Energy & Buildings 215 (2020) 109892

Contents lists available at ScienceDirect

Energy & Buildings

journal homepage: www.elsevier.com/locate/enbuild

The good, the bad, and the ugly: Data-driven load profile discord identification in a large building portfolio

June Young Park[a,b,*], Eric Wilson[a], Andrew Parker[a], Zoltan Nagy[b]

[a] Buildings and Thermal Sciences Center, National Renewable Energy Laboratory, Golden, CO, USA
[b] Intelligent Environments Laboratory, Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, Austin, TX, USA

ARTICLE INFO

Article history:
Received 25 October 2019
Revised 13 February 2020
Accepted 19 February 2020
Available online 22 February 2020

Keywords:
Smart meter
Portfolio analysis
Load profile
Data mining
Discord detection

ABSTRACT

Reducing the overall energy consumption and associated greenhouse gas emissions in the building sector is essential for meeting our future sustainability goals. Recently, smart energy metering facilities have been deployed to enable monitoring of energy consumption data with hourly or subhourly temporal resolution. This unprecedented data collection has created various opportunities for advanced data analytics involving load profiles (e.g., building energy benchmarking programs, building-to-grid integration, and calibration of urban-scale energy models). These applications often need preprocessing steps to detect daily load profile *discords*, such as: 1) outliers due to system malfunctions (*the bad*) and 2) irregular energy consumption patterns, such as those resulting from holidays (*the ugly*) compared to normal consumption patterns (*the good*). However, current preprocessing methods predominantly focus on filtering using statistical threshold values, which fail to capture the contextual discords of daily profiles. In addition, discord detection algorithms in building research are often aimed at finding individual building-level discords, which are not suitable at a large scale. Thus, in this paper, we develop a method for automated load profile discord identification (*ALDI*) in a large portfolio of buildings (more than 100 buildings). Specifically, *ALDI* 1) uses the matrix profile (MP) method to quantify the similarities of daily subsequences in time series meter data, 2) compares daily MP values with typical-day MP distributions using the Kolmogorov-Smirnov test, and 3) identifies daily load profile discords in a large building portfolio. We evaluate *ALDI* using the metering data of both an academic campus and a residential neighborhood. Our results demonstrate that *ALDI* efficiently discovers measurement errors by system malfunctions and low energy consumption days in the academic campus portfolio, and it detects unique load shape patterns likely driven by occupant behavior and extreme weather conditions in the residential neighborhood.

1. Introduction

The building sector contributes to more than 40% of energy consumption and 38% of greenhouse gas emissions in the United States [1]. With rapid urbanization, the associated energy consumption and environmental emissions in buildings are projected to increase at unprecedented levels globally [2]. Therefore, it is highly important to apply energy-efficient and sustainable strategies to the building stock.

The development of sensing and computing systems has created promising opportunities in the built environment in particular [3]. One example is advanced metering infrastructure, which measures and stores electricity energy consumption data in hourly or subhourly resolution for a building. In the United States, more than 70 million smart electricity meters were installed as of 2016 [4]; this provides the potential for utilities, customers, and researchers to have a better understanding of how energy is consumed in buildings and advanced applications (e.g., portfolio analysis, discord detection, load profiling, customer classification, and intelligent energy management systems).

1.1. Smart meter data analytics for building portfolios

The main objective of this paper is analyzing a large building portfolio, which typically consists of more than 100 buildings. If these buildings are located in the same geographical region and/or connected to the same electrical grid, an owner of a portfolio can analyze and develop a strategy for improving energy performance for the property as a whole [5].

An example portfolio application is extracting meaningful knowledge of building energy performance (e.g., daily load shape)

* Corresponding author at: Intelligent Environments Laboratory, Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, Austin, TX, USA.

E-mail address: juneyoungpark@utexas.edu (J.Y. Park).

https://doi.org/10.1016/j.enbuild.2020.109892
0378-7788/

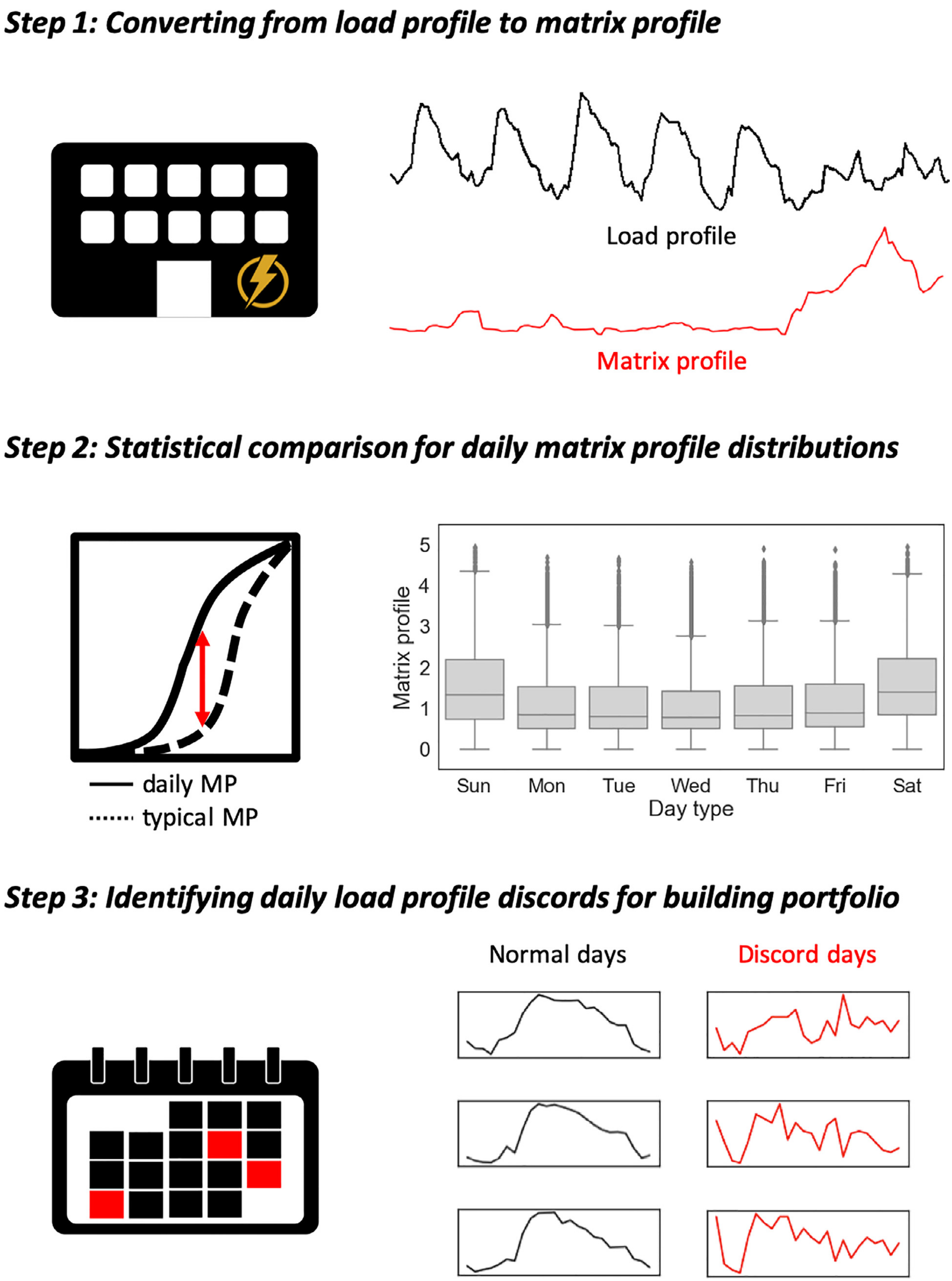

Fig. 1. Overview of *ALDI* method.

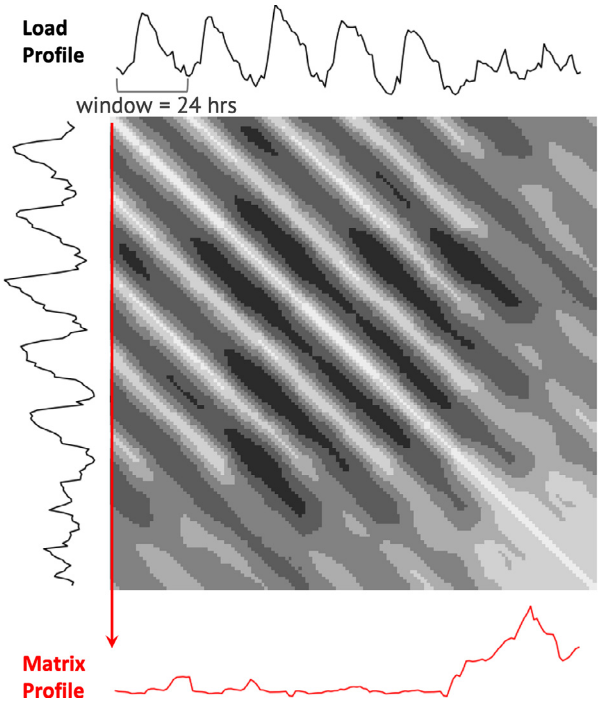

Fig. 2. This figure shows the procedure used to calculate matrix profile for load profile: (1) Construct a z-normalized Euclidean distance matrix of all pairs of subsequences of time window, 24 h (grey scale heatmap: the z-normalized Euclidean distance), (2) Calculate the distance of the nearest neighbor under z-normalized Euclidean distance (red line plot: matrix profile). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

from a large building portfolio using various clustering techniques. Such load profile information has been used for multiple applications: customer classification [6], portfolio analysis [7,8], and building energy benchmarking [9,10]. For instance, Park et al. discovered three fundamental load shapes from a large and diverse smart metering data set (3,829 buildings); then, they grouped buildings according to their dominant load shape profiles [10]. Their benchmarking study (grouping buildings with similar energy performance) confirmed that buildings can be classified by their energetic behaviors rather than solely by static building classification information (e.g., usage type, floor area, location).

Smart metering data also can be utilized for community-level studies [11–14]. Jain et al. analyzed hourly resolution energy consumption data to develop a distributed energy sources management plan [11]. Other researchers extracted daily load profile patterns from metering data and demonstrated the integration of renewable energy sources with individual building loads [12,13]. Nutkiewicz et al. used smart metering data for urban energy modeling and calibrated their simulation model with the metering data [14].

These studies demonstrate the importance of smart metering data analytics in various aspects of buildings research. Specifically, smart metering infrastructure allows researchers to monitor the temporal aspects of building performance and extract fine-grained information. For this reason, the quality of smart metering data has become an important issue, because erroneous meter readings and abnormal daily patterns (e.g., holidays) can distort aggregated data and reduce its accuracy and effectiveness for various applications [15]. Therefore, ideal preprocessing should: 1) remove outliers (i.e., measurement error by system malfunctions), and 2) differentiate unique daily load shapes (i.e., by occupant behavior, dif-

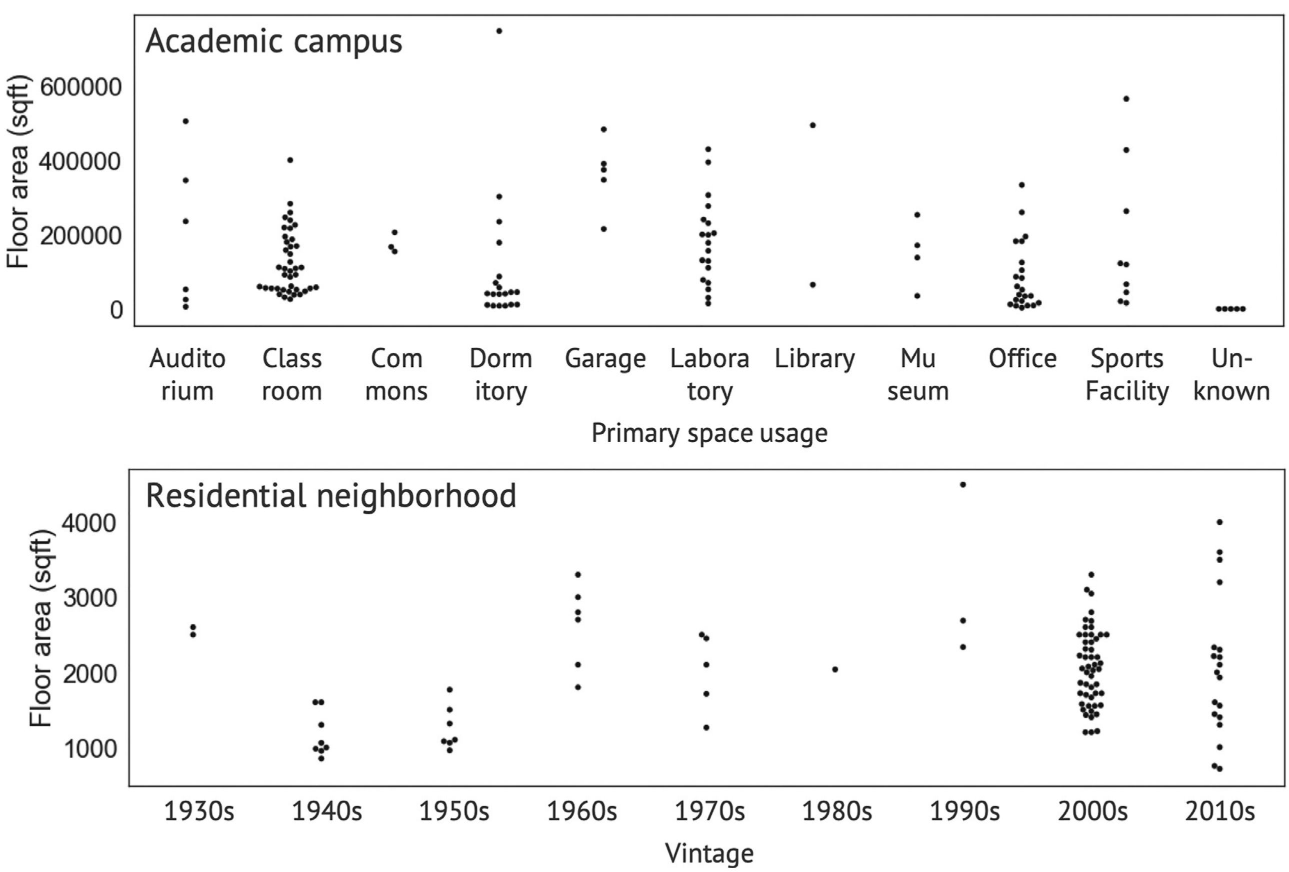

Fig. 3. Metadata summary for university campus (top) and residential neighborhood (bottom) building portfolios.

ferent operation schedules) to characterize daily load shapes and utilize them accurately.

