
Building-level aggregated load profiles are then calculated for business days, closing days and vacation days which are intermediate occupancy days. The calculation of BALP for these three day types results from the observation of three different clusters corresponding to those day types in a previous work of the authors on the clustering of non-residential building load profile [40].

# 3. Results and discussion

## 3.1. Collected data statistics

Several measurement campaigns were performed in different office buildings for load profile measurements, 113 different appliances were monitored for a total of 767 days. In total, 6408 load curves were collected to account for the diversity of used devices and the occupants'

behaviors. We show in Figure 3 the distribution of the collected load profiles per appliance type and day-type.




(a) (b)

*Figure 3: Distribution of load profiles per appliance (a) and per day type (b).*

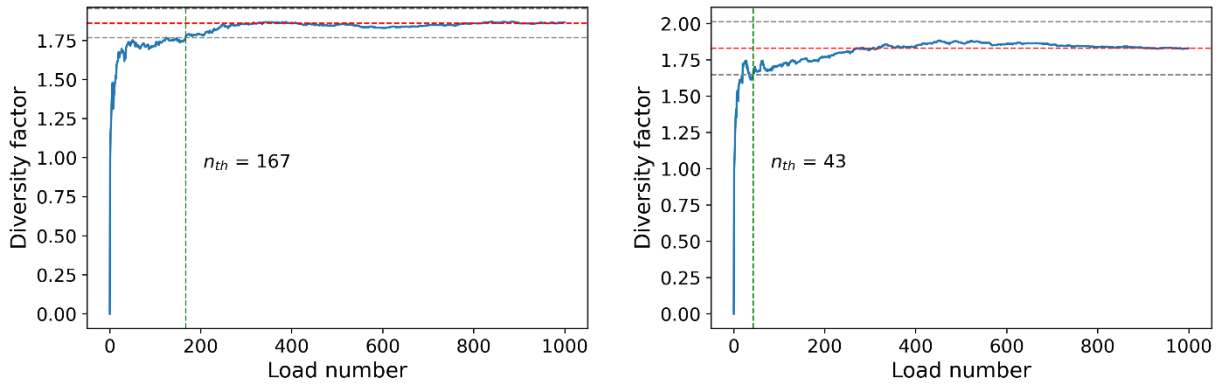
We can note that 14 different types of appliances were monitored during the different measurement campaigns. The majority of collected load profiles, 91% for instance, were measured on typical office equipment: desktop computers, laptops, computer screens, printers, etc. A much smaller part of the dataset, 9% in our case, were measured on less frequently used appliances such as coffee machines, kettles, telephones, water fountains and office presentation overhead projectors.

The measurement campaigns covered different day types: business days, closing days and vacation days. Closing days include weekends and statutory holidays. Vacation days include school holidays. In this work, we consider three day types instead of the commonly used two-day type classification since we have shown in a previous work that vacation day load profiles can be distinguished from closing and business day load profiles for several office buildings [35]. Indeed, occupancy during holidays is often lower than a typical business day while being larger than zero. The distribution of collected load profiles with respect to day types is as follows: 46% of business

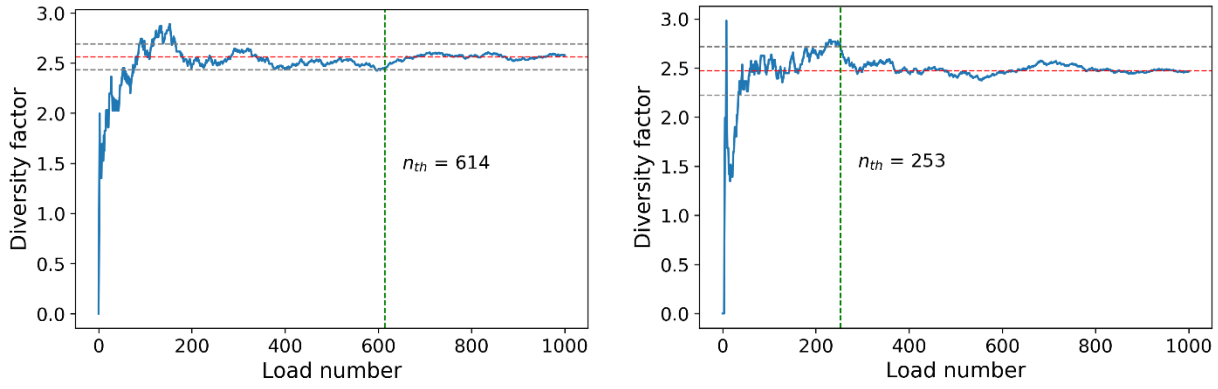
days, 32% of closing days and 22% of vacation days. When comparing the distribution of the days measured and the distribution of days in the French calendar, it can be concluded that the days of measurement are an adequate representation of the distribution of the different day types in a year.

## 3.2. Single appliance aggregated load profiles stabilization threshold

We calculated the diversity factor defined in equation 1 for all the appliances present in our dataset. Since the calculation of the DF is based on a random sampling of appliance load profiles, 50 runs have been performed for the DF calculation to assess the dependence of the results to the random sampling. As mentioned in section 2.5, the stabilization threshold $n_{th}$ depends on the stabilization interval width $[DF(N) - \epsilon, DF(N) + \epsilon]$, the stabilization threshold was calculated for $\epsilon = 0.05$ *and* $\epsilon = 0.1$ for each appliance type. Obtained results for two different appliances, desktop computers and kettles for instance, are plotted in Figure 4 for $\epsilon = 0.05$ and $\epsilon = 0.1$. Similar results for all appliances are provided in Appendix B. The dotted red line represents $DF(N)$ and the area in between the dotted grey lines indicates the stabilization interval $[DF(N) - \epsilon, DF(N) + \epsilon]$. The green dotted line indicates the stabilization threshold $n_{th}$.




(a) (b)




(c) (d)

*Figure 4: Diversity factors for $\epsilon = 0.05$ (left) and $\epsilon = 0.1$ (right) for desktop computers (a)(b) and kettles (c)(d).*

As can be observed in all subfigures of Figure 4, the DF tends to and stabilizes around a constant value for $n > n_{th}$. This is particularly obvious for desktop computers, computer screens, and laptops. Obtained results are summarized in Table 1 where we provide an average value of $n_{th}$ calculated after 50 runs for all appliances. The values of $n_{th}$ for the 50 runs are provided in Appendix C, Figure A.3 which enable to assess the dispersion of this value as a function of the random sampling of aggregated daily load profiles. For common office equipment such desktop computers, computer screens and laptops, a stabilization threshold lower than 300 for a stabilization interval width of $0.1 \times DF(N)$ and up to 127 for a stabilization interval width of $0.2 \times DF(N)$ are obtained. Such numbers of appliances is likely to be encountered in average-sized office buildings [42].

| **Appliance** | **Stabilization threshold (average of 50 runs)** | |
|---|---|---|
| | $\epsilon = 0.05$ | $\epsilon = 0.1$ |
| Computer screen | 93 | 32 |
| Desktop computer | 167 | 43 |

| | | |
|---|---|---|
| Laptop with one screen | 348 | 127 |
| Coffee machine | 709 | 417 |
| Individual heater | 368 | 129 |
| Multifunction device | 443 | 204 |
| Kettle | 616 | 258 |

*Table 1: Average stabilization thresholds $n_{th}$ for 50 runs for all appliances for $\epsilon = 0.05$ and $\epsilon = 0.1$.*

However, for other appliances such as multifunction devices, coffee machines and kettles which are often used as shared resources, or individual heaters which are much less frequently encountered, the stabilization occurs for a threshold too large compared to the number of such appliances that can be found in a reasonably sized building. For example, for specific devices such as kettles, stabilization is even much harder which explains the very large stabilization threshold. This could be explained by the discontinued use of these appliances throughout the day resulting in discontinuous load profiles with very sharp peaks a few times a day. Consequently, the diversity of individual kettle daily load profiles is very large. Therefore, the aggregation of a large number of appliances is required in order to observe a stabilized aggregated load profile. As can be shown later when discussing building-level aggregated load profiles, realistic scenarios of office buildings do not enable reaching the stabilization threshold for all appliances, especially for shared appliances.

We show in Figure 6 some examples of aggregated load profiles at the stabilization threshold per appliance type and per day type. Similar aggregated load profiles calculated for all appliance types are provided in Appendix D.

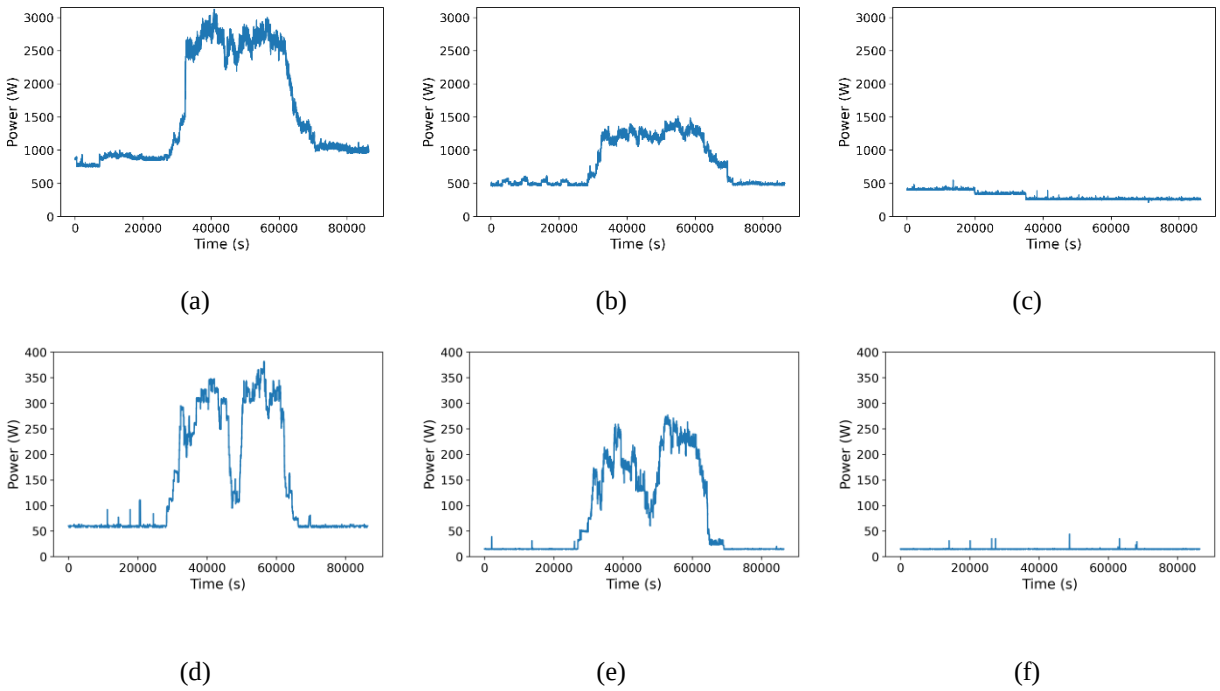



*Figure 6: Aggregated load profiles at $n_{th}$ per day type (left: business day, center: vacation day, right: closing day) for desktop computers (a)-(c) and single computer screens (d)-(f). The time, in seconds, shown in the x axis is representative of a 24-hour period.*

For all appliances, as can be expected, the business day load profiles exhibit the highest power levels, while closing days show the lowest electric demand levels. Electric power demand in weekdays in a vacation period exhibit intermediate levels between typical business days and closing days because of a lower occupancy since employees may take time off according to their children's vacations.

In addition to the load profiles, one can extract the occupancy information from aggregated load profiles of individual appliances such as personal computers or screens. Indeed, we observe the load increase in the morning suggesting a gradual arrival of the employees, a valley at noon around lunch break and a gradual decrease in the afternoon corresponding to the gradual departure of the employees.

The aggregated load profiles of some shared devices such as coffee machines and kettles show multiple peaks throughout the day. These appliance load profiles exhibit numerous short peak loads at specific times but are used less frequently than personal devices. Those peaks are often observed at the beginning of the day and at the beginning of the afternoon, as well as sporadically throughout the day. From the energy consumption point of view, the weight of those devices is not significant despite their very large peak loads because of their short duration of use.

Aggregated load profiles of desktop computers, laptops, computer screens and multifunction devices show non-zero power demand on closing days and during the night, up to 25% of their business day maximum load for desktop computers for example (Figure 6-a). During those times, these appliances could have possibly been left on standby mode. Optimizing the management of those appliances by a systematic shutdown during non-working hours, for example, could be a step towards reducing an office building's electric energy consumption. Even though this point has been highlighted decades ago [43], it still seems not to be systematically implemented yet.

Aggregated load profiles such as those shown in Figure 6 and Appendix D can also be very useful for building energy modeling enhancement. Indeed, beyond a certain number of occupants, for most of the appliances, one aggregated load profile could be used to simulate the electric power demand of the appliances' stock in an office building. However, this strategy cannot be used for small buildings or for appliances with a high stabilization threshold such, where a random sampling of single appliances load curves can be used. We further explore this idea in the following section.

## 3.3.Building level aggregated load profiles per building type

In this section, we calculate and discuss the building-level aggregated load profiles (BALP) for different office building types and sizes and for the above-mentioned three day types. A BALP is

a load profile including the load profiles of a number of different appliances according to a given distribution. The different distributions were obtained based on a field survey which collected and analyzed the responses of more than 1000 office building occupants in France in 2019. An automatic clustering [8] of the responses provided three main clusters resulting in three office buildings types considered in the following paragraphs. Each building type is characterized by a set of appliances per occupant ratio for all considered appliances. Those ratio values range from 0 to 1. For example, for an individual appliance such as a personal computer in a building where all occupants have their own personal computer, the ratio is equal to 1. For a shared device, such as a multifunction device for example shared between n users, the ratio is equal to 1/n. All ratios calculated and used in the present work are provided in Appendix A. The highest ratios for building type 1 are for desktop computers, computer screens and laptops. The highest ratios for building type 2 are for laptops and computer screens. For building type 3, the highest ratios are for coffee machines, kettles, desktop computers and laptops. For each building type, we consider two building sizes, represented here by the number of occupants n and compared to the stabilization thresholds of existing appliances. We consider $n_1 < n_{th,min}$ and $n_2 \geq n_{th,max}$for the three building types where $n_{th,min}$ is the smallest stabilization threshold amongst the appliances present in each building type and $n_{th,max}$ is the largest stabilization threshold in the same building type. We show in Figure 7 and 8 the aggregated load profiles for these two cases for the three buildings types and the three day types.

**Case 1: below the stabilization threshold** $(n_1 < n_{th,min})$

For this case, any number below the stabilization threshold of all the appliances could be considered. We considered $n_1 = 100$ in our calculations to obtain the aggregated load profiles for each building type shown in Figure 7.

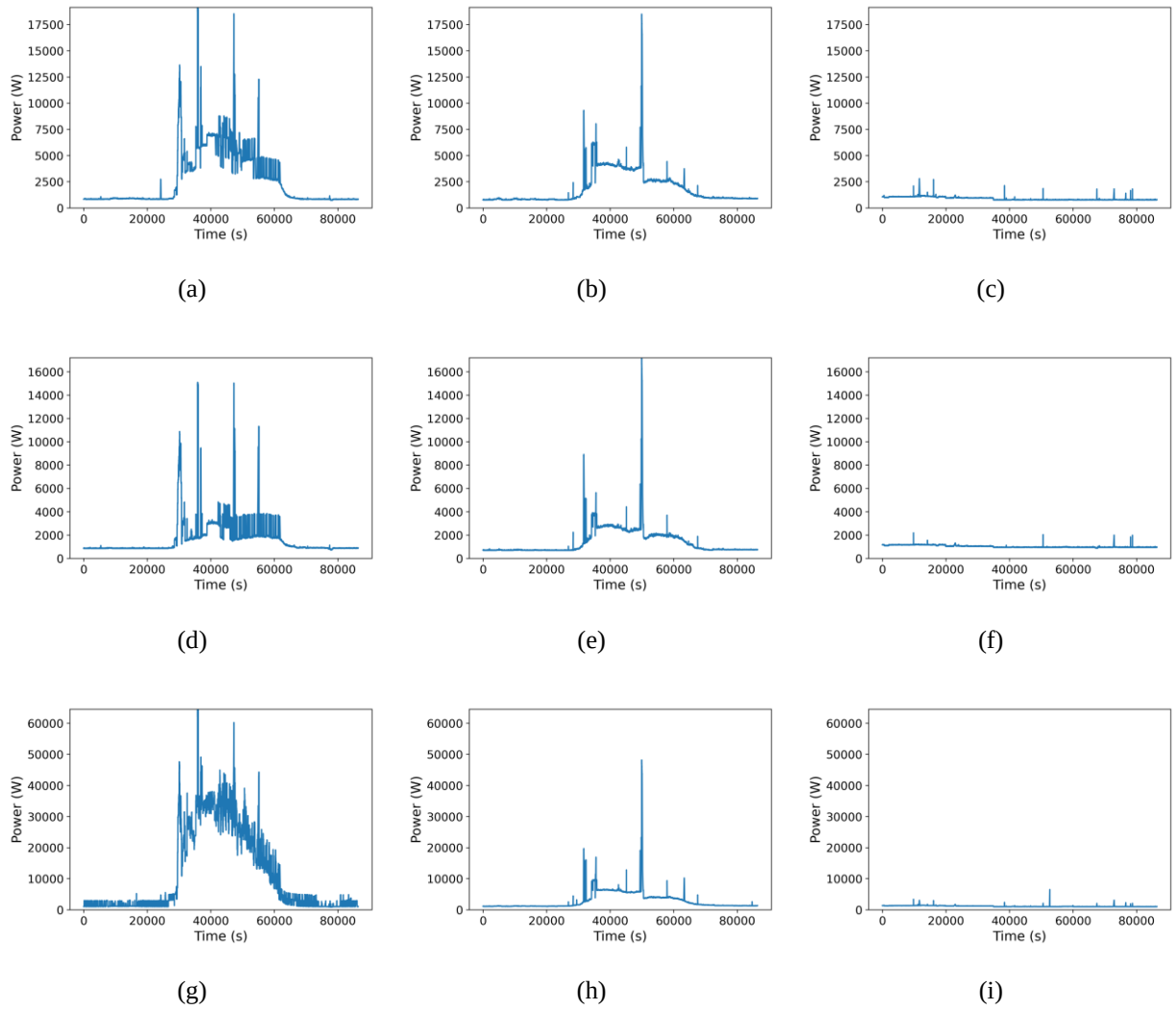


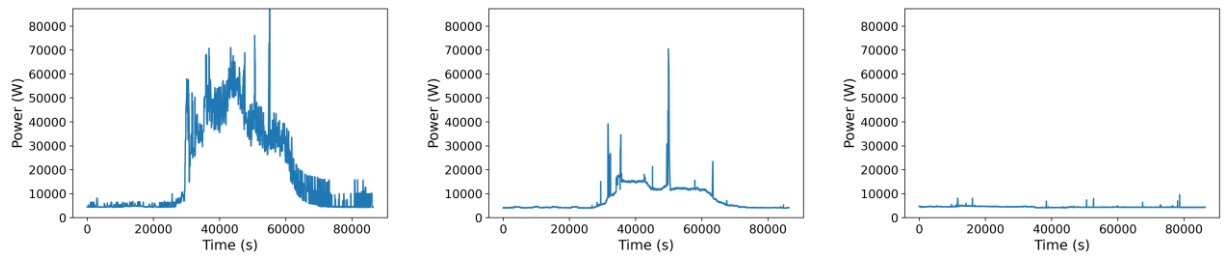

*Figure 7: Aggregated load profiles for $n_1$ = 100 occupants per day type (left: business day, center: vacation day, right: closing day) for building type 1 (a)-(c), building type 2 (d)-(f) and building type 3 (g)-(i).*

We observe the same trend of a decreasing electric power load as for SAALP when switching from business days, to vacation days to closing days. We also observe non-nil electric load during closing days for buildings types 1 and 2 where the dominating plug appliances are individual office appliances such as computers and screens.