1.2. Outlier and discord detection in smart meter data

As a conventional approach, researchers use visual inspection on their time series load profiles to identify outliers. The most intuitive way of detecting outliers is calculating the 25% (Q1) and 75% quartiles (Q3) and the interquartile range (IQR), and then eliminating high ($\geq Q3 + 1.5 \times IQR$) and low measurements ($\leq Q1 - 1.5 \times IQR$). There are numerous filtering techniques for identifying outliers in other domains (e.g., economy, finance, mathematics, statistics) with various definitions of thresholds [16], such as using the fixed threshold filter [17], the standard deviation filter [18], the recursive filter [19], and the moving threshold filter [20].

Adopting such filtering methods is inappropriate for building portfolio analysis, however, because these methods predominantly focused on calculating statistical outliers (extreme values) rather than the energy consumption patterns of each day. Considering that various applications require not only filtering extreme values but also contextual meanings (e.g., different operation schedule, capturing unique occupant behavior) of daily load profiles, it is necessary to have a framework to investigate the contextual backgrounds of a building portfolio.

Discord (and motif) detection can be a viable approach to providing such contextual information. With time series data, there are both frequently and rarely occurring patterns, and we consider the former to be motifs and the latter to be discords [5]. In the building energy context, a motif load profile means a dominant energy consumption pattern, whereas a discord load profile indicates either measurement error from the system malfunctioning or an uncommon daily energy consumption pattern. This uncommon and/or faulty pattern may be caused by internal (occupant behavior) and/or external factors (weather conditions).

Various discord detection algorithms have been developed by the computer science community to achieve faster and more accurate anomaly detection [21]. Smart metering infrastructure has been deployed predominantly during the last decade, and this recent trend led to the development and implementation of various discord detection techniques in building energy research. Researchers have used various clustering techniques to detect motif signals for energy consumption data [22,23]. Similarly, autoencoder-based methods have been developed to identify

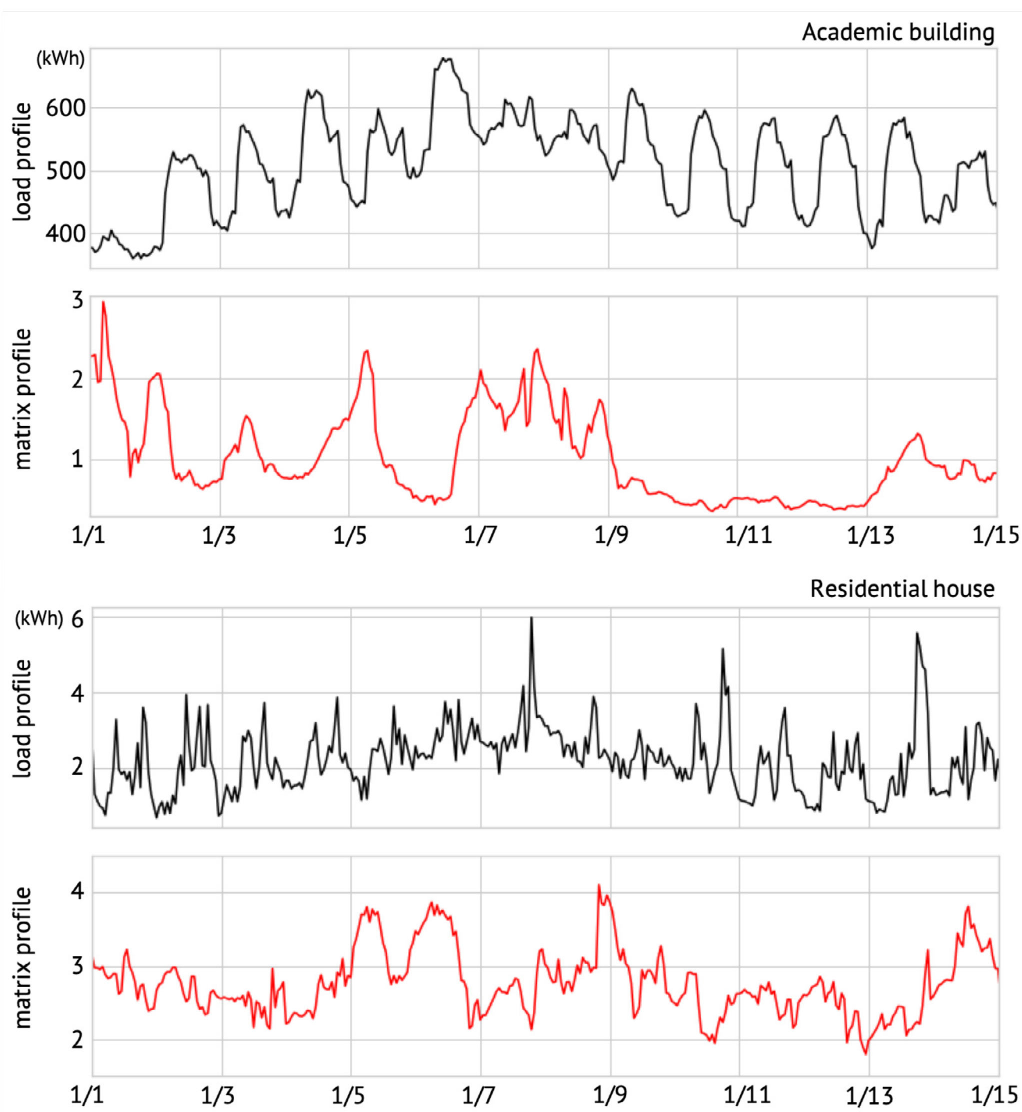

Fig. 4. Load profile and matrix profile for the first two weeks of 2017 (top: academic building, bottom: residential house).

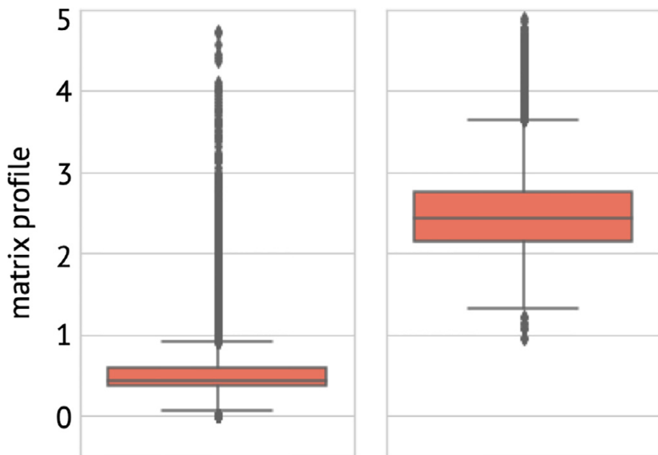

Fig. 5. Distribution of matrix profile values for one year, 2017 (left: an academic building; right: a residential house).

the discord patterns in time series energy consumption data in buildings [24,25]. Miller et al. adopted a Symbolic ApproXimation (SAX) [26] method to identify discord days [27]; researchers added adaptive features to the SAX method to find discord days according to input building energy consumption patterns [28,29].

Although these studies successfully identified the motif and discord profiles, they were mostly looking at individual building-level discords, and only 1–2 buildings were evaluated in each study. Considering that smart grid applications typically manage more than 100 buildings [30], there is significant need for a fast and automated discord detection solution for large building portfolios. In addition, the previous techniques require parameters to be tuned, which affects the performance of the discord detection. For example, clustering time series data sometimes generates arbitrary profiles due to inappropriate parameter selection [31]; autoencoder-based methods require even more complex parameters in neural network architecture. The performance of the SAX-based method is highly dependent upon the segmentation of y-axis values (i.e., the selections of letter size) and the division of x-axis time scale (i.e., the numbers of letters) [27].

1.3. Contribution

In this paper, we introduce a data-driven framework, *ALDI*, to automatically identify daily load profile discords for a large building portfolio. With rapid urbanization and increased size of building portfolios, it is necessary to detect daily discord profiles efficiently. *ALDI* calculates matrix profile (MP) values, which is a relatively simple and fast data mining technique to search for similar pairs within a large amount of time series data [32]. Moreover, *ALDI* is a parameter-light approach; the user only needs to specify the confidence level (i.e., p-value) of discord detection. As a case study, we tested *ALDI* with actual building energy meter data to identify discord days in large building portfolios.

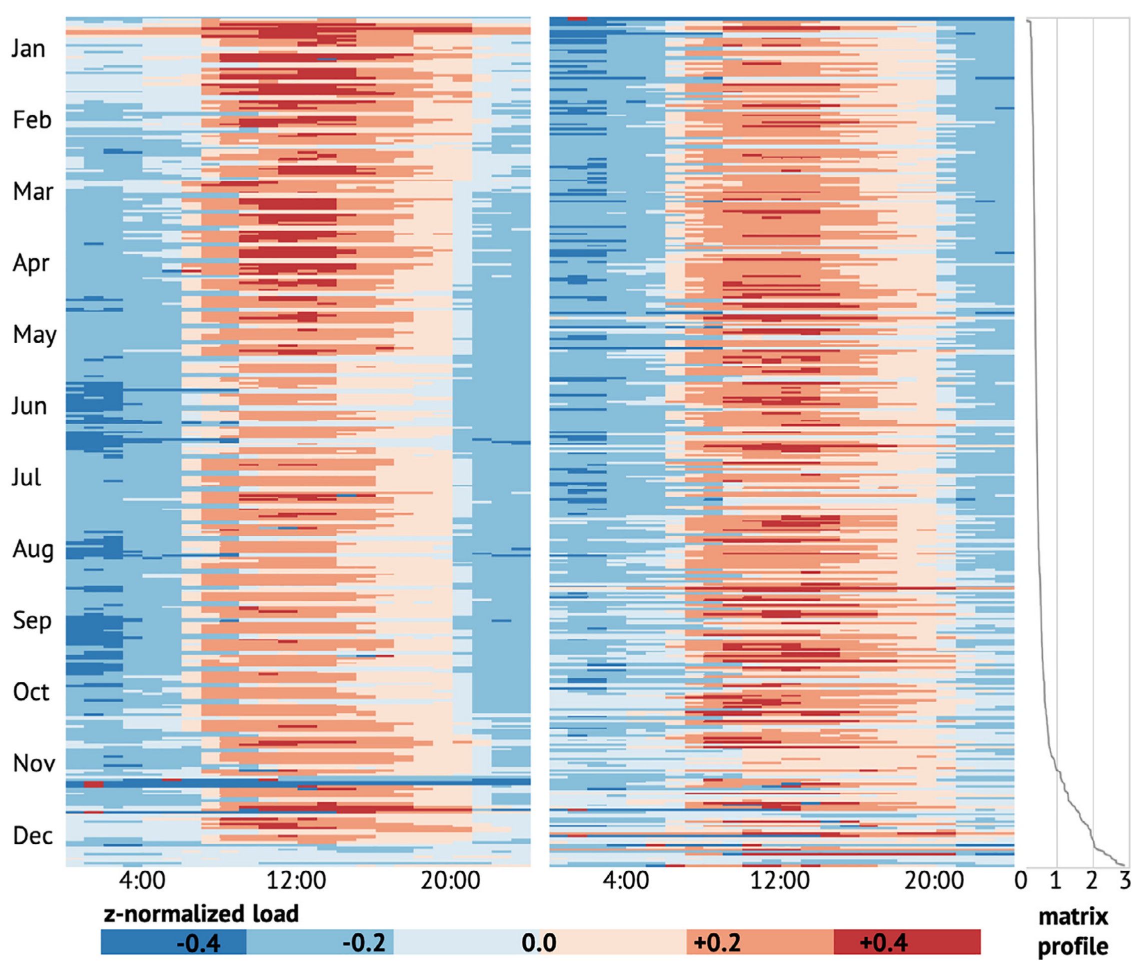

Fig. 6. Comparison between z-normalized daily load profiles ordered by date and matrix profile for example academic building (left: from 1/1/2017 to 12/31/2017; right: from low to high matrix profile values).

1.4. Organization

The remainder of this paper is structured as follows. Section 2 explains *ALDI*, which automatically identifies discord days in a building portfolio and our test data set (academic campus and residential neighborhood). Section 3 reports the identified discords in two building portfolios. Section 4 discusses the implications, potential applications, and limitations of our approach. Finally, Section 5 concludes the paper.

2. Methodology

2.1. ALDI: *Automated load profile discord identification*

As shown in Fig. 1, using *ALDI* consists of three steps: 1) *ALDI* reads the load profiles (hourly electrical energy consumption time series data) and converts the input data into the matrix profile (MP) to discover similar daily load shapes. MP calculation is a novel data mining technique used to search for similar pairs within time series data [32]. 2) Applying the previous calculation to multiple buildings, *ALDI* calculates the daily MP values and groups them by the typical day types (Monday–Sunday). Then, it conducts statistical tests to compare how individual days' MP distributions are similar (or dissimilar) against the typical days' MP distributions using the Kolmogorov-Smirnov (KS) test [33]. 3) Last, we qualitatively evaluate the hypothesis test results to identify the discord types for each portfolio.

2.2. Matrix profile calculation for building load profile

For the first step, *ALDI* requires at least 1 year of time series data. This is because MP calculation compares diurnal patterns of load profiles and it can capture seasonal variation of the load shape patterns. Also, given that metering data will continue to accumulate in the future, 1 year worth of data should be easily accessible. The desired input resolution is hourly (or subhourly), which has some temporal patterns throughout a year. It should be noted that *ALDI* can accept raw metering data without preprocessing. This is

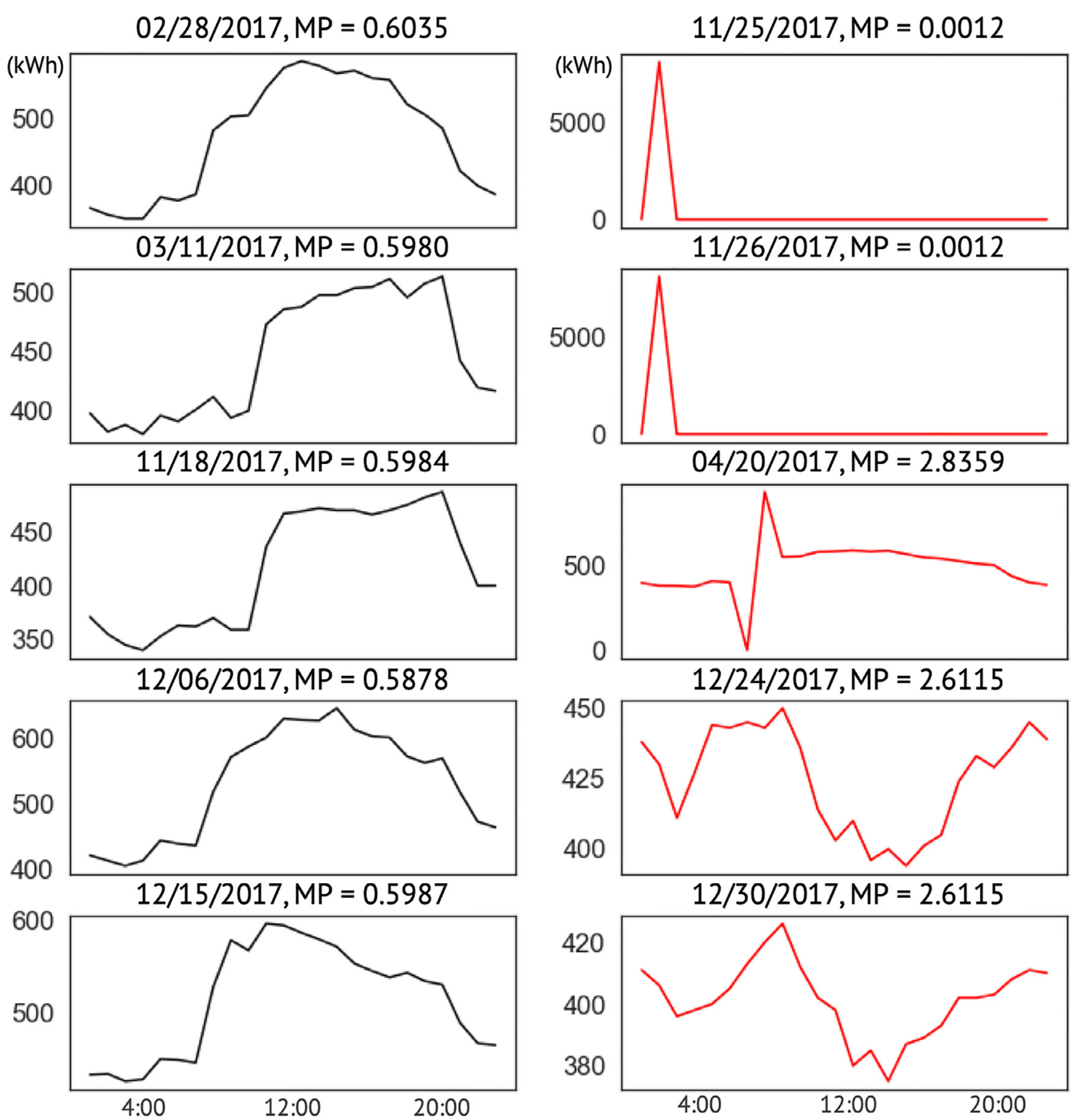

Fig. 7. Load profiles with different matrix profiles for example academic building (left [black]: medium matrix profiles; right [red]: extremely low and high matrix profiles; each subplot has a different *y*-axis scale for visualization purposes). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

because it automatically identifies such measurement error as one of the discords.

After reading the original time series load profile inputs, *ALDI* calculates MP values. This calculation method was originally developed by [32]. To illustrate the MP method, we sample electricity consumption data (1 week, Monday–Sunday) from a random building (Fig. 2).

First, *ALDI* makes an identical copy of the load profile. Then, it calculates the z-normalized Euclidean distances (ZED) of the input profile with the subsequence length, m, of all pairs. Because our objective is investigating daily load shape pattern, we set $m = 24$. We calculate the ZED of two time series ($T_{i,m}$ and $T_{j,m}$) as,

$$\mathrm{ZED}_{i,j} = \sqrt{2m\left(1 - \frac{QT_{i,j} - m\mu_i\mu_j}{m\sigma_i\sigma_j}\right)} \tag{1}$$

where μ_i and σ_i are the mean and standard deviation of $T_{i,m}$, μ_j and σ_j are the mean and standard deviation of $T_{j,m}$, and $QT_{i,j}$ is the dot product of $T_{i,m}$ and $T_{j,m}$ [32]. Essentially, z-normalization lets us focus on the shape of the profile rather than the magnitude, as the resulting mean and standard deviation would be closed to 0 and 1, respectively [34].

To acquire the MP values M_i (red line plot in Fig. 2) of a given time series, *ALDI* takes the minimum values of the ZEDs along the horizontal lines of the heatmap in Fig. 2, i.e.,

$$\mathrm{M}_i = \{\min(\mathrm{ZED}_{i,j}) \mid j \in [0, 168[\} \tag{2}$$

where $i \in [0, 168[$ is the hour of the input data. In brief, the MP calculation finds the ZEDs of the nearest neighborhood, i.e., the distance of the most similar time series pattern.

As shown in the heatmap in Fig. 2, the diagonal line indicates 0 distances, which is clear because the ZED of the two identical profiles is always 0. There are also relatively short distances (brighter colors) for the next diagonal lines, i.e., the starting points (midnight) of the rest of the weekdays. However, the starting points of the weekends have relatively long distances (darker color). The MP calculation result (red line plot in Fig. 2) clearly indicates the observed pattern. For example, we found relatively low MP values until the end of Friday, which means the shape of daily load profile is rather repetitive during weekdays. Then, it increases on Saturday and shows the maximum value on Sunday. This means that the shapes of the weekend profiles are different in general and the Sunday profile is the most unique of the input profiles.

Because we are comparing daily profiles, we previously chose m as 24 data points, and the first value of the daily load profile must start at midnight. Therefore, we select the midnight MP values as representative values on each day for a particular building (*bldg*). With such settings, we can compare the similarities among daily

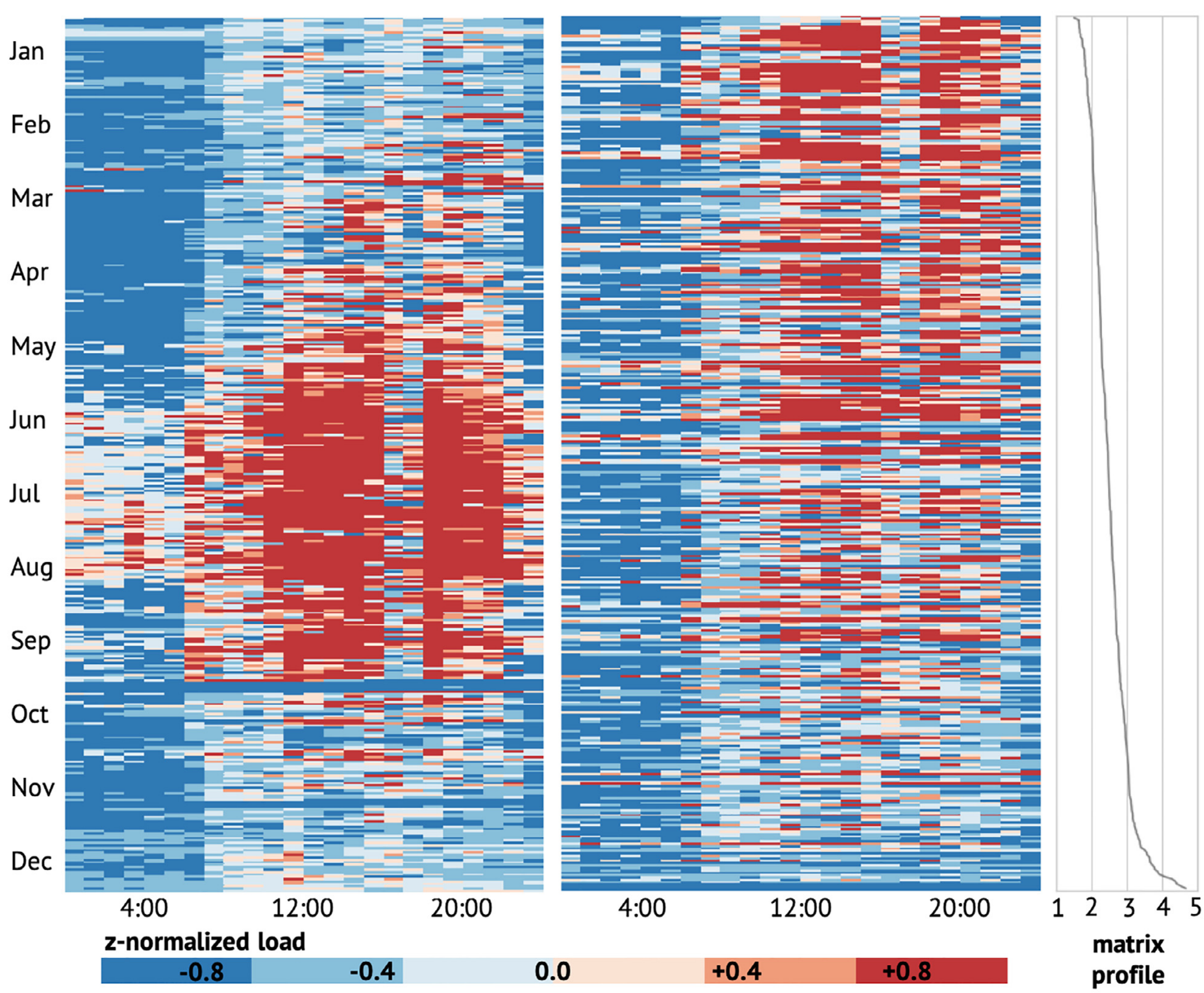

Fig. 8. Comparison between z-normalized daily load profiles ordered by date and matrix profile for example residential building (left: from 1/1/2017 to 12/31/2017; right: from low to high matrix profile values).

profiles. Let $h_{midnight}$ be the set of hours at midnight for each day; we then define $\mathrm{MP}_{bldg}(d)$ as,

$$\mathrm{MP}_{bldg}(d) = \{\mathrm{M}_i | i \in h_{midnight}\} \tag{3}$$

2.3. Statistical comparison of daily matrix profile distributions

To identify the portfolio-level discords, we collect the MP values for all buildings in a particular portfolio (prt). Now, $\mathrm{MP}_{prt}(d)$ is described as,

$$\mathrm{MP}_{prt}(d) = \{\mathrm{MP}_{bldg}(d) | bldg \in prt\} \tag{4}$$

then, we divide $\mathrm{MP}_{prt}(d)$ according to typical day types. Let $date_{Mon}$–$date_{Sun}$ be the set of dates of Monday–Sunday, respectively. Then the MP collections for the seven unique typical days are represented as,

$$\begin{cases} MP_{prt}(Monday) = \{MP_{prt}(d) | d \in date_{Mon}\} \\ \ldots \\ MP_{prt}(Sunday) = \{MP_{prt}(d) | d \in date_{Sun}\} \end{cases} \tag{5}$$

The next step is determining discord days based on the hypothesis test result of the KS test [33]. This statistical test evaluates the equality of the two probability distributions. It calculates a distance, D, between cumulative distribution functions of two samples and quantifies the statistical significance (p-value) of such distance.