We observe similar maximum power-demand levels for buildings 1 and 2 while building 3 exhibits a much larger maximal power. Indeed, buildings 1 and 2 are characterized by similar appliance to occupant ratios This can be explained by the large appliance per occupant ratio for heat production appliances such as individual heaters, kettles and coffee machines in this building type. Those appliances' nominal electric power demand is much larger than that of typical office equipment which explains the particularly large load levels in Figure 7-g.

**Case 2: above the stabilization threshold** $(n_2 \geq n_{th,max})$

Using Table 1 as a reference, any number of occupants larger than the largest stabilization threshold should be sufficient for all appliances SAALP to stabilize and therefore to obtain a stabilized BALP. A simple calculation shows that 2 573 occupants are required for this purpose (Appendix A). However, as discussed in the previous sections, the dominating plug appliances in an office building are individual appliances, except for buildings with on plug individual electric heating as shown in Figure 7-g. As a reasonable assumption, we consider the stabilization threshold of desktop computers, laptops and computer screens only. Consequently, 500 occupants are considered in the following since it enables reaching those appliances' stabilization thresholds. Obtained BALP are shown in Figure 8. Similar trends to those of Figure 7, *i.e.* for buildings with 100 occupants, are observed here when comparing different day types and different buildings types.



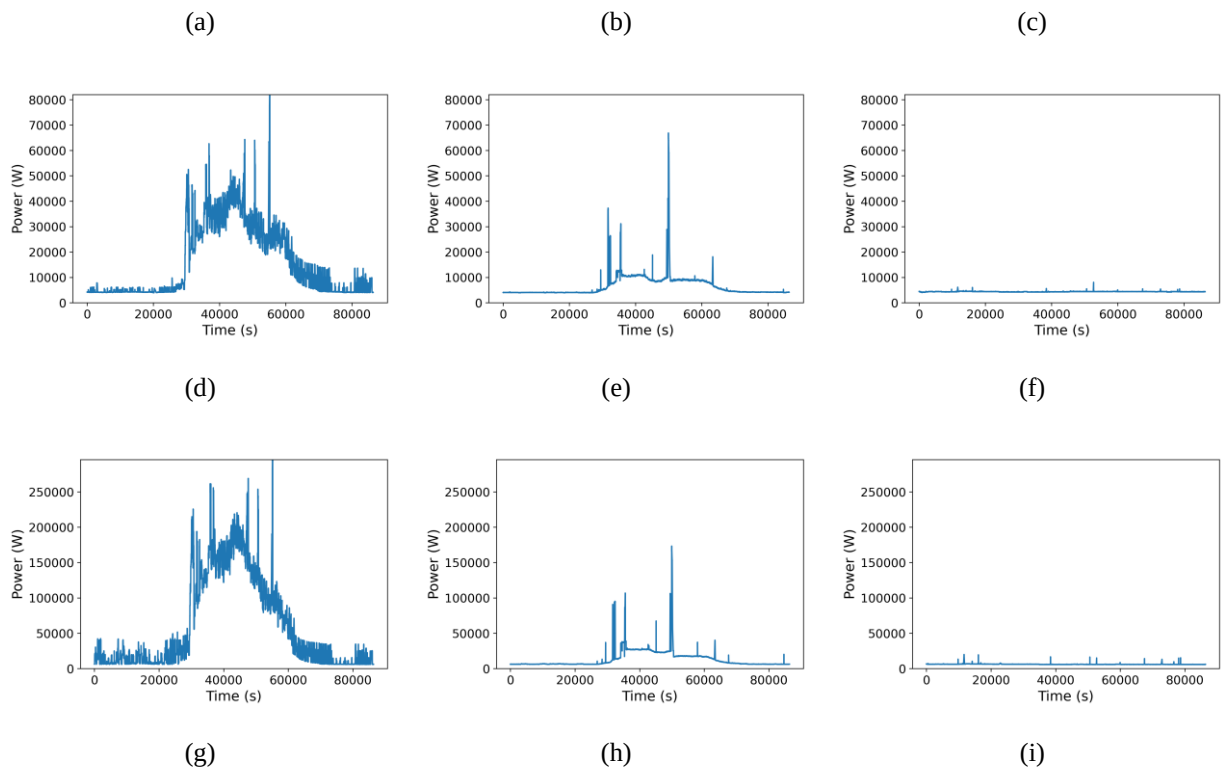



*Figure 8: Aggregated load profiles for $n_2$ = 500 occupants per day type (left: business day, center: vacation day, right: closing day) for building type 1 (a)-(c), building type 2 (d)-(f) and building type 3 (g)-(i)*

To compare the aggregated load profiles for different building occupancy numbers ($n = n_1 \; and \; n = n_2$), we calculated the difference between the normalized BALP for $n_1$ = *100 and* $n_2$ = *500* for the different buildings types and different day types. The use of normalized load profiles to calculate the difference enables the comparison of load profiles shapes only without considering the difference in loads base levels which also depend on the number of occupants and would obviously diverge with increasing occupancy contrary to what is expected for the load profile shape. Obtained results are shown in Figure 9.

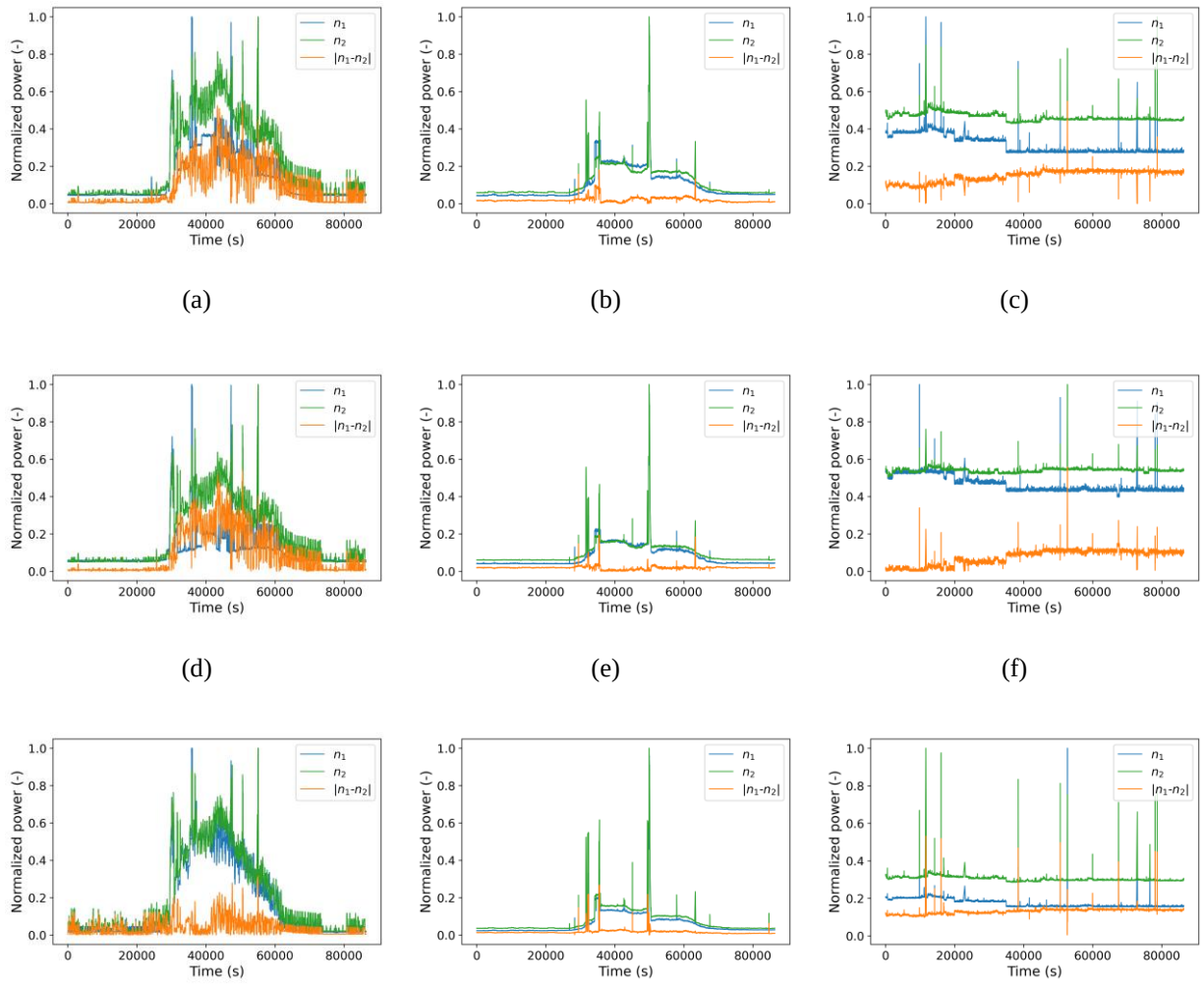



*Figure 9: Difference between the normalized aggregated load profiles for $n_1$ = 100 and $n_2$ = 500 occupants per day type (left: business day, center: vacation day, right: closing day) for building type 1 (a), building type 2 (b) and building type 3 (c)*

We clearly observe a non-negligible difference between building 1 and 2 BALP for $n_1$ and $n_2$ occupants while the difference is much closer to zero for building type 3. We can conclude that the BALP does not stabilize for building 1 and 2 for $n = n_1$ while, by construction, we are sure for this stabilization for $n = n_2$. Consequently, to generate aggregated load curves for $n \leq n_{th,max}$

for those two building types, a random sampling is recommended in order to account for the diversity of the appliance's electric loads and the use of the stabilized BALP is not recommended. For building type 3, the very small difference between calculated BALP for $n = n_1$ and $n = n_2$ is because this building type's aggregated load profile is dominated by heat generation equipment such as individual heaters which exhibit very small diversity since they are often operated continuously. Therefore, the diversity of typical office equipment load profiles which require a large number of appliances to observe BALP stabilization has less importance here. One should note that we do not account here for the obvious seasonal effect that will reduce heater usage during the summer and warm days. This third category of buildings can probably be distinguished from the other two in cold days when heating load is not negligible.

On vacation and closing days for all building types, the normalized aggregated load profiles for both cases are quite similar or almost flat, which explains the very small difference between $n_1$=100 and $n_2 = 500$. This is probably due the lower diversity of power demand during these days. Consequently, a simple strategy consisting in using a stabilized load profile multiplied by the number of occupants can be adopted for these day types which may be useful to simplify such data computational time and provide simple realistic input to building energy models.

## 4. Conclusions

We presented in this paper a study of aggregated plug load profiles in office buildings based on more than 6400 field-measured office appliances load profiles and a field survey with more than 1000 responses from office buildings occupants in France to quantify the different appliance distributions in office buildings. First, we considered single appliance type aggregated load profiles. We showed that the aggregated load profile shapes stabilize for a number of aggregated appliances large enough and that the stabilization threshold depends on the appliance type.

Aggregated load profiles of individual appliances like desktop computers, laptops and screens are more likely to stabilize in buildings of a few hundred occupants, which enables the use of stabilized load profiles to simulate realistic plug load profiles for such buildings, while for other appliances, such as multifunction devices, the stabilization threshold is so high that its aggregated load profile is not likely to stabilize in a reasonably sized building. In this case, a random sampling of a specific number of appliance load profiles is more suitable for an aggregated load profile generation. This stabilization threshold is calculated and discussed for all considered appliances for the first time in the literature. These two main conclusions were then implemented in a tool for building-level aggregated plug load profile generation. The analysis of the survey results enables the definition of three buildings categories exhibiting different appliance distributions. We generated aggregated plug load profiles for these three building categories for different day types and sizes. We show different dependencies of the aggregated load profile shapes to the buildings sizes and day types which suggest different possible strategies for aggregated load profile generation to account for the diversity of appliances and the occupants' behaviors. An important future step of the present research will be the validation of the generated load curves through field load curves monitoring for different buildings at the building level as well as the appliances stock level inside one specific building. This validation will require other measurement campaigns or open dataset of relevant data for cross validation. Obtained results might be useful for a large community of researchers and engineers interested in the energy consumption and load profiles of office buildings for various applications such as demand response, renewable energy generation systems sizing and energy efficiency.

## 5. Acknowledgments

Authors would like to thank W. Wu and H. Hoxha for contributions to the Python codes used for data processing and D. Da Silva and T. Guiot for fruitful discussions and help.

## 6. References


[1] International Energy Agency, « Data & Statistics », 2018.

[2] International Energy Agency, « Tracking Buildings – Analysis », 2020. https://www.iea.org/reports/tracking-buildings (consulté le avr. 20, 2020).

[3] B. N. Silva, M. Khan, et K. Han, « Towards sustainable smart cities: A review of trends, architectures, components, and open challenges in smart cities », *Sustain. Cities Soc.*, vol. 38, nº August 2017, p. 697-713, 2018, doi: 10.1016/j.scs.2018.01.053.

[4] D. R. Obey, « Text - H.R.1 - 111th Congress (2009-2010): American Recovery and Reinvestment Act of 2009 », 2009. https://www.congress.gov/bill/111th-congress/house-bill/1/text

[5] U.S. Energy Information Administration (EIA), « How many smart meters are installed in the United States, and who has them? », 2018. https://www.eia.gov/tools/faqs/faq.php?id=108&t=3

[6] European Commission - Joint Centre | Smart Electricity Systems and Interoperability, « Smart Metering deployment in the European Union | JRC Smart Electricity Systems and Interoperability ». https://ses.jrc.ec.europa.eu/smart-metering-deployment-european-union

[7] A. Foucquier, S. Robert, F. Suard, L. Stéphan, et A. Jay, « State of the art in building modelling and energy performances prediction: A review », *Renew. Sustain. Energy Rev.*, vol. 23, p. 272-288, juill. 2013, doi: 10.1016/J.RSER.2013.03.004.

[8] M. Bourdeau, X. qiang Zhai, E. Nefzaoui, X. Guo, et P. Chatellier, « Modeling and forecasting building energy consumption: A review of data-driven techniques », *Sustain. Cities Soc.*, vol. 48, p. 101533, juill. 2019, doi: 10.1016/j.scs.2019.101533.

[9] O. T. Masoso et L. J. Grobler, « The dark side of occupants' behaviour on building energy use », *Energy Build.*, vol. 42, nº 2, p. 173-177, févr. 2010, doi: 10.1016/j.enbuild.2009.08.009.

[10] M. S. Piscitelli, S. Brandi, et A. Capozzoli, « Recognition and classification of typical load profiles in buildings with non-intrusive learning approach », *Appl. Energy*, vol. 255, nº August, p. 113727, 2019, doi: 10.1016/j.apenergy.2019.113727.

[11] I. P. Panapakidis, T. A. Papadopoulos, G. C. Christoforidis, et G. K. Papagiannis, « Pattern recognition algorithms for electricity load curve analysis of buildings », *Energy Build.*, vol. 73, p. 137-145, avr. 2014, doi: 10.1016/j.enbuild.2014.01.002.

[12] M. Quintana, P. Arjunan, et C. Miller, « Islands of misfit buildings: Detecting uncharacteristic electricity use behavior using load shape clustering », *Build. Simul.*, nº October, 2019, doi: 10.13140/RG.2.2.11489.86883.

[13] F. McLoughlin, A. Duffy, et M. Conlon, « A clustering approach to domestic electricity load profile characterisation using smart metering data », *Appl. Energy*, vol. 141, p. 190-199, mars 2015, doi: 10.1016/j.apenergy.2014.12.039.

[14] M. Richard, H. Fortin, A. Poulin, et M. Leduc, « Daily load profiles clustering : a powerful tool for demand side management in medium-sized industries », *ACEEE Summer Study Energy Effic. Ind.*, p. 160-171, 2017.

[15] S. Yilmaz, J. Chambers, et M. K. Patel, « Comparison of clustering approaches for domestic electricity load profile characterisation - Implications for demand side management », *Energy*, vol. 180, p. 665-677, août 2019, doi: 10.1016/j.energy.2019.05.124.

[16] G. Chicco, R. Napoli, et F. Piglione, « Load pattern clustering for short-term load forecasting of anomalous days », in *2001 IEEE Porto Power Tech Proceedings*, 2001, vol. 2, p. 217-222. doi: 10.1109/PTC.2001.964745.
[17] A. Satre-Meloy, M. Diakonova, et P. Grünewald, « Cluster analysis and prediction of residential peak demand profiles using occupant activity data », *Appl. Energy*, vol. 260, févr. 2020, doi: 10.1016/j.apenergy.2019.114246.
[18] A. Lavin et D. Klabjan, « Clustering time-series energy data from smart meters », *Energy Effic.*, vol. 8, nº 4, p. 681-689, juill. 2015, doi: 10.1007/s12053-014-9316-0.
[19] E. C. Bobric, G. Cartina, et G. Grigoraş, « Clustering techniques in load profile analysis for distribution stations », *Adv. Electr. Comput. Eng.*, vol. 9, nº 1, p. 63-66, 2009, doi: 10.4316/aece.2009.01011.
[20] F. Spertino, G. Chicco, A. Ciocia, S. Corgnati, P. Di Leo, et D. Raimondo, « Electricity consumption assessment and PV system integration in grid-connected office buildings », in *2015 IEEE 15th International Conference on Environment and Electrical Engineering, EEEIC 2015 - Conference Proceedings*, juill. 2015, p. 255-260. doi: 10.1109/EEEIC.2015.7165548.
[21] G. Chicco, « Overview and performance assessment of the clustering methods for electrical load pattern grouping », *Energy*, vol. 42, nº 1, p. 68-80, juin 2012, doi: 10.1016/j.energy.2011.12.031.
[22] J. Yang *et al.*, « k-Shape clustering algorithm for building energy usage patterns analysis and forecasting model accuracy improvement », *Energy Build.*, vol. 146, p. 27-37, juill. 2017, doi: 10.1016/j.enbuild.2017.03.071.
[23] U. DOE, « An assessment of energy technologies and research opportunities », *Quadrenn. Technol. Rev. U. S. Dep. Energy*, 2015.
[24] I. Metzger, A. Kandt, et O. VanGeet, « Plug load behavioral change demonstration project », National Renewable Energy Lab.(NREL), Golden, CO (United States), 2011.
[25] M. K. Kim, Y.-S. Kim, et J. Srebric, « Impact of correlation of plug load data, occupancy rates and local weather conditions on electricity consumption in a building using four back-propagation neural network models », *Sustain. Cities Soc.*, vol. 62, p. 102321, nov. 2020, doi: 10.1016/j.scs.2020.102321.
[26] Residential End Use Monitoring Program (REMP), « General Plug Loads Data Collection and Analysis », E3 : Equipment Energy Efficiency, avr. 2012.
[27] R. American Society of Heating, A.-C. Engineers, et Ashrae, *2013 ASHRAE Handbook: Fundamentals*. ASHRAE, 2013.
[28] Chartered Institution of Building Services Engineers, *Energy efficiency in buildings: CIBSE guide F*. 2012.
[29] S. Gilani et W. O'Brien, « Quantification of building energy performance uncertainty associated with building occupants and operators », présenté à 7th International Building Physics Conference, IBPC18, sept. 2018.
[30] C. Carpino, E. Loukou, P. Heiselberg, et N. Arcuri, « Energy performance gap of a nearly Zero Energy Building (nZEB) in Denmark: the influence of occupancy modelling », *Build. Res. Inf.*, vol. 48, nº 8, p. 899-921, nov. 2020, doi: 10.1080/09613218.2019.1707639.
[31] S. Poll et C. Teubert, « Pilot Study of a Plug Load Management System: Preparing for Sustainability Base », in *2012 IEEE Green Technologies Conference*, avr. 2012, p. 1-6. doi: 10.1109/GREEN.2012.6200945.