In detail, we first convert $\mathrm{MP}_{prt}(d)$ to empirical cumulative distribution functions ($F_n(x)$), i.e.,

$$F_n(x) = \frac{1}{n}[\text{Number of observations} \le x] \tag{6}$$

where x is the daily MP values of an input portfolio, and n is the number of the daily MP values for each set. Then, we calculate the largest vertical distance, D, of the two empirical cumulative distribution functions, i.e.,

$$D = \max_{1 \le i \le n}\left(F(x_i) - \frac{i-1}{n}, \frac{i}{n} - F(x_i)\right) \tag{7}$$

Subsequently, we conduct a hypothesis test based on our observation of distances, D. In the KS test, the null (H_0) and the alternative

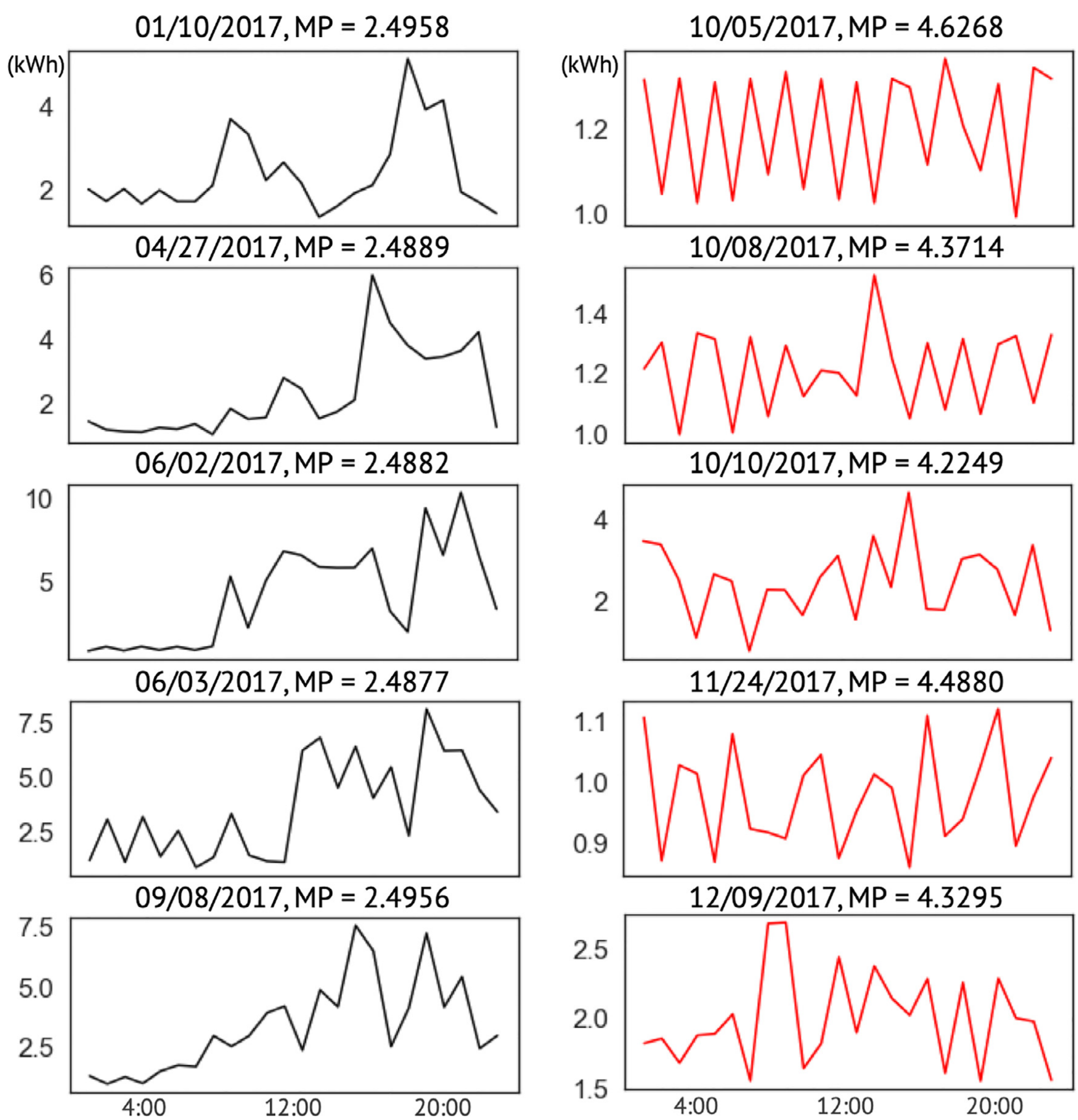

Fig. 9. Load profiles with different matrix profiles for example residential building (left [black]: medium matrix profiles; right [red]: high matrix profiles; each subplot has a different y-axis scale for visualization purposes). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

hypotheses (H_A) are,

$$\begin{cases} H_0 : \text{two input data samples follow the same distribution} \\ H_A : \text{two input data samples do not follow the same} \\ \quad\;\; \text{distribution} \end{cases} \tag{8}$$

To test the hypotheses, we decide the confidence level using p-value. This value is the significance level for accepting the null hypothesis (H_0) for all values less than the assigned p-value.

We implement this statistical tool for discord operation day detection. Our main assumption is that if the MP distributions of the portfolio on a particular day are distinct from the ones of the typical day, then we consider that day a discord day. Because the daily MP values indicate the distance of the nearest load profile, it is fair to use the MP distributions to evaluate the abnormality of load profiles. Also, ZED calculation normalizes the differences of magnitude in energy consumption. Therefore, we can collect information from all buildings to calculate the distributions of portfolio. We compare $\mathrm{MP}_{prt}(d)$ and $\mathrm{MP}_{prt}(Monday)$–$\mathrm{MP}_{prt}(Sunday)$ in our KS test. For example, as the day type of January 1, 2017, ($d = 1$) is Sunday, we conduct the KS test for $MP_{prtf}(1)$ and $MP_{prtf}(Sunday)$. This comparison was evaluated for every day in the year ($d \in [1, 365]$).

After the KS test, we calculate the sets of both distances, D, and calculated probabilities (p-value) for every day, i.e.,

$$\begin{cases} D = \{\text{distances of input day}(d)|d \in [1, 365]\} \\ p = \{p\text{-values of input day}(d)|d \in [1, 365]\} \end{cases} \tag{9}$$

these sets are also used for characterizing daily load profile types.

2.4. Identifying typical and discord daily load profiles

With a desired significance level (p-value), we can further group each day by normal and discord operation days. Specifically, if the input day's p-value is less than the threshold, then such day is assigned as $Day_{discord}$; otherwise it is part of Day_{normal}. It is important to note that our framework can provide an insight for facility managers to determine the significance level depending upon their objectives and portfolios, rather than fixing a certain p-value for discord identification. For example, if a facility manager wants to identify extreme load profiles (e.g., meter malfunctioning and/or fault), then they can assign a very small p-value. To describe the importance of the significance level in portfolio management, we report the numbers of discord days by different p-values.

With an example p-value, we group each day into Day_{normal} and $Day_{discord}$. Both day types are mapped into the calendar view and the discord days are qualitatively analyzed. In addition, we identify the typical load shape pattern of each portfolio from the load profiles of Day_{normal}. Ultimately, the discord load profiles are visualized and compared with the identified normal load profiles in order to infer the error types.

2.5. Case study data source and summary

To evaluate the proposed framework, we selected two metering data sets. Table 1 summarizes our test data sources. The data format is saved as a comma-separated values (CSV) file, and it contains hourly electricity consumption data for 2017. Both portfolios have more than 100 buildings and are located in Austin, Texas

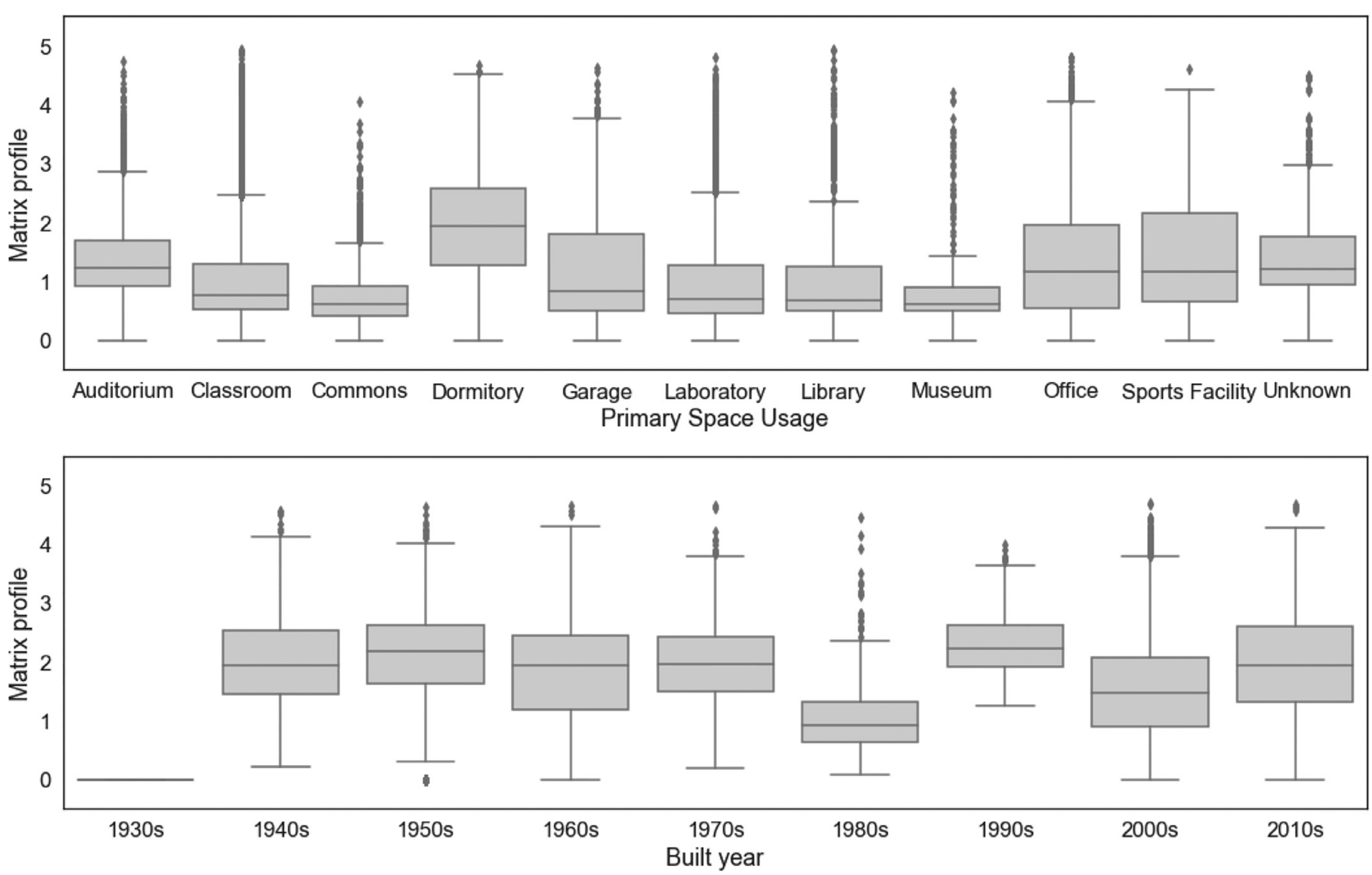

Fig. 10. Matrix profile distributions by meta information (top: academic campus portfolio by building primary space usages; bottom: residential neighborhood portfolio by built years).

Table 1
Summary of data sources.

Data Set	Location	Nr. of Bldgs	Type	Date Range	Ref.
UT-Austin Energy Portal	Austin, TX	128	Various types; academic campus	01.01 - 12.31.2017	[35]
Pecan Street Dataport	Austin, TX	102	Residential houses	01.01 - 12.31.2017	[36]

(30.2672° N, 97.7431° W). We intentionally selected our test data sets from both an academic campus and a residential neighborhood to evaluate the applicability for various building types.

Fig. 3 shows the distributions of the floor area by primary space usages (PSU) and built-year information, respectively. The academic campus portfolio has 11 different PSUs, and the major ones are classroom (40 buildings), office (22 buildings), laboratory (19 buildings), and dormitory (19 buildings). The average floor area of the buildings in this portfolio is 138,398 ft^2. All 128 buildings are located in the downtown campus of the University of Texas at Austin, and they are connected to a single electrical grid, which is operated by a facility management department. Residential neighborhood portfolio consists of only one building type (single-family homes) with photovoltaic equipment, and they are mainly constructed after the year 2000 (70 buildings). The average floor area of the houses in this portfolio is 2,010 ft^2, and the houses are all located in the same neighborhood (Mueller district, Austin).

To summarize, this section described the proposed data analytic framework, *ALDI* (Fig. 1). In brief, *ALDI* calculates MP values to quantify the similarities among daily load profiles and statistically compares the daily MP distributions with typical days to find load profile discords. Then, users can qualitatively evaluate the building portfolio with respect to the characteristics of daily load shape patterns. We also introduced the two building portfolios (academic campus and residential neighborhood) in this section. The following section presents the results of applying *ALDI* to the two portfolios.

3. Results

First, we randomly selected a building from each portfolio to calculate the MP values. With this result, we explain the implications of MP in building energy contexts. After the single building analysis, we further apply the proposed framework to a broader scale (building portfolio) as our main result. Using the KS test, the proposed framework identifies discord days in terms of load shape pattern. We qualitatively investigate the load profiles in order to infer the potential reasons of discords.

3.1. Single building analysis with matrix profile

Fig. 4 indicates the load profiles and MP values of two randomly selected buildings from the academic campus and residential neighborhood. In the academic building, the weekday load profiles are generally repetitive compared to the weekend ones, which results in low weekday and high weekend MP values. However, the shape of the load profile on January 5, (Thursday) is dissimilar to other weekdays. Consequently, this generates a small peak on January 5, in the MP plot. Looking at the residential house, it is relatively hard to capture clear patterns. This is because the original load profile has no clear patterns, which is related to the difference between schedule-driven (academic building) and enclosure-driven buildings (residential house). Nonetheless, the MP values indicate similarities between other daily load profiles. For example,

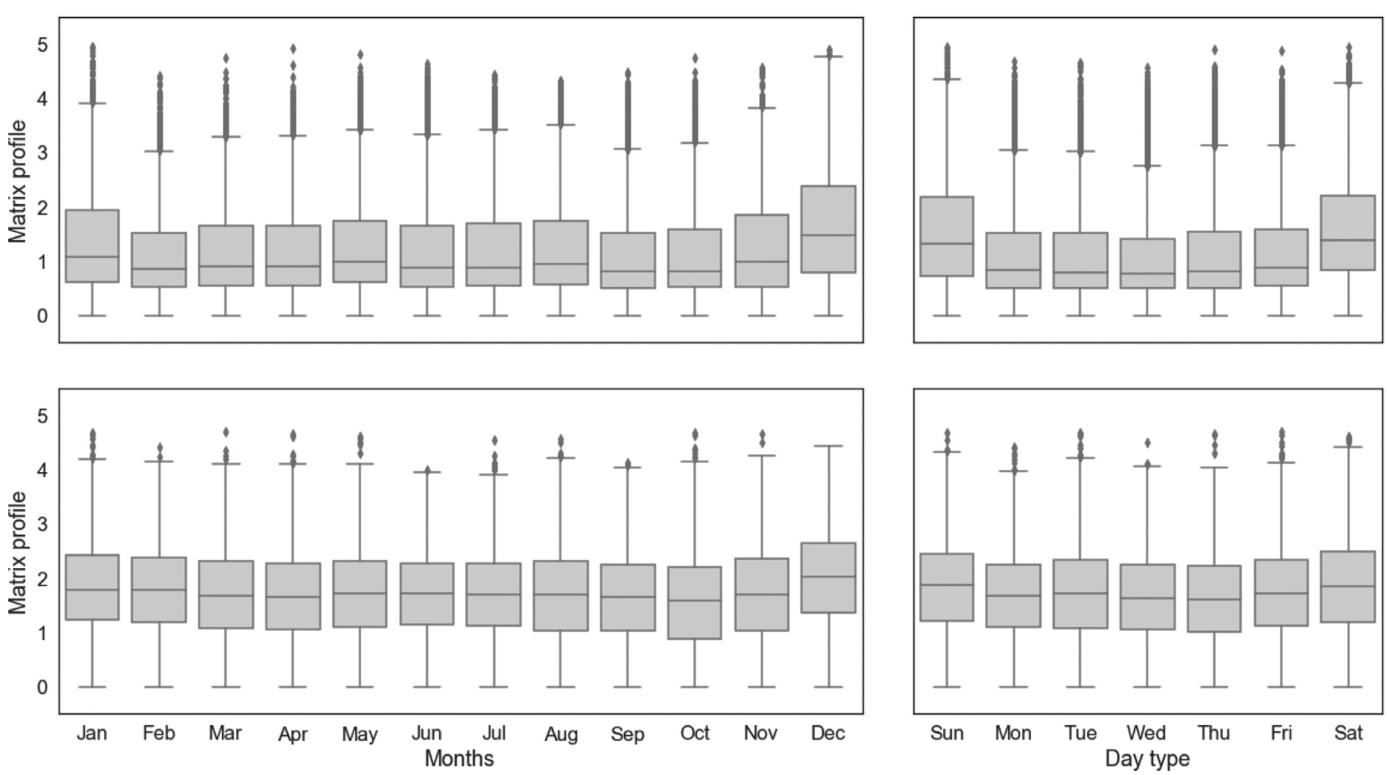

Fig. 11. Matrix profile distributions by month and day (top: academic campus portfolio; bottom: residential neighborhood portfolio).

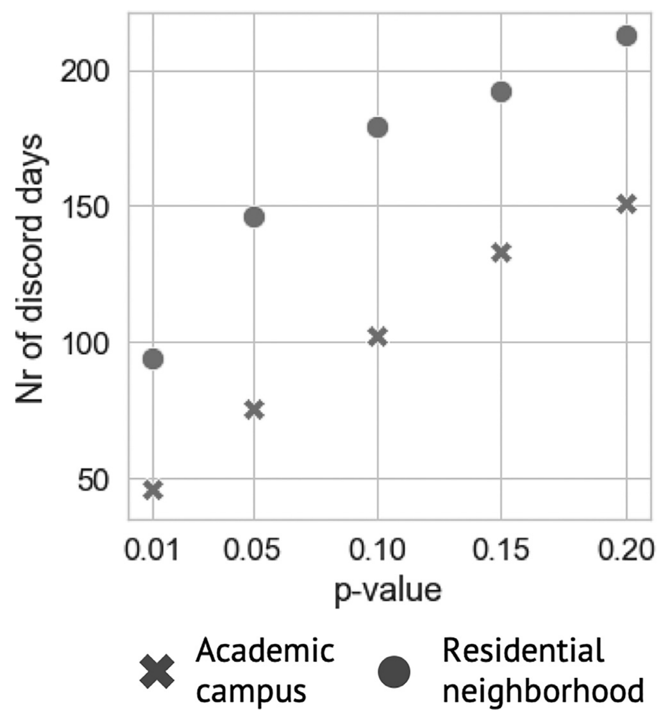

Fig. 12. Numbers of discord days by different significance levels.

the daily load profiles of January 5, 6, and 8 are unique shapes compared to other days.

Fig. 5 compares the MP distributions of the two buildings. Clearly, the IQR location of the academic building is lower than the residential house. This observation confirms our previous result—daily load profiles in the academic building are more similar compared to the residential house. However, it should be noted that we found a number of statistical outliers ($\geq Q3 + 1.5 \times IQR$) because this building has extremely large metering values (e.g., $\pm$ 50000 kWh) as metering malfunctions. In addition, there are a few points with 0 as MP values, which indicates that such days are identical profiles (by the definition of the MP calculation). We will revisit and discuss this error type in the next subsection.

As stated in Section 2, we further assign the daily MP values (midnight values) for each day. Figs. 6 and 8 explain how the daily MP value orders load profiles. For example, individual rows in the heatmap visualize the z-normalized daily load profiles. The left heatmap indicates load profiles with a chronological order (top: January 1; bottom: December 31). Alternatively, the right heatmap orders the same load profiles by the MP values (top: minimum MP; bottom: maximum MP). In addition, we selected 10 load profiles of the high and low MP values to visualize and investigate the actual load profiles (Figs. 7 and 9).

We find relatively high peak loads for spring and fall in Fig. 6. Because the selection was from the academic campus portfolio, this observation makes sense according to the academic calendar. There were also clear distinct load shapes in late November (e.g., high peak at 1 a.m. and steady values) and December (e.g., low values for a day). These days might be potential discord days in further steps. The right heatmap clearly distinguishes these days by the MP values. The majority of the daily MP values are less than 1 (upper portion), and they are typical load shape patterns of this building (left subplots in Fig. 7). Notably, the two minimum MP values are 0, which confirms that these are the exact same load profiles. In fact, these were malfunctions during Thanksgiving break (two right upper subplots in Fig. 7). The bottom portion of the right heatmap (higher than 1 in the daily MP values) indicates that the load profiles of the high MP values have no similar pairs (three right lower subplots in Fig. 7).

Fig. 8 shows different energy consumption patterns in a residential house. Clearly, we find intensive energy consumption patterns during summer season. This is because there is no predefined operation schedule in this residential neighborhood; rather, occupant behavior governs the building operation, which is highly dependant on outdoor weather conditions. Nevertheless, the MP values order each daily load profile by its similarities among other daily profiles. The lowest daily MP value is 1.7, and we found most daily MP values to be less than 3 (right heatmap in Fig. 8), which indicates that individual load profiles are rather unique (left subplots in Fig. 9). On the other hand, the high MP value load profiles have small spikes, and it is difficult to find similar patterns in this building (right subplots in Fig. 9).

3.2. Statistical test for discord detection

We now present our main result using the two building portfolios. First, we explore the distributions of the MP values by meta (e.g., PSU, vintage) and date information (e.g., month, day, typical day type). Next, we conduct a KS test to identify rejected and non-rejected days by comparing the MP distributions.

Fig. 10 indicates the distributions of the MP values by different building information (i.e., PSU type and built year information for the academic campus and residential neighborhood portfolio). The MP values of the dormitory buildings are high in general. This is because the energy consumption patterns of dormitory buildings are rather unique and are based on occupant behavior of energy use; other PSU-type buildings are controlled by the predefined operation schedule under campus facility management. The median MP values of the commons, garage, laboratory, library, and museum are relatively lower than ones of the auditorium, office, and sports facility because the latter PSUs have dynamic schedules. In addition, the minimum MP values are all 0, and there are extremely large outliers ($\geq Q3 + 1.5 \times IQR$) for all PSU-type buildings, which are related to the meter malfunctions. Finally, it is difficult to observe the relationship between the MP values and built year information.

Next, we investigate the daily MP distribution with month and day type (Fig. 11). Full details with date information are visualized in Figs. 18 and 19 in the Appendix. For the academic campus portfolio, we found relatively high daily MP values during winter (January, November, and December) and weekends. These high daily MP values suggest that there are more discord days in the winter season because of long holidays, and the operation schedule for the academic campus is rather flexible during weekends. In the residential neighborhood, we also found relatively high MP values in December, but the distributions of weekends are similar to ones of weekdays. We also used the distributions of typical days to differentiate discord days.

Again, the main assumption of discord detection is that if the daily MP distributions of a portfolio on a certain day are different from the distributions of typical day type, then that day is a candidate for a discord operation. The KS test quantifies this comparison by distance, D, and p-value. In the Appendix, Figs. 20 and 21 visualize the full detailed results of the KS test (D, p-value) with a calendar view. In the next section, we explain this result with respect to discord daily load profiles.

3.3. Qualitative investigation on load profiles

3.3.1. Number of discords by p-value

As stated in Section 2, the significance level (p-value) should be decided by facility managers based on their perspectives.

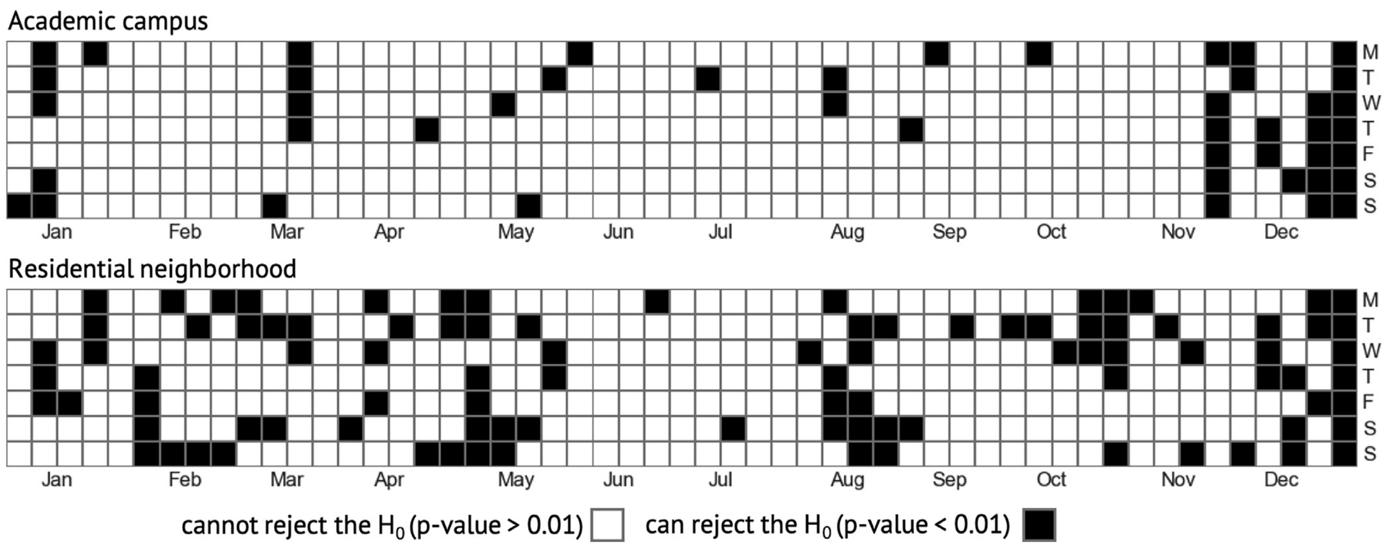

Fig. 13. Hypothesis test results with p-value of 0.01.

Fig. 12 indicates the number of detected discord days by different p-values. We found that there are more discord days with a larger p-value. Also, we found more discord days in the residential neighborhood portfolio. This suggests that the load profiles of residential buildings vary by occupant behaviors and weather conditions, whereas the academic buildings are rather consistent with the predefined campus schedules.

To demonstrate further, we set the significance level (p-value) at 0.01, meaning that if the associated p-value of input day is less than 0.01, we assign those days as the rejected days ($Day_{discord}$); otherwise, they are grouped as nonrejected days (Day_{normal}). With the p-value of 0.01, there are 46 and 94 discord days within the academic campus and residential neighborhood portfolio, respectively.

3.3.2. Discord and normal operation days

According to Fig. 13, most of the discords days are during the winter season (e.g., New Year's, Thanksgiving, and Christmas holi-

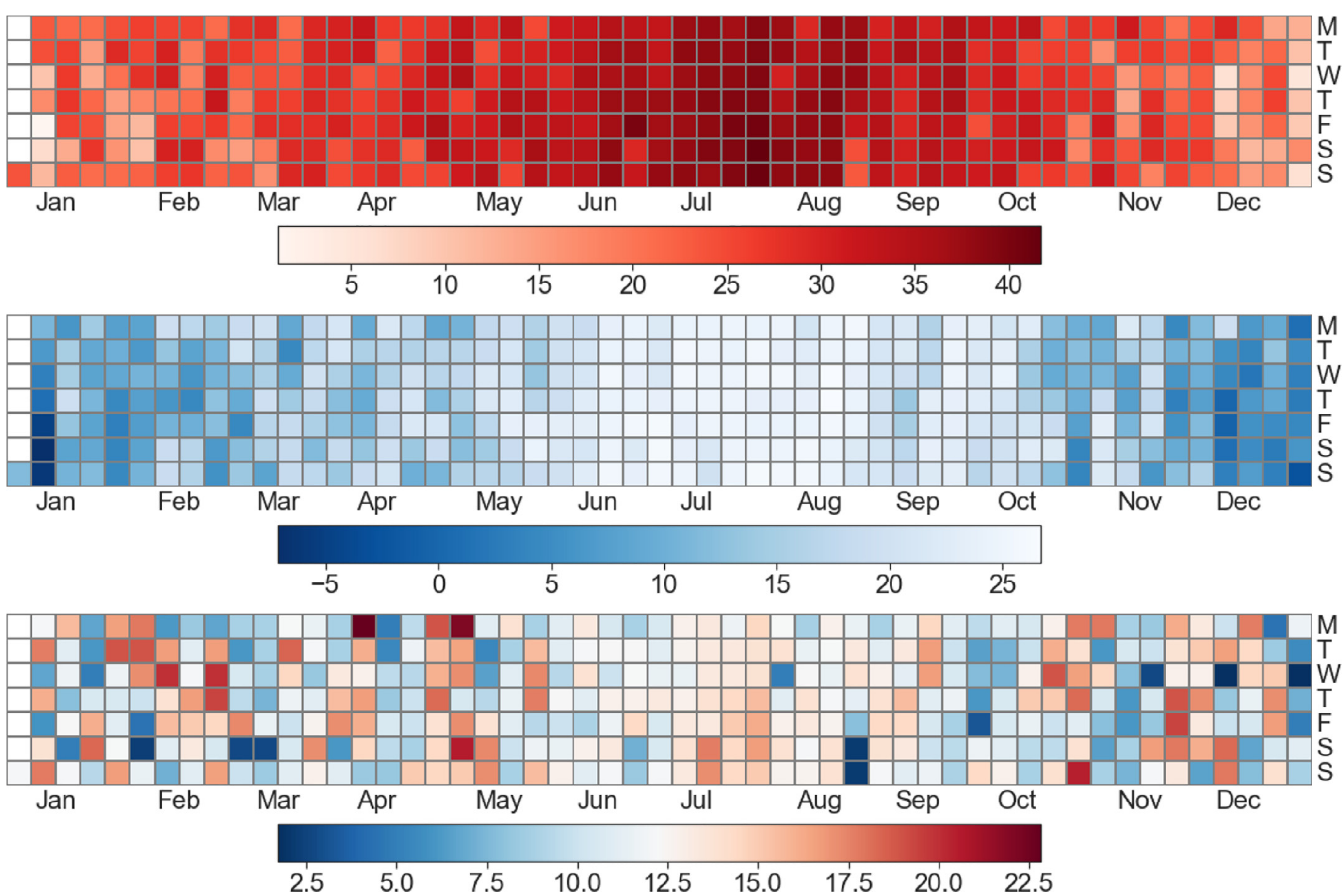

Fig. 14. Weather data visualization for Austin, Texas, in 2017 (top: maximum; middle: minimum; bottom: delta temperature in Celsius).