[32] R. Tuttle, K. Trenbath, K. Maisha, et A. LeBar, « Assessing and Reducing Plug and Process Loads in Office Buildings », National Renewable Energy Lab.(NREL), Golden, CO (United States), 2020.
[33] D. Kaneda, B. Jacobson, P. Rumsey, et R. Engineers, « Plug load reduction: The next big hurdle for net zero energy building design », in *ACEEE Summer Study on Energy Efficiency in Buildings*, 2010, p. 120-130.
[34] D. Mitra, N. Steinmetz, Y. Chu, et K. S. Cetin, « Typical occupancy profiles and behaviors in residential buildings in the United States », *Energy Build.*, vol. 210, p. 109713, 2020, doi: https://doi.org/10.1016/j.enbuild.2019.109713.
[35] E. Wilson, C. Engebrecht-Metzger, S. Horowitz, et R. Hendron, « 2014 building america house simulation protocols », National Renewable Energy Lab.(NREL), Golden, CO (United States), 2014.
[36] A. C. Menezes, A. Cripps, R. A. Buswell, et D. Bouchlaghem, « Benchmarking small power energy consumption in office buildings in the United Kingdom: A review of data published in CIBSE Guide F », *Build. Serv. Eng. Res. Technol.*, vol. 34, nº 1, p. 73-86, févr. 2013, doi: 10.1177/0143624412465092.
[37] A. C. Menezes, A. Cripps, R. A. Buswell, J. Wright, et D. Bouchlaghem, « Estimating the energy consumption and power demand of small power equipment in office buildings », *Energy Build.*, vol. 75, p. 199-209, juin 2014, doi: 10.1016/j.enbuild.2014.02.011.
[38] H. B. Gunay, W. O'Brien, I. Beausoleil-Morrison, et S. Gilani, « Modeling plug-in equipment load patterns in private office spaces », *Energy Build.*, vol. 121, p. 234-249, juin 2016, doi: 10.1016/j.enbuild.2016.03.001.
[39] « Current Cost - Reducing your energy bills so you can live a greener life! » http://currentcost.com/product-envir-specifications.html (consulté le févr. 16, 2021).
[40] M. Bourdeau *et al.*, « Classification of daily electric load profiles of non-residential buildings », *Energy Build.*, vol. 233, p. 110670, févr. 2021, doi: 10.1016/j.enbuild.2020.110670.
[41] « Insee, enquête Emploi en continu sur l'année 2010. », Institut national de la statistique et des études économiques (INSEE), 2010.
[42] J. Michaels, « 2018 Commercial Buildings Energy Consumption Survey (CBECS) », U.S. Energy Information Administration, Office of Energy Statistics, U.S. Department of Energy, Washington, DC, nov. 2020. Consulté le: févr. 25, 2021. [En ligne]. Disponible sur: https://www.eia.gov/consumption/commercial/pdf/CBECS%202018%20Preliminary%20Results%20Flipbook.pdf
[43] C. A. Webber, J. A. Roberson, M. C. McWhinney, R. E. Brown, M. J. Pinckard, et J. F. Busch, « After-hours power status of office equipment in the USA », *Energy*, vol. 31, nº 14, p. 2823-2838, nov. 2006, doi: 10.1016/j.energy.2005.11.007.

# 7. Appendices

## 7.1. Appendix A: Appliance / occupant ratio per appliance type for office buildings

To determine the different building types and their appliance per occupant ratios, an inquiry was conducted among 1033 participants. 109 different building profile types, based on a building's workforce size, location and main activity, were represented in the responses. In order to identify a smaller (reasonable) number of typical building profiles, the k-means clustering algorithm was applied to the 109 building profiles. The key feature for cluster formation is the appliance per occupant ratio for each appliance type. Indeed, it is the key information needed for our aggregated load curves generation algorithm. Four clusters were obtained, however, one of the fours clusters contained one single value. Therefore, only three clusters were conserved. As a result, three "typical" building types, as well as their appliance per occupant ratios were determined. The appliance per occupant ratios for each cluster is the average of the ratios of the individuals in the cluster.

From the 109 building profiles, 47% are represented in building type 1, 44% in building type 2 and 9% in building type 3.

| **Appliance** | **Building type 1** | **Building type 2** | **Building type 3** |
|---|---|---|---|
| Computer screen | 0.890965 | 0.207358 | 1.1675 |
| Desktop computer | 0.608468 | 0.142564 | 0.840833 |
| Coffee machine | 0.295835 | 0.246464 | 1.1325 |
| Laptop computer with one screen | 0.55412 | 0.912062 | 0.7725 |
| Individual heater | 0.266988 | 0.200852 | 0.973333 |
| Multifunction device | 0.181958 | 0.109208 | 0.291667 |
| Kettle | 0.265219 | 0.270015 | 0.893333 |

*Table A1: Appliance per occupant ratios for each building type*

The following table summarizes the number of appliances per appliance type in a building with 500 occupants. At 500 occupants, the stabilization threshold of computer screens, desktop computers, laptops with one screen and coffee machines are reached. Through further calculations,

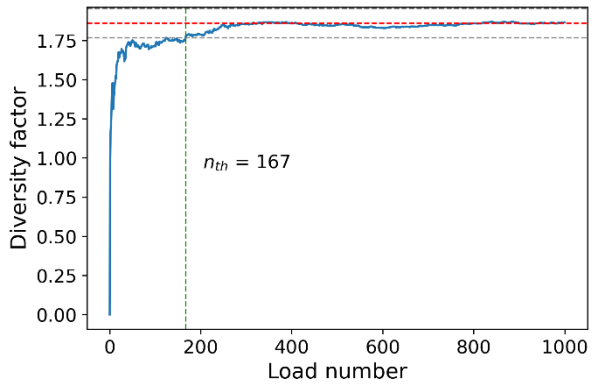
to reach the stabilization threshold of all the appliances and for each building type, the office building must have 1 692 occupants.

| Appliance | Stabilization threshold | Number of appliances in a 500 occupant building | | |
|---|---|---|---|---|
| | | Building type 1 | Building type 2 | Building type 3 |
| Computer screen | 32 | 445 | 104 | 584 |
| Desktop computer | 43 | 304 | 71 | 420 |
| Coffee machine | 417 | 148 | 123 | 566 |
| Laptop with one screen | 127 | 277 | 456 | 386 |
| Individual heater | 129 | 133 | 100 | 487 |
| Multifunction device | 204 | 91 | 55 | 146 |
| Kettle | 258 | 133 | 135 | 447 |

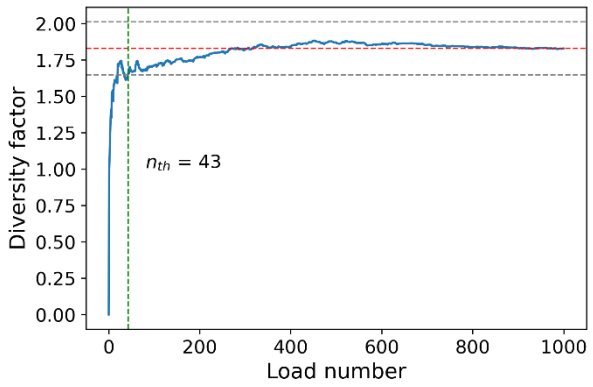
*Table A2: Number of appliances in a 500-occupant building*

## 7.2. Appendix B: Diversity factor stabilization for different appliances

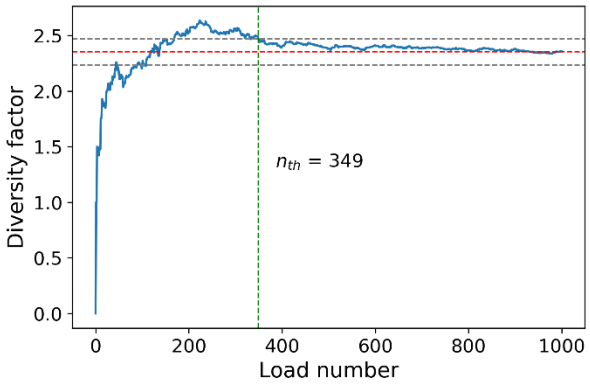
The following figure of diversity factors stabilization provides additional information with respect to other appliances not shown in Figure 4.