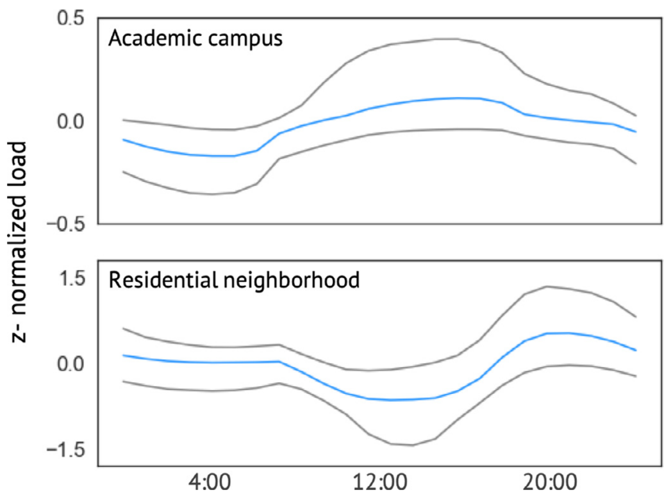

Fig. 15. Normal load profiles for each portfolio (two grey lines indicate 25% and 75% quartiles, and blue line shows median of all normal load profiles; each subplot has a different *y*-axis scale for visualization purposes). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

days in 2017) for the academic campus portfolio. In particular, the distances, *D*, of around Thanksgiving, Christmas, and April 20, are relatively high. With this observation, we can infer that such days were operated by very different schedules compared to their normal records, which is also partially confirmed in Fig. 7.

Similarly, in the residential neighborhood portfolio, there are some discord days during the first and last week in 2017, when people have special events. We also observed some discord days resulting from weather conditions, i.e., hot, cold, and swing season. Fig. 14 visualizes the outdoor weather data for 2017 in Austin, Texas. For example, the difference between the daily maximum and minimum temperature is relatively high in May and October (swing seasons) and those days are detected as discords. On the other hand, the discord days in February and August were the coldest and hottest days, respectively, in 2017. The observations in Figs. 21 and 14 suggest that the energy consumption of the residential neighborhood portfolio has a relationship with occupant behavior and outdoor weather conditions.

Additionally, we compare the distances, *D*, to determine the dominant error type for each portfolio. According to Figs. 20 and 21, there are generally more days with longer distances for the academic campus compared to the residential neighborhood portfolio. In other words, the load shapes for discord days

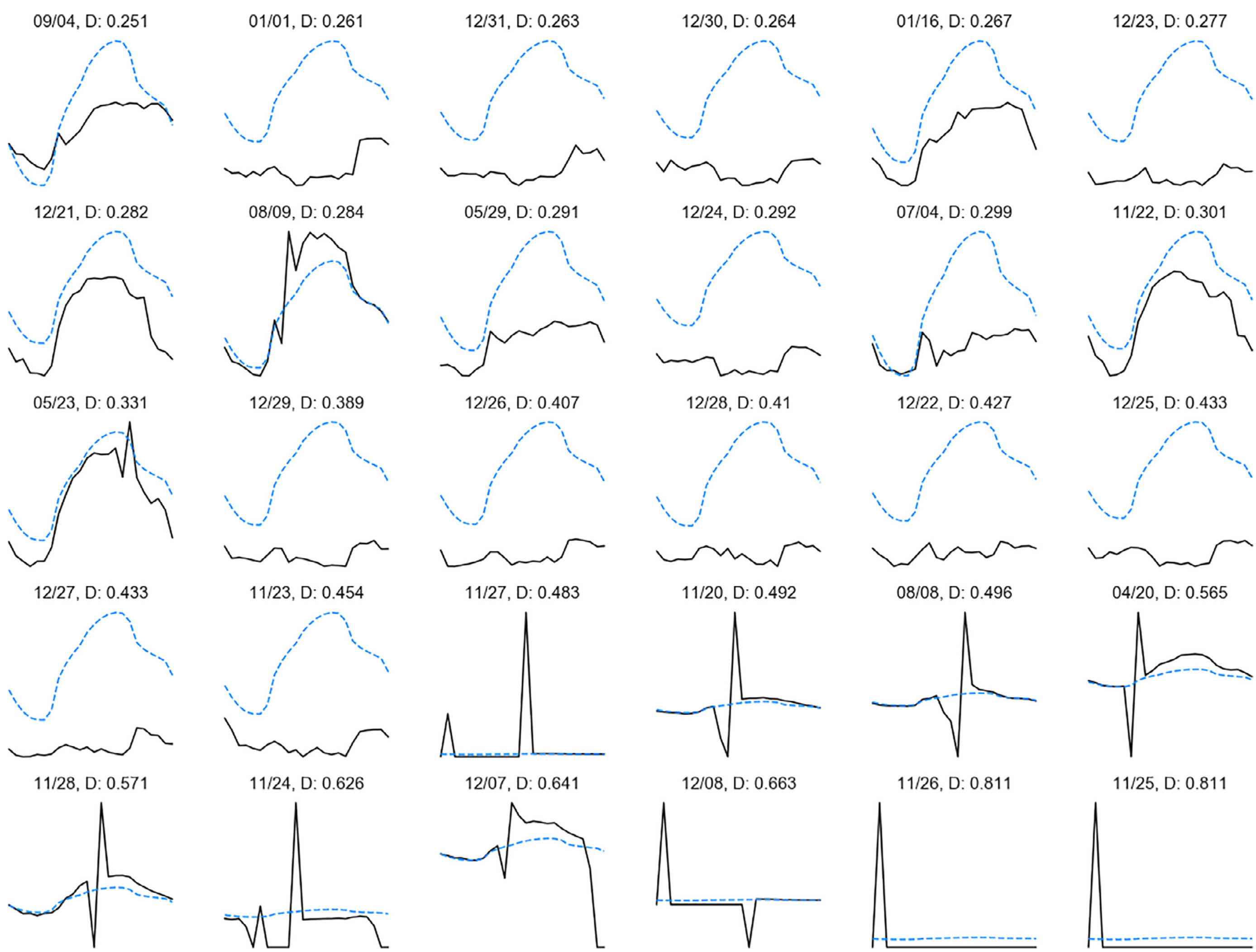

Fig. 16. Academic campus load profile discords ordered by D (black lines indicate individual discord profiles [50% quartile], whereas the dashed blue line shows the typical normal profile of the portfolio [Fig. 15]; each subplot has a different *y*-axis scale for visualization purposes). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

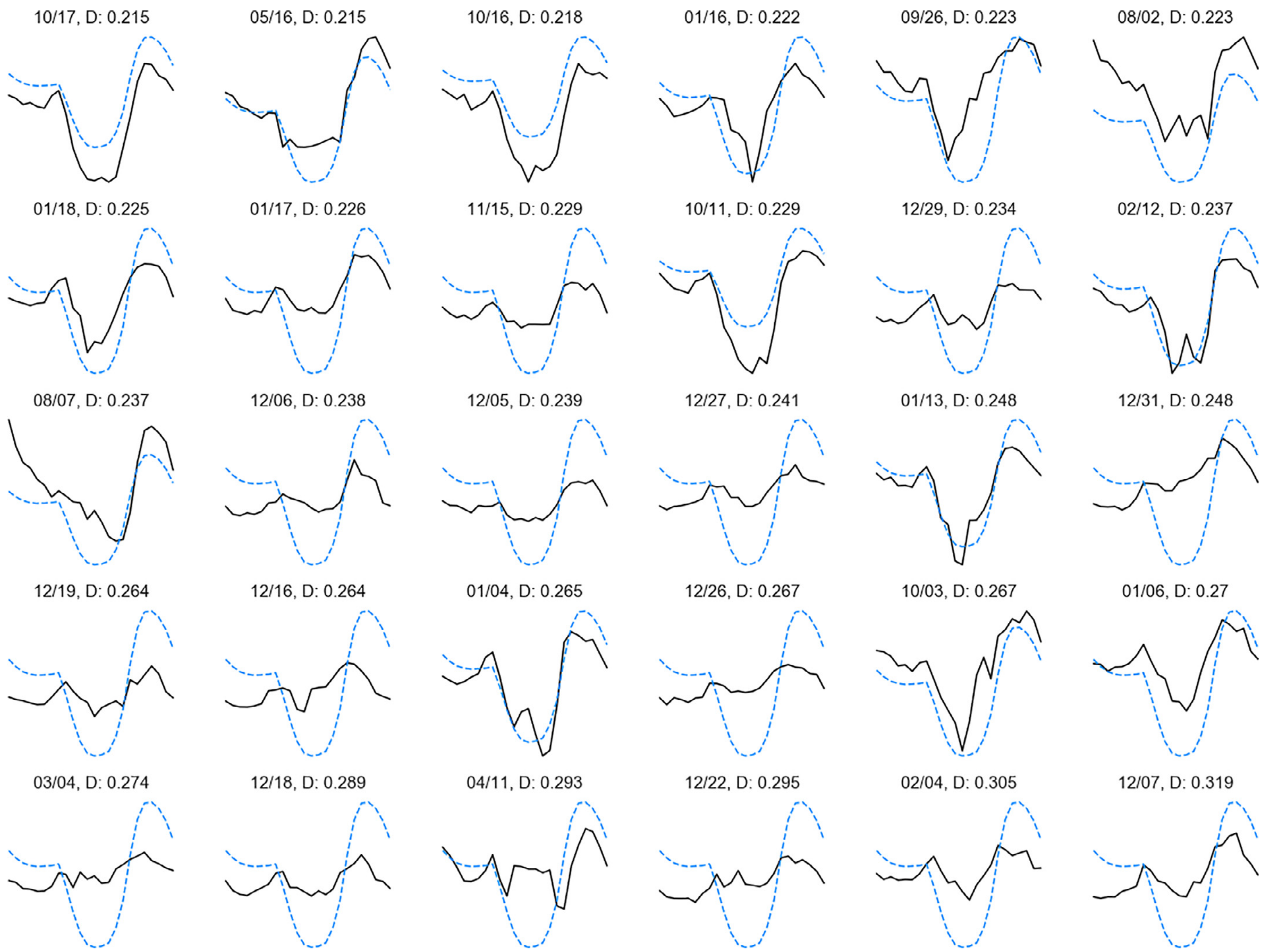

Fig. 17. Residential neighborhood load profile discords ordered by D (black lines indicate individual discord profiles [50% quartile], whereas the dashed blue line shows the typical normal profile of the portfolio [Fig. 15]; each subplot has a different *y*-axis scale for visualization purposes). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

are very unique compared to the load shapes for typical days, and these differences are very large in the academic campus portfolio. So, to summarize, the academic campus portfolio has fewer discord days with very unique shapes of discord days.

Fig. 15 illustrates the typical load shape pattern from the z-normalized load profiles of Day_{normal}. In general, the academic buildings consumed less energy before 6 a.m. and started to consume more energy after that. This suggests that most of the campus buildings follow the predefined operation schedule. There was a shallow peak in the afternoon, with a larger band of Q1 and Q3, mainly due to various PSUs. The energy consumption decreased after 6 p.m., which is regular off time in academic campus buildings. The residential neighborhood portfolio indicates a very different consumption pattern. There was steady energy consumption before 6 a.m., and very low energy consumption was detected around noon; then, it sharply increased around 4 p.m. (i.e., coming back from work or other activity). Similarly, there were two large bands of Q1 and Q3; this suggests that occupants in this portfolio behaved very differently around noon and evening. It is also important to note that the IQR band is larger in the residential neighborhood compared to the academic campus. This observation reconfirms that an individual house in the residential neighborhood shows very different energy consumption patterns.

Finally, we compared the discord profiles with the normal load profiles (Fig. 15). In Figs. 16 and 17, 30 discords are ordered by the distance, *D*, from the KS test and compared with the normal load profile. Note that each individual profile is visualized with a different *y*-axis scale in order to inspect the shape differences on each discord day.

For the academic portfolio (Fig. 16), it is clear that the top 10 discords (bottom) have meter malfunctions on those days. For example, the normal profile (dashed blue) appears in a steady shape because of extremely high peaks on those discord days. Alternatively, the second major discord indicates very low energy consumption compared to the normal load profile. This is mainly because those days are around New Year's, Thanksgiving, and Christmas, and the academic buildings were not operated under a regular schedule. The remaining discords follow the shape of the normal load profile in general, but have proportional discordance (e.g., spikes, drops, and offsets) in the profiles (May 23, and August 9).