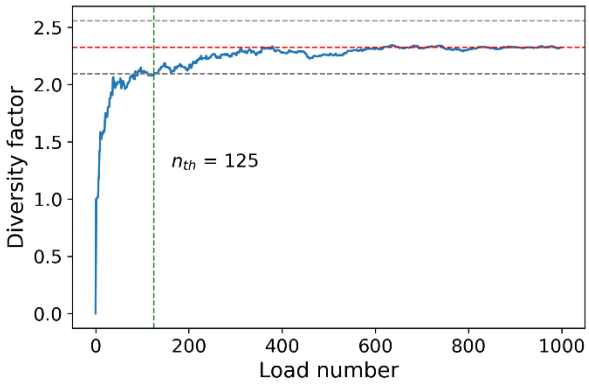

(a)



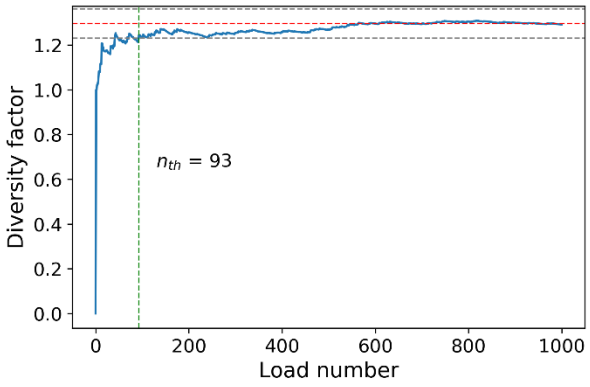

(b)


Diversity factor
Load number
$n_{th}$ = 349

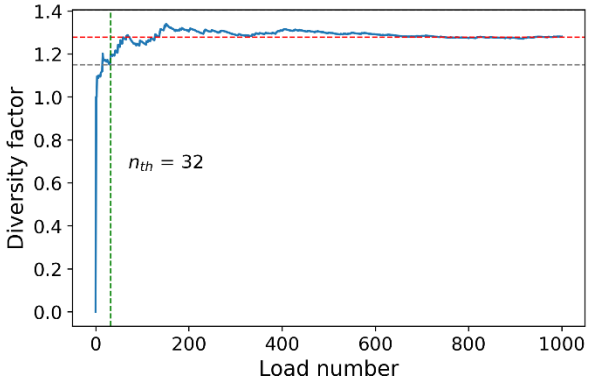
(c)


Diversity factor
Load number
$n_{th}$ = 125

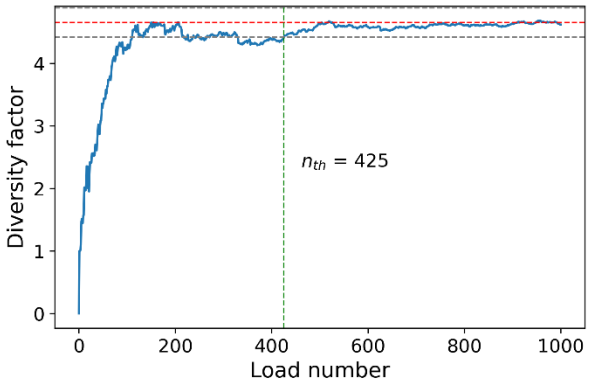
(d)


Diversity factor
Load number
$n_{th}$ = 93

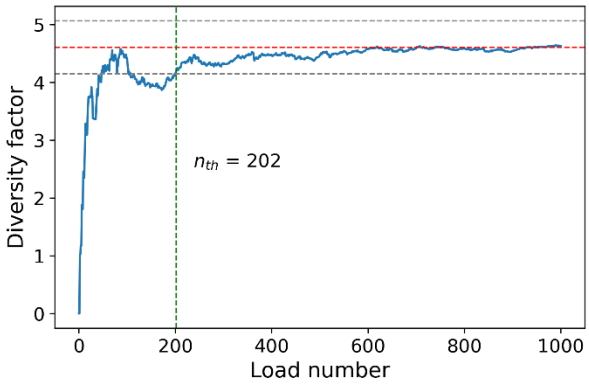
(e)


Diversity factor
Load number
$n_{th}$ = 32

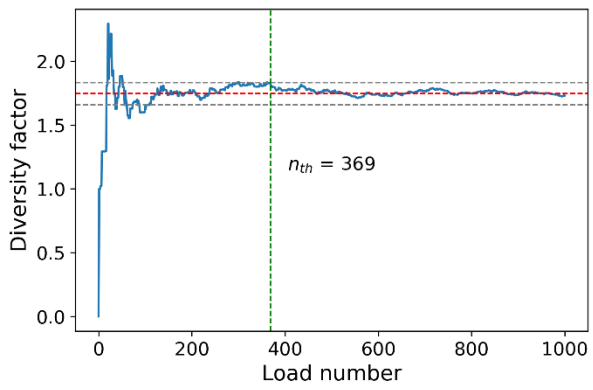
(f)


Diversity factor
Load number
$n_{th}$ = 425

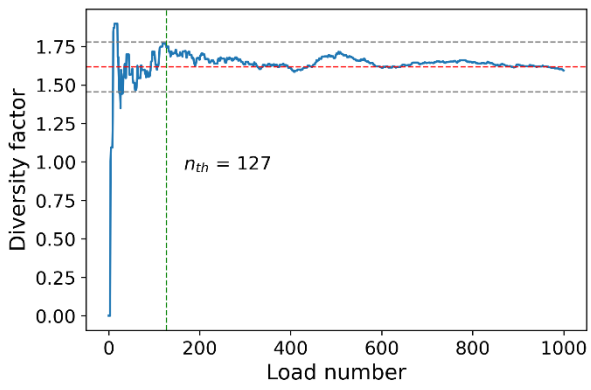
(g)


Diversity factor
Load number
$n_{th}$ = 202

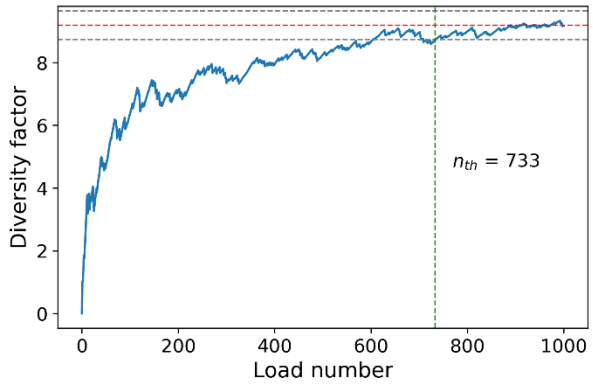
(h)



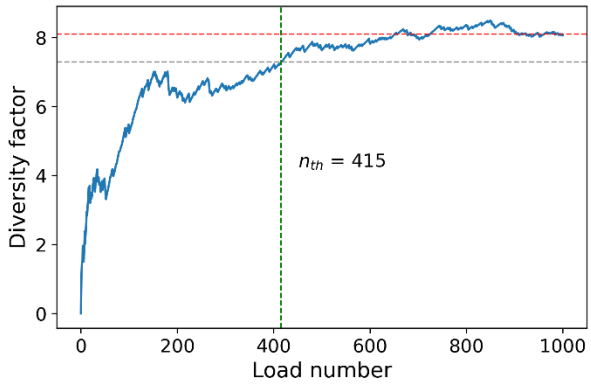

(i)



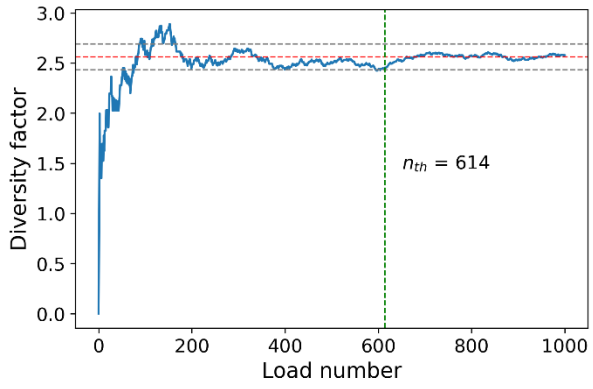

(j)



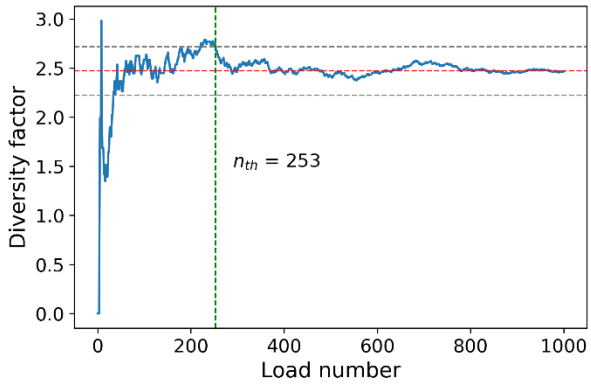

(k)




(l)




(m)




(n)

*Figure B1: Diversity factors for $\epsilon = 0.05$(left) and $\epsilon = 0.10$(right) for desktop computers (a)(b), laptops with one screen (c)(d), computer screens (e)(f), multifunction devices (g)(h), individual heaters (i)(j), coffee machines (k)(l) and kettles (m)(n)*

## 7.3. Appendix C: Stabilization threshold and random sampling

There is a significant gap between the stabilization thresholds of $\epsilon = 0.05$ and $\epsilon = 0.10$. To determine which $\epsilon$ had more uniform results, the stabilization thresholds for the 50 runs for each appliance and each $\epsilon$ were plotted in the following scatter plots.

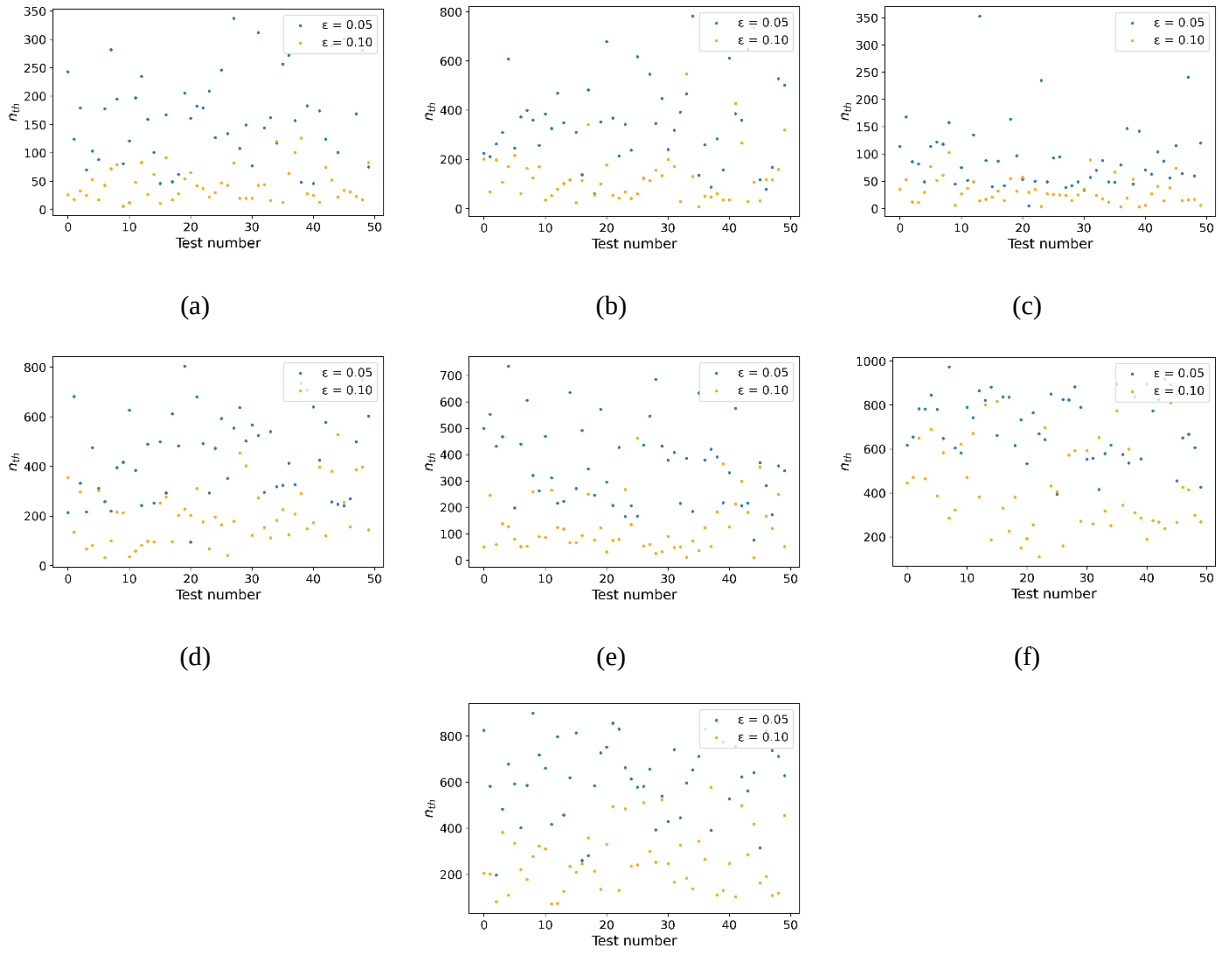



*Figure C1: Dispersion of $n_{th}$ for $\epsilon = 0.05$ and $\epsilon = 0.10$ for desktop computers (a), laptops with one screen (b), screens (c), multifunction devices (d), individual heaters (e), coffee machines (f) and kettles (g) for 50 calculations*

For desktop computers, laptops with one screen, and screens, the diversity stabilization thresholds for $\epsilon = 0.10$ show more uniform results, than for $\epsilon = 0.05$. It is unclear which $\epsilon$ generates more

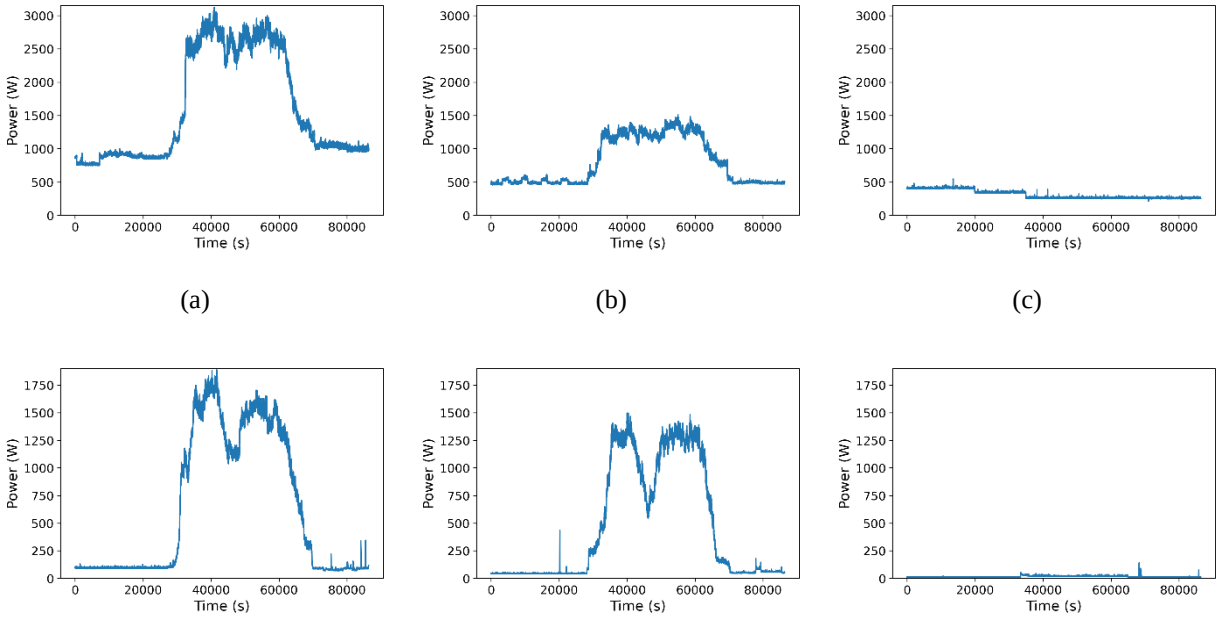
uniform results for individual heaters, multifunction devices, kettles and coffee machines. Since those appliances, either do not stabilize or their stabilization threshold is unrealistic, they can be disregarded. The diversity stabilization thresholds for $\epsilon = 0.10$ will be considered. We show in table A3 the average and standard deviation of the stabilization threshold obtained for all appliances for $\epsilon = 0.05$ and $\epsilon = 0.1$.

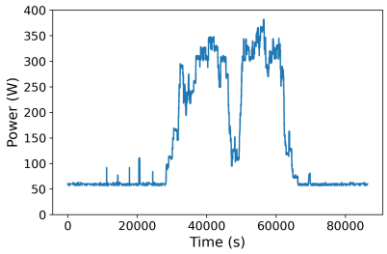
| Appliance | Stabilization threshold | | | |
|---|---|---|---|---|
| | $\mu(\epsilon = 0.1)$ | $\sigma(\epsilon = 0.1)$ | $\mu(\epsilon = 0.05)$ | $\sigma(\epsilon = 0.05)$ |
| Screen | 31,7 | 23,0 | 92,8 | 61,3 |
| Desktop Computer | 43,3 | 29,1 | 166,6 | 79,7 |
| Laptop | 127,3 | 106,0 | 348,3 | 178,7 |
| Heater | 128,6 | 102,1 | 367,9 | 155,1 |
| mfd | 204,1 | 118,7 | 442,9 | 171,6 |
| Kettle | 257,6 | 132,5 | 615,9 | 169,8 |
| Coffee maker | 417,4 | 194,3 | 709,5 | 147,1 |

*Table A3 : stabilization threshold average and standard deviation for all appliances for $\epsilon = 0.05$ and $\epsilon = 0.1$. obtained after 50 runs for each appliance.*

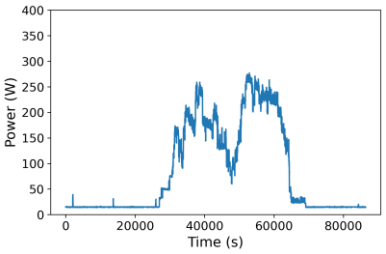
## 7.4. Appendix D: Single Appliance type Aggregated Load Profiles

The following figure of single appliance type aggregated load profiles is complementary to Figure 6.