In the residential neighborhood portfolio, the meter malfunction discord days were not detected. The shape difference was the main discord type. For example, the majority of the discord energy consumption patterns in Fig. 17 dropped around noon and increased afterward, but the magnitude of such patterns was low. Similar to the discords in the academic portfolio, we also observed proportional discord days (February 12, August 2, and October 16). Moreover, some days (April 11, August 7, and December 31) show totally different energy consumption patterns.

4. Discussion

In this research, we focused on the discord detection of multiple buildings as a portfolio analysis. As demonstrated in our case study, this type of analysis can provide important pre- and/or post-insights on a large set of smart metering data. Methodologically, *ALDI* weaves three components to identify discords in a large portfolio. First, the MP calculation is adopted for evaluating the similarity of each daily load profile in terms of its shape. This calculation method is relatively simple and has been evidenced by its computational efficiency compared to other discord detection techniques [37]. Considering that we will have to manage even more buildings in the near future, this will require an instantaneous calculation speed for discord detection. The MP-based framework would be the most suitable approach for this type of situation. The second step is a KS test. The associated p-value of this test is not firmly fixed; we discussed the implication of p-value selection for other facility managing purposes. The last component is visualization and evaluation of load profiles based on the hypothesis testing result. For facility managers, this final outcome is particularly useful, because they can intuitively understand different error types in their portfolio and develop proper maintenance plans for robust and resilient energy management systems.

4.1. Future applications

Our proposed framework is versatile for various applications introduced in Section 1. Specifically, it can be used as a potential preprocessing step to not only filter malfunctioning data points but also to characterize individual load profiles. According to a nationwide survey from 448 building management professionals [38], the adoption of data analytics and simulation is essential for the future of building energy management. In particular, rapid urbanization requires building and portfolio managers to have management skills for a large group of buildings [39]. Considering this, *ALDI* would be an ideal tool for diagnosing multiple buildings in terms of discord days.

In addition, the automation of detecting motif and discord types of individual load profiles can facilitate building-to-grid applications. To implement a demand response scenario [40] and renewable energy integration [41] on a grid, the owner/manager of such a portfolio needs to characterize temporal energy consumption patterns. For example, if the owner/manager of the micro-grid knows the typical dates and shapes of motif and discords of the portfolio, this information can be used for planning an energy management program.

4.2. Limitations of the framework

One possible improvement in *ALDI* is developing a systematic way of selecting p-value. Regardless, we intentionally open p-value selection for future usage of *ALDI*. Instead of using a fixed p-value, we report the discord detection results by different p-values. It should be noted that the p-value selection (0.01) in Section 3 is not necessarily our recommendation, but we wanted to demonstrate *ALDI* with a reasonable value. In addition, we understand that managers of large building portfolios have their own interests or problems, which require totally different p-values. Although we found that ALDI identifies more discord days with higher p-values, the main discord types with large D are same in Figs. 16 and 17. This suggests that facility managers can find the main discords by any reasonable p-value, but they can increase the p-value to discover other minor discord days.

The main assumption of our discord detection is that buildings within the same portfolio have similar discord load shape patterns, because they are located in the same geographical region (similar behavior patterns and outdoor weather conditions) and are connected with the same electrical grid and metering facility (same holiday schedules and system maintenance plan). Because of this assumption, the proposed framework is more suitable for an academic campus compared to a residential neighborhood. Because the buildings in the academic campus have similar schedules, our framework identified distinct discords. In addition, it sometimes failed to identify each individual building-level discord (local event or system fault for a specific building). However, the main contribution of *ALDI* is detecting the discords within a portfolio for community-level applications. For building-level studies, discord detection techniques [27–29] introduced in Section 1 are recommended. Finally, it should be noted that *ALDI* requires daily load profiles from buildings, which potentially limits the future applications of building energy management for buildings with low-resolution meters (e.g., buildings with only monthly or yearly metering data).

The last step of *ALDI* is visualizing discord profiles with normal profiles and quantitatively evaluating discord types. Although we identified the discord types for each portfolio, it should be noted that this evaluation process still relies on engineering decisions. To fully automate and robustly identify the discord types, we recommend future study integrating labels (discord types) and load profile data points to train the discord detection machine learning models. In fact, there are no such labels for every day load profiles; this reversely means that *ALDI* could provide useful insights for identifying those labels for further training of supervised learning models. In addition, our engineering decision used the distance, D, and p-value as metrics of quantifying the discrepancy of the discord profiles. For example, 10 load shapes with high distances in Fig. 16 were clearly measurement errors from metering system malfunctioning. However, the other error types (e.g., low energy consumption and normal profile with spikes) were not highly correlated with either distance, D, or p-values. This observation suggests that we should develop a robust metric for the quantification of the discord level against the normal profiles.

5. Conclusion

In this paper, we proposed *ALDI*, a data-driven framework to detect discord operation days at a portfolio level. *ALDI* consists of MP calculation, a KS test, and visual analytics. As a case study, we demonstrated *ALDI* with two actual energy meter data sets, from an academic campus and a residential neighborhood. Our results indicate that the academic buildings have lower MP values compared to residential houses, which suggests that the academic buildings were operated with predefined and consistent schedules, but they also have extreme load shape patterns mainly because of malfunctioning of the metering facility. Regarding discord days, we identified that the majority were national holidays in both portfolios, and the residential neighborhood portfolio had some discord operations days because of outdoor weather conditions (hot, cold, and swing season). *ALDI* provided comprehensive evaluations of our portfolios, and this type of knowledge is critical for integrating buildings with a smart grid.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by U.S. Department of Energy

Office of Energy Efficiency and Renewable Energy Building Technologies Office. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

CRediT authorship contribution statement

June Young Park: Conceptualization, Methodology, Formal analysis, Data curation, Visualization, Writing - original draft, Writing - review & editing. **Eric Wilson:** Conceptualization, Resources, Funding acquisition, Writing - review & editing, Supervision. **Andrew Parker:** Conceptualization, Resources, Funding acquisition. **Zoltan Nagy:** Writing - review & editing, Supervision.

Acknowledgement

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Building Technologies Office. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

Appendix

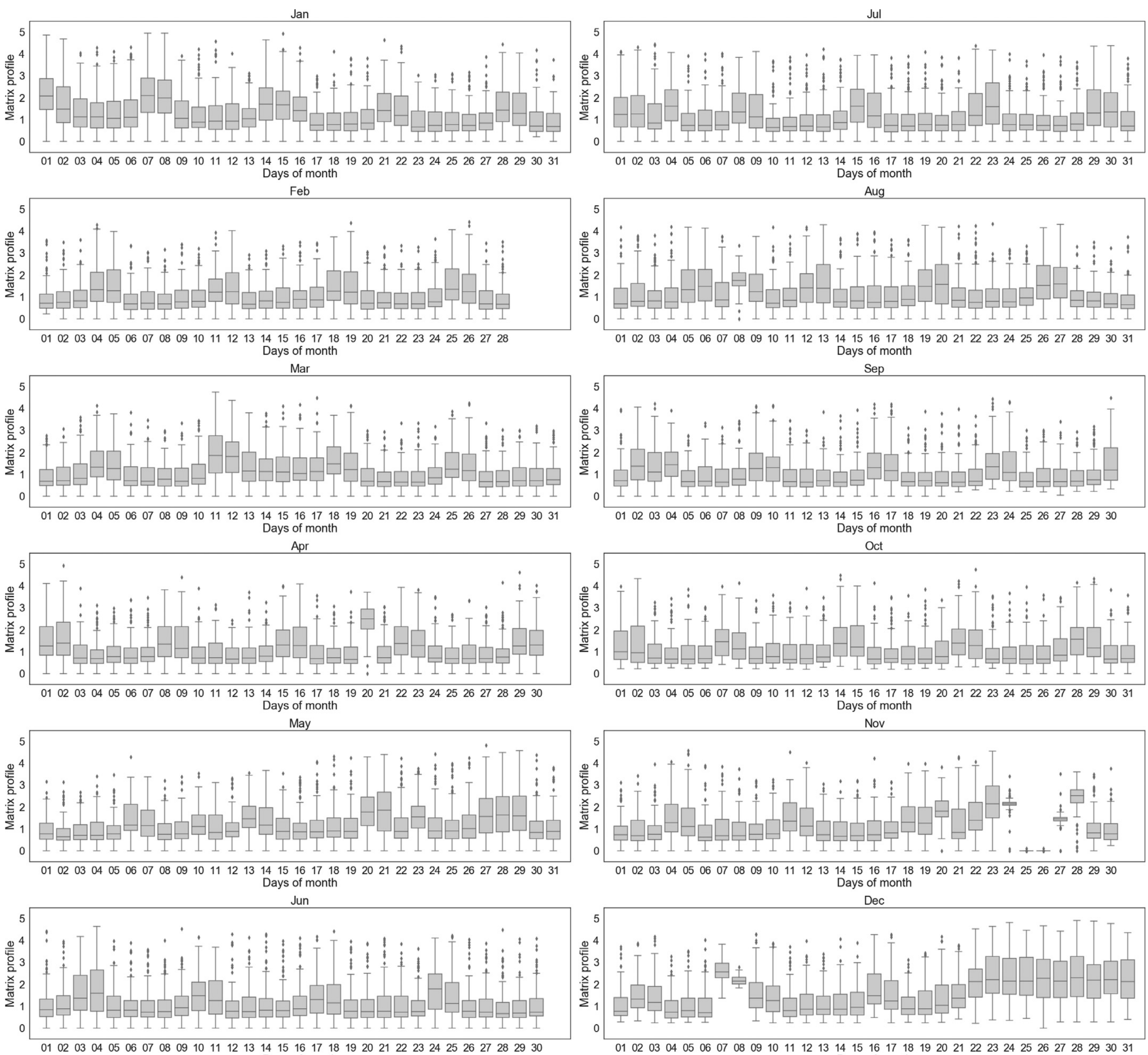

Fig. 18. Matrix profile distributions of academic campus portfolio by date (left: January–June, right: July–December); there are extraordinary MP values during nonsemester periods.

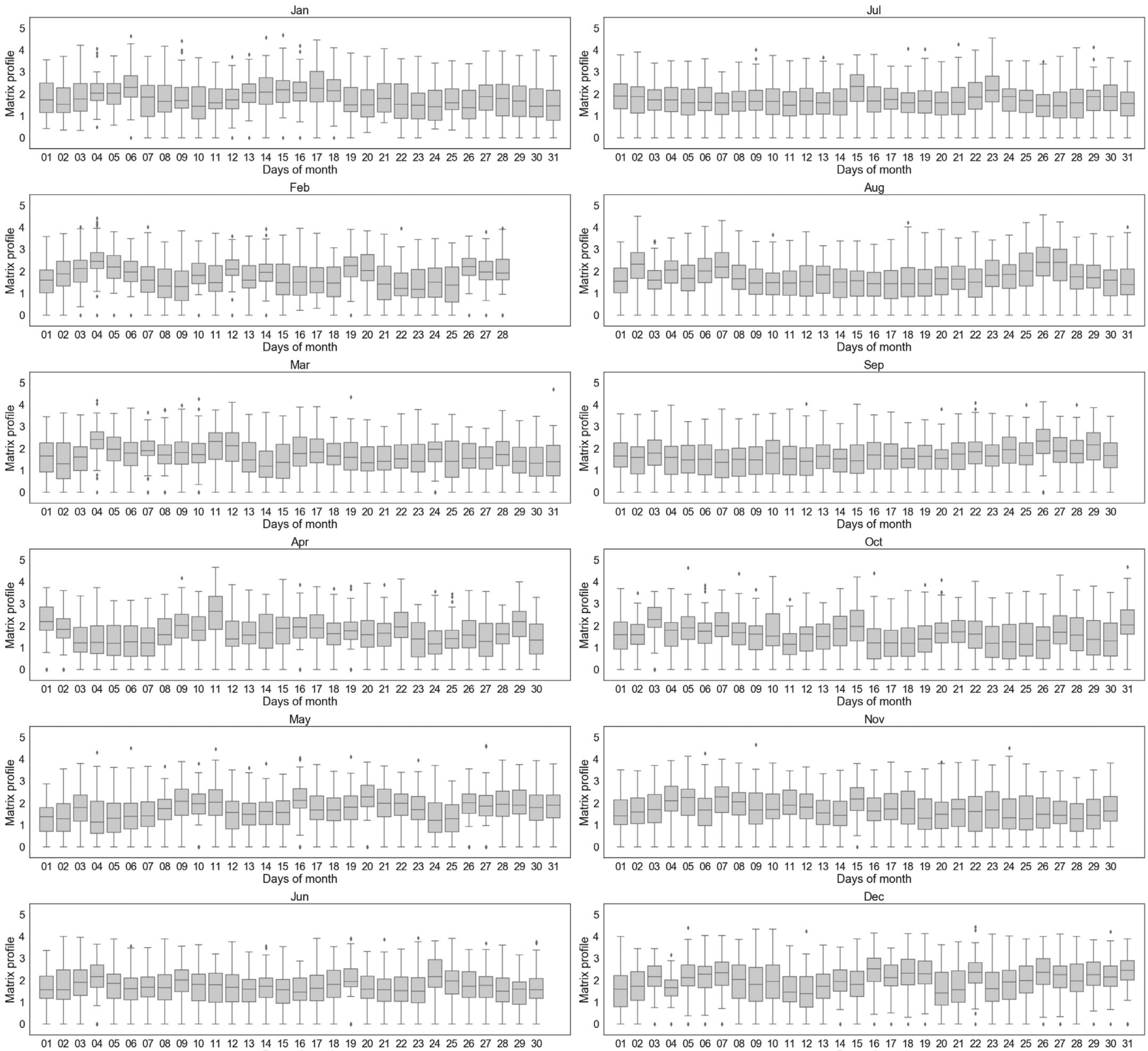

Fig. 19. Matrix profile distributions of residential neighborhood portfolio by date (left: January–June, right: July–December); there are extraordinary MP values during holiday and swing season periods.