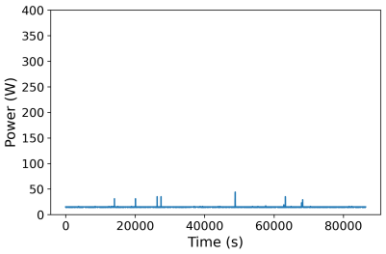




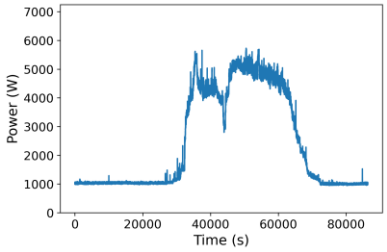

(g)



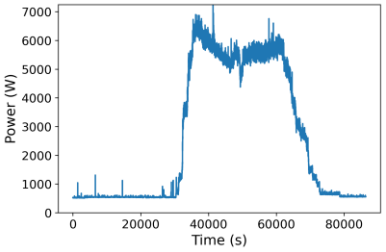

(h)



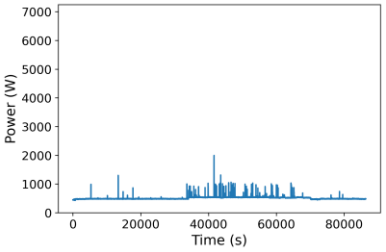

(i)



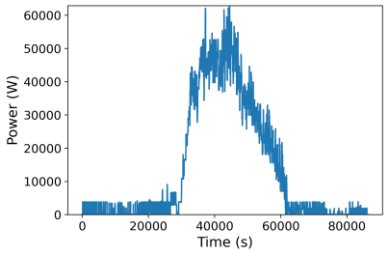

(j)



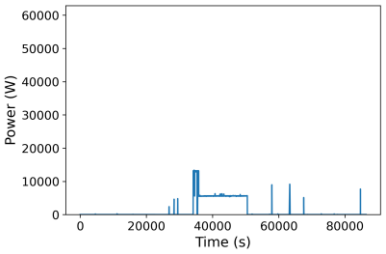

(k)



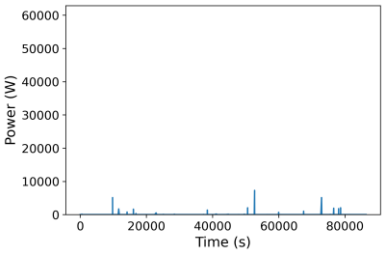

(l)



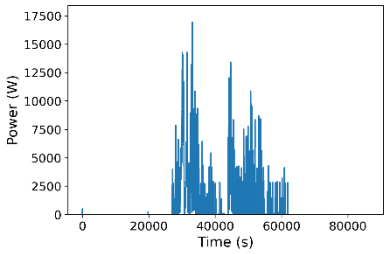

(m)



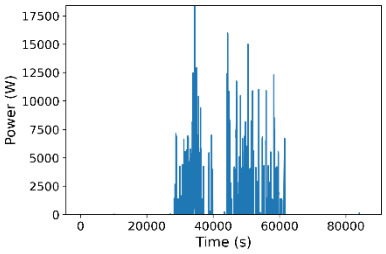

(n)



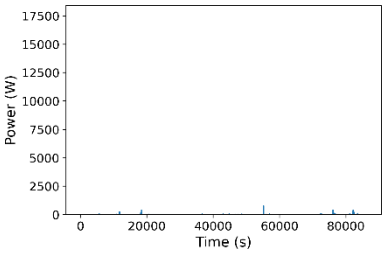

(p)



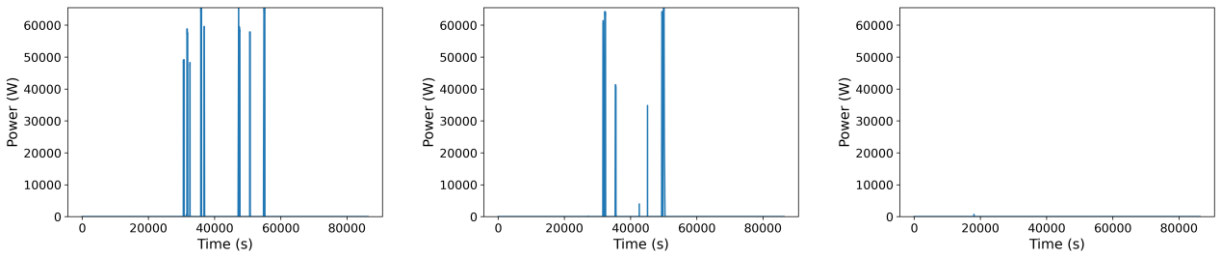

(q)




(r)




(s)




(t) (u) (v)

*Figure D1: Aggregated load profiles at $n_{th}$per day type (left: business day, center: vacation day, right: closing day) for desktop computers (a)-(c), laptops with one screen (d)-(f), computer screens (g)-(i), multifunction devices (j)-(l), individual heaters (m)-(o), coffee machines (p)-(s) and kettles (t)-(v)*

## 7.5. Appendix E: Representativity of the survey responses distribution with respect to office buildings stock in France

The inquiry was conducted during 2018-2019 with office workers to determine different building profiles based on five parameters: the company's main activity, workforce size, building location, employee status and whether or not the building was the head office. The inquiry had 1038 participants in France. The participants were asked questions based on their workplace characteristics (main activity, size, location, etc.), work hours (work schedule, break times, etc.) and equipment types (office and personal equipment).

The distribution of the work sectors in the inquiry responses were compared to that of the INSEE (National Institute of Statistics and Economic Studies)[1] to determine if the data collected by the inquiry is representative of the main activities in office buildings in France.

[1] https://www.insee.fr/fr/statistiques/2569348?sommaire=2587886

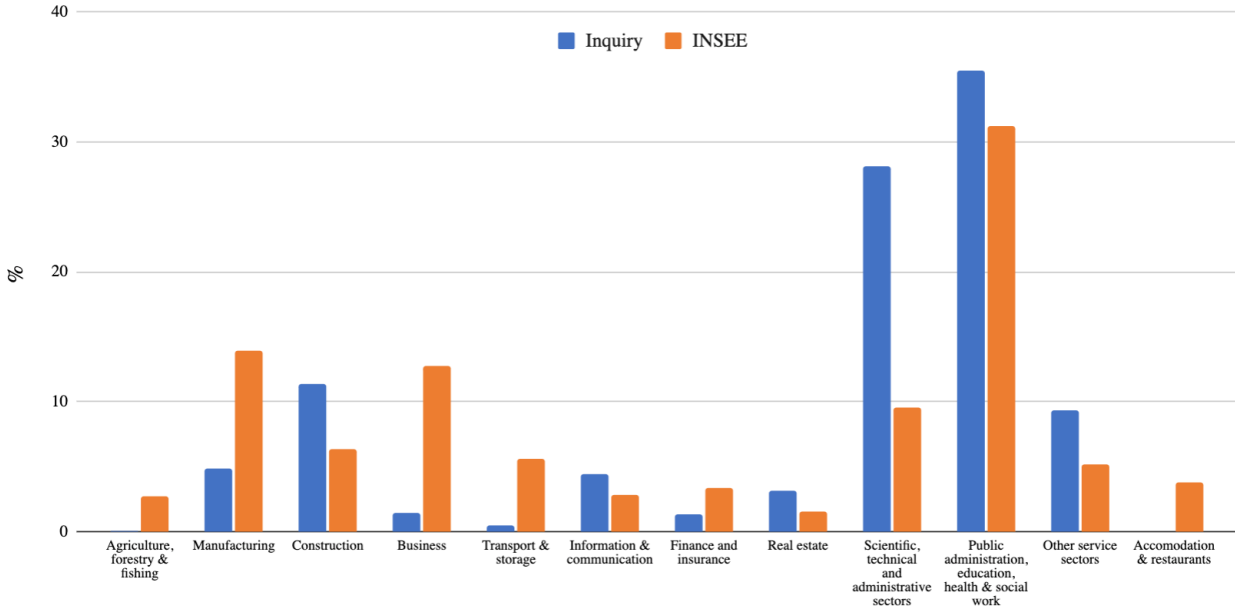



*Figure E1: Distribution of the work sectors according to the inquiry responses and the INSEE.*

At a first glance, the distributions are different. 92.1% of the present inquiry responses were mainly from people working in sectors where their main workplace is an office building. In the INSEE statistics, the same sectors represent 56.6% of the population. The inquiry mainly targeted office buildings occupants which explains the discrepancy. Since we are focusing in office buildings, the covered population is relevant for the present research and its diversity largely covers the main office buildings activities.

## 7.6. Appendix F: Experimentally measured load profiles statistics and pre-processing

In this appendix, we provide additional information on the collected data statistics such as the location, size and workforce, the number of monitored appliances, the number of measurement days, etc.

We also provide additional information on the data pre-processing step described in section 2.2. In particular, we provide the thresholds used to clean data and remove the outliers from measured load profiles.

### 7.7.1. Measured load profiles statistics

Load profile measurements were carried out on three separate buildings in France, Paris Metropolitan Area. The measured load profile statistics and the characteristics of these buildings are detailed in the following table.

| **Building** | **Location** | **Main activity** | **Number of occupants** | **Area ($m^2$)** | **Number of monitored appliances** | **Number of monitoring days** |
|---|---|---|---|---|---|---|
| B1 | Paris Metropolitan area | Education, research, public administration | 2500 | 30 000 | 45 | 236 |
| B2 | Paris Metropolitan area | Production and distribution of electricity, gas, steam and air conditioning | 44 | 1000 | 36 | 104 |
| B3 | Paris Metropolitan area | Research and development, public administration | 636 | 48000 (including exeperimental facilities) | 32 | 427 |

*Table F1 : key information on the buildings where load profiles were measured and the statistics of measured load profiles.*

### 7.7.2. Pre-processing: filtering and outliers removal

In the pre-processing stage, outliers in the load profile data were filtered out. These outliers are data points that exceed a power threshold considered as the maximum reachable load for each appliance. The thresholds are user-defined based on the observation of typical load profiles of the dataset and are provided in the following table:

| **Appliance type** | **Threshold load (W)** |
|---|---|
| Laptop | 200 |
| Desktop computer | 500 |
| Computer screen | 60 |
| Multifunction device (printer) | 2500 |
| Laptop with 1 screen | 300 |
| Coffee maker | 2000 |
| Kettle | 2500 |
| Personal heater | 3000 |

*Table F2 : threshold for outliers detection per appliance type.*