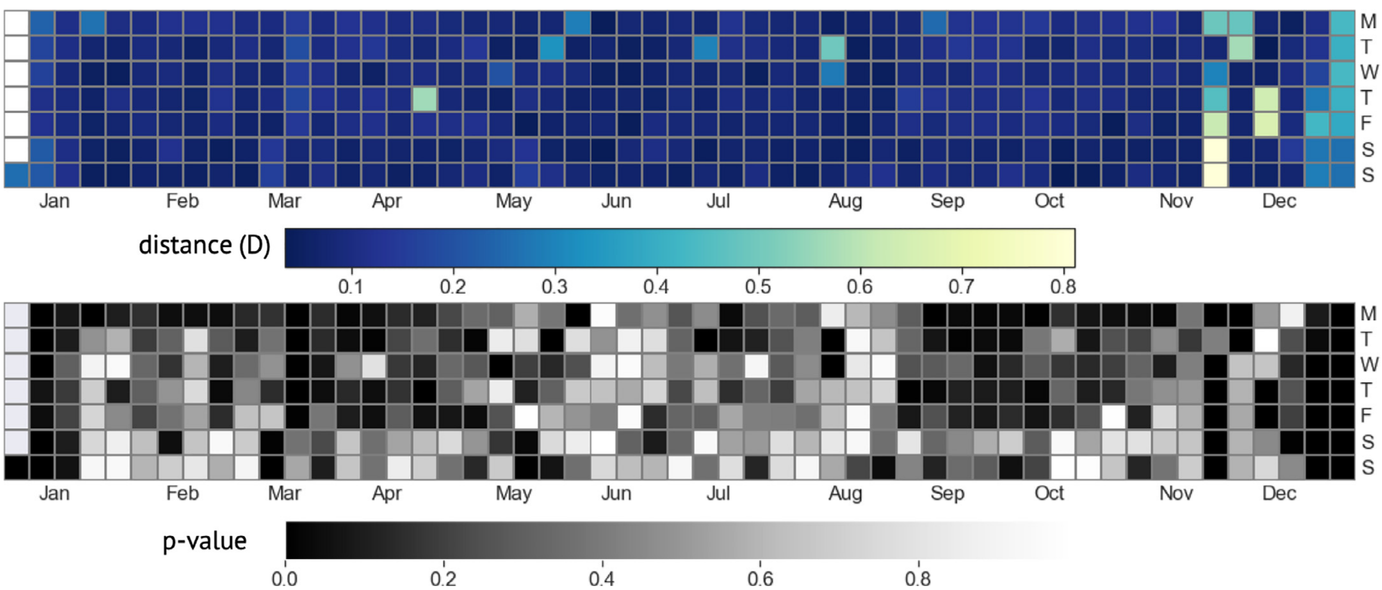

Fig. 20. Calendar view (row: 7 day types; column: week of month) of Kolmogorov-Smirnov test results for academic campus portfolio (top: distance (D); bottom: *p*-value).

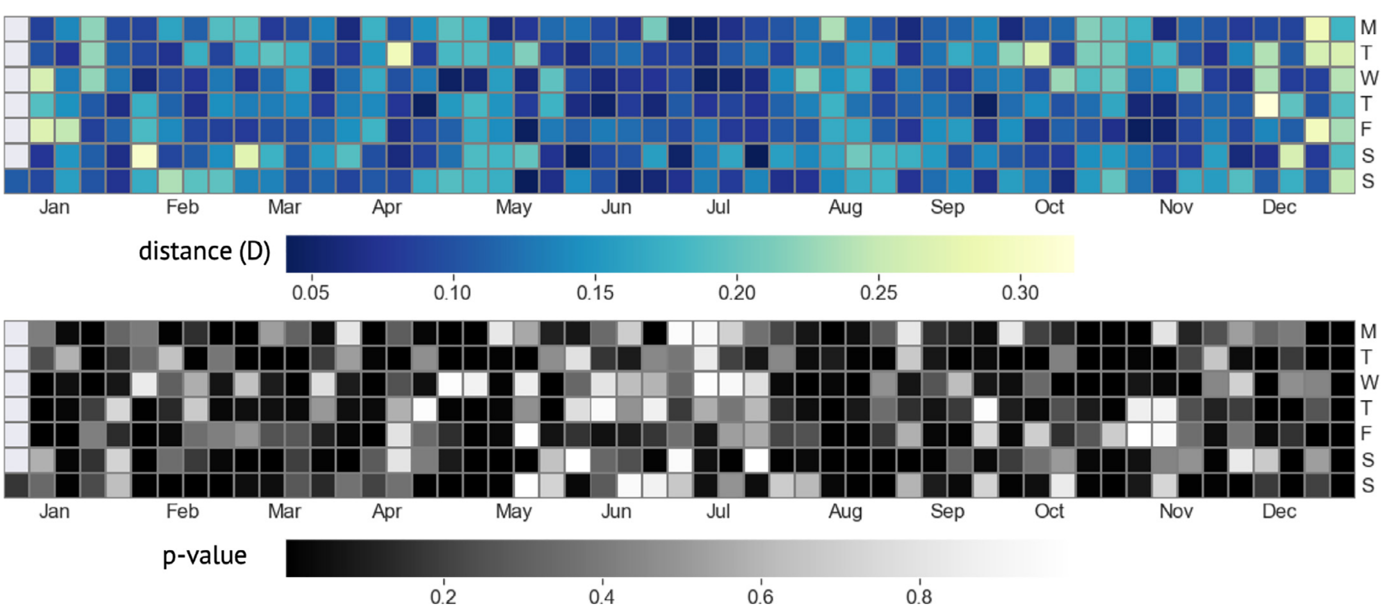

Fig. 21. Calendar view (row: 7 day types; column: week of month) of Kolmogorov-Smirnov test results for residential neighborhood portfolio (top: distance (D); bottom: *p*-value).

References

[1] U. D. of Energy, Buildings energy data book, 2011, (https://openei.org/doe-opendata/dataset/buildings-energy-data-book). Accessed: 2018-06-20.

[2] U. E. I. Administration, International energy outlook, 2018, (https://www.eia.gov/outlooks/ieo/). Accessed: 2018-06-20.

[3] J.Y. Park, Z. Nagy, Comprehensive analysis of the relationship between thermal comfort and building control research-a data-driven literature review, Renew. Sustain. Energy Rev. 82 (2018) 2664–2679.

[4] U.E.I. Administration, How many smart meters are installed in the united states, and who has them?, 2016, (https://www.eia.gov/tools/faqs/faq.php?id=108&t=3). Accessed: 2018-06-20.

[5] C. Miller, Z. Nagy, A. Schlueter, A review of unsupervised statistical learning and visual analytics techniques applied to performance analysis of non-residential buildings, Renewa. Sustain. Energy Rev. 81 (2018) 1365–1377.

[6] M.S. Piscitelli, S. Brandi, A. Capozzoli, Recognition and classification of typical load profiles in buildings with non-intrusive learning approach, Appl. Energy 255 (2019) 113727.

[7] K. Li, Z. Ma, D. Robinson, J. Ma, Identification of typical building daily electricity usage profiles using gaussian mixture model-based clustering and hierarchical clustering, Appl. Energy 231 (2018) 331–342.

[8] E. Calikus, S. Nowaczyk, A. Sant'Anna, H. Gadd, S. Werner, A data-driven approach for discovering heat load patterns in district heating, Appl. Energy 252 (2019) 113409.

[9] X. Luo, T. Hong, Y. Chen, M.A. Piette, Electric load shape benchmarking for small-and medium-sized commercial buildings, Appl. Energy 204 (2017) 715–725.

[10] J.Y. Park, X. Yang, C. Miller, P. Arjunan, Z. Nagy, Apples or oranges? Identification of fundamental load shape profiles for benchmarking buildings using a large and diverse dataset, Appl. Energy 236 (2019) 1280–1295.

[11] R.K. Jain, J. Qin, R. Rajagopal, Data-driven planning of distributed energy resources amidst socio-technical complexities, Nat. Energy 2 (8) (2017) 17112.

[12] S. Xu, E. Barbour, M.C. González, Household segmentation by load shape and daily consumption, in: Proceedings of ACM SigKDD 2017 Conference, Halifax, Nova Scotia, Canada, 2017.

[13] J. Every, L. Li, D.G. Dorrell, Leveraging smart meter data for economic optimization of residential photovoltaics under existing tariff structures and incentive schemes, Appl. Energy 201 (2017) 158–173.

[14] A. Nutkiewicz, Z. Yang, R.K. Jain, Data-driven urban energy simulation (due-s): aframework for integrating engineering simulation and machine learning methods in a multi-scale urban energy modeling workflow, Appl. Energy 225 (2018) 1176–1189.

[15] W. Chen, K. Zhou, S. Yang, C. Wu, Data quality of electricity consumption data in a smart grid environment, Renew. Sustain. Energy Rev. 75 (2017) 98–105.
[16] D.O. Afanasyev, E.A. Fedorova, On the impact of outlier filtering on the electricity price forecasting accuracy, Appl. Energy 236 (2019) 196–210.
[17] E. Fanone, A. Gamba, M. Prokopczuk, The case of negative day-ahead electricity prices, Energy Econ. 35 (2013) 22–34.
[18] A. Cartea, M.G. Figueroa, Pricing in electricity markets: a mean reverting jump diffusion model with seasonality, Appl. Math. Finance 12 (4) (2005) 313–335.
[19] L. Clewlow, C. Strickland, Energy Derivatives: Pricing and Risk Management, Lacima Publ., 2000.
[20] S. Borovkova, F.J. Permana, Modelling electricity prices by the potential jump-diffusion, in: Stochastic Finance, Springer, 2006, pp. 239–263.
[21] V. Chandola, A. Banerjee, V. Kumar, Anomaly detection: a survey, ACM Comput. Surv. (CSUR) 41 (3) (2009) 15.
[22] J.E. Seem, Pattern recognition algorithm for determining days of the week with similar energy consumption profiles, Energy Build. 37 (2) (2005) 127–139.
[23] M. Afzalan, F. Jazizadeh, J. Wang, Self-configuring event detection in electricity monitoring for human-building interaction, Energy Build. 187 (2019) 95–109.
[24] D.B. Araya, K. Grolinger, H.F. ElYamany, M.A. Capretz, G. Bitsuamlak, An ensemble learning framework for anomaly detection in building energy consumption, Energy Build. 144 (2017) 191–206.
[25] C. Fan, F. Xiao, Y. Zhao, J. Wang, Analytical investigation of autoencoder-based methods for unsupervised anomaly detection in building energy data, Appl. Energy 211 (2018) 1123–1135.
[26] J. Lin, E. Keogh, L. Wei, S. Lonardi, Experiencing sax: a novel symbolic representation of time series, Data Min. Knowl. Discov. 15 (2) (2007) 107–144.
[27] C. Miller, Z. Nagy, A. Schlueter, Automated daily pattern filtering of measured building performance data, Autom. Constr. 49 (2015) 1–17.
[28] J.A. Fonseca, C. Miller, A. Schlueter, Unsupervised load shape clustering for urban building performance assessment, Energy Procedia 122 (2017) 229–234.
[29] A. Capozzoli, M.S. Piscitelli, S. Brandi, D. Grassi, G. Chicco, Automated load pattern learning and anomaly detection for enhancing energy management in smart buildings, Energy 157 (2018) 336–352.
[30] C. Tu, X. He, Z. Shuai, F. Jiang, Big data issues in smart grid–a review, Renew. Sustain. Energy Rev. 79 (2017) 1099–1107.
[31] Harrington, P. (2012). *Machine learning in action*. Manning Publications Co.
[32] C.-C.M. Yeh, Y. Zhu, L. Ulanova, N. Begum, Y. Ding, H.A. Dau, D.F. Silva, A. Mueen, E. Keogh, Matrix profile i: all pairs similarity joins for time series: a unifying view that includes motifs, discords and shapelets, in: 2016 IEEE 16th International Conference on Data Mining (ICDM), IEEE, 2016, pp. 1317–1322.
[33] F.J. Massey Jr, The Kolmogorov-Smirnov test for goodness of fit, J. Am. Stat. Assoc. 46 (253) (1951) 68–78.
[34] D.Q. Goldin, P.C. Kanellakis, On similarity queries for time-series data: constraint specification and implementation, in: International Conference on Principles and Practice of Constraint Programming, Springer, 1995, pp. 137–153.
[35] UTAustin, Utilities & energy management, energy portal, n.d., (https://utilities.utexas.edu/). Accessed: 2019-7-1.
[36] PecanStreet, Pecan street dataport, n.d., (http://www.pecanstreet.org/). Accessed: 2019-7-1.
[37] Y. Zhu, C.-C.M. Yeh, Z. Zimmerman, K. Kamgar, E. Keogh, Matrix profile xi: Scrimp++: time series motif discovery at interactive speeds, in: 2018 IEEE International Conference on Data Mining (ICDM), IEEE, 2018, pp. 837–846.
[38] C. Srivastava, Z. Yang, J.R. K., Understanding the adoption and usage of data analytics and simulation among building energy management professionals: anationwide survey, Build. Environ. 157 (2019) 139–164.
[39] C. Nägeli, A. Farahani, M. Österbring, J.-O. Dalenbäck, H. Wallbaum, A service-life cycle approach to maintenance and energy retrofit planning for building portfolios, Build. Environ. (2019) 106212.
[40] J.R. Vázquez-Canteli, Z. Nagy, Reinforcement learning for demand response: a review of algorithms and modeling techniques, Appl. Energy 235 (2019) 1072–1089.
[41] J. Yan, Y. Zhai, P. Wijayatunga, A.M. Mohamed, P.E. Campana, Renewable energy integration with mini/micro-grids, Appl. Energy 201 (2017) 241–244.