Energy & Buildings 177 (2018) 184–196

Contents lists available at ScienceDirect

Energy & Buildings

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

Building electricity consumption: Data analytics of building operations with classical time series decomposition and case based subsetting

Ethan M. Pickering*, Mohammad A. Hossain, Roger H. French, Alexis R. Abramson

Case School of Engineering, Case Western Reserve University, Cleveland, OH 44106, USA

ARTICLE INFO

Article history:
Received 22 February 2018
Revised 13 June 2018
Accepted 25 July 2018
Available online 9 August 2018

Keywords:
Building energy efficiency
Classical time series decomposition
Data analytics
Commercial buildings
Virtual energy audit

ABSTRACT

The commercial building sector consumes approximately one-fifth of U.S. total energy and exhibits significant operational inefficiencies, leaving a great opportunity to implement various energy-efficiency measures. However, conventional energy audit techniques are expensive, time-consuming, and frequently inaccurate. Conversely, classical time series decomposition of smart meter (i.e. 15 min interval) building electricity consumption provides quick, inexpensive, and useful insights to building operation and characteristics. Paired with complementary time series datasets such as outdoor temperature and solar irradiation, specific insights into HVAC scheduling, daily operational variation, and the relative impact of temperature and solar radiation were quantitatively assessed. This work analyzes six commercial buildings and identifies various building characteristics, including the potential for savings of over 700 MWh valued at $92,000 per year from building rescheduling alone. With access to only whole building smart meter data, these results are obtained virtually and instantaneously, making the case for a rigorous data analytics approach to unlock the potential of building energy efficiency.

1. Introduction

Energy is an ever important issue capturing the attention of all major countries, as evident from the 2015 United Nations Climate Change Conference in Paris, France. A massive consumer of energy is the building sector where there is significant opportunity to reduce energy waste [1]. The United States uses approximately 100 quads of energy each year, with about 40% attributed to building uses such as heating, cooling, lighting, and electronics [2,3]. The U.S. Department of Energy (DOE) has recognized this potential and is committed to reducing commercial building energy use 20% by 2020, as well as citing a long term goal of 50% reduction in overall use [4]. The "Prioritization Tool", developed in the Building Technologies Office (BTO) of the DOE, has calculated the energy savings that could be achieved from currently available and emerging technologies [2,3]. The analysis revealed that implementation of cost-effective technologies available today could lead to a 30% reduction of energy use in buildings by 2030. Accounting for emerging technologies estimated to become cost-effective within 5 years, the energy savings reaches 55% [2,3]. This is equivalent to saving up to $300 billion per year if investments in energy efficiency are made strategically.

Considering the long lifetimes and slow replacement rates of buildings, energy retrofits are necessary to appreciably reduce energy consumption in existing building stock [5]. These retrofits typically are identified and quantified in terms of cost and payback time through conventional energy audits or building information modeling. Conventional audits entail a physical walk-through of a building, which can include performing leak tests, infrared imaging, blower door tests, equipment sub-metering, extensive sensoring, and more. Therefore, these audits require a team of individuals to survey an entire building, can result in high costs, and can take days to weeks to perform [6]. Further, studies have found that the recommendations from these audits can vary drastically between energy audit companies [7]. Considering the time, cost, and variability in energy recommendations, building managers frequently question the economic benefit and ultimately may refrain from mobilizing their company to conduct conventional energy audits [6–9]. With building information modeling, one can gain an even deeper understanding of a building's retrofit potential. These physics-based models, such as EnergyPlus, BLAST, DOE-2.1E, TRNSYS-TUD, and ESP-r among others, can assess energy efficiency, but require thousands of inputs [10–13]. These models are often time-consuming and cumbersome, requiring substantial calibration to historical consumption data through various refinement parameters [14,15]. Despite the large number of required inputs and other detailed information, various parameters, such as the solar irradiance gain or the convective heat transfer coefficient,

* Corresponding author at: California Institute of Technology, 1200 East California Blvd., MC104-44, Pasadena, CA 91106, United States.
E-mail address: pickering@caltech.edu (E.M. Pickering).

https://doi.org/10.1016/j.enbuild.2018.07.056
0378-7788/

frequently deviate (up to 30% from actual values) and result in serious accuracy limitations and varied energy recommendations [11,12,16,17]. Although solutions can be cost-effective and payoff within several years, building efficiency faces great obstacles in implementation. The lack of progress could be blamed on risk aversion and distrust of return-on-investment (ROI) estimates [3]. The industry needs a transformational solution to efficiently diagnose problems, build trust in solutions, accelerate their implementation, and validate their future economic return.

Data analytics for building energy audits has the potential to provide this new solution for energy efficiency, with a renewed approach leveraging Big Data, or multiple time series datasets at high resolution including building energy consumption and weather [18]. Big Data is of growing interest, enabling a more rigorous analytics approach that has the potential to uncover new insights to building energy efficiency [19]. Select data analytics approaches applied to buildings have been employed recently to measure the energy savings associated with building retrofits [20–22] and efficiency programs, but substantial progress is yet to be made [23–25]. For example, techniques such as anomaly detection, clustering, and time series analysis [26–30] have been used to study building operations and behavior. Anomaly detection has helped identify irregular behaviors [29] and anomalous consumption in real-time [27,28], pointing to potential acute problems in the building (albeit not general operational characteristics [31]). Clustering techniques, although limited in their application to-date, have allowed for classification into various usage categories (e.g. days of the week or occupied/unoccupied times) [26]. Time series methods, such as using spectral time series or Fourier decomposition analyses, have helped reveal more specific behaviors of the building [32–34].

Many of the limitations of these rudimentary data analytics techniques can be overcome using classical time series decomposition to identify more specific operational characteristics of buildings and point to opportunities for energy savings. Classical time series decomposition techniques have been used in a variety of applications including studies of economics, weather, energy, and other time varying phenomena, but with limited application to buildings [35,36,38,37]. One example is the use of classical time series decomposition to develop building consumption forecasting models, but this approach does not pinpoint the presence of building systems or identify efficiency measures [39,40]. The work described herein augments this type of analysis by applying classical decomposition to provide unique insight to building operation.

We will discuss the data characteristics, cleaning and data assembly, followed by discussion of the time series analysis methods used. The results and discussion of applying these time series methods to six commercial buildings in two climate zones are then presented, showing day of the week effects, heating and cooling characteristics and the impact of solar thermal load on a building during heating seasons.

2. Data cleaning and assembly of dataframes

The analysis presented describes the operational signatures of six commercial buildings in two locations: San Jose, California (SJ,CA) and Richardson, Texas (RN,TX). Table 1 describes each building by their location, size, type, Köppen Geiger climate classification [41–43], and HVAC characteristics. Building data was collected from utility electricity meters (kWh) taken at 15 min intervals for approximately 2 years. The "Electric HVAC" column in Table 1 refers to whether building electricity is used to operate the main heating or cooling system. Electricity used for heating and cooling is an important characteristic since only electricity (i.e. not natural gas) data is analyzed here. Nonetheless, even though some of the buildings do employ natural gas systems, there are still several HVAC equipment in these buildings (e.g. air handlers) that use electricity and can be analyzed accordingly. Hourly weather data was collected from publicly available National Oceanic and Atmospheric Administration (NOAA) datasets (typically within 25 miles from building location) [44]. In addition, 30 min interval weather datasets with a 3.5 km spatial resolution provide comparison weather datasets and a more localized weather data source. This 30 min data is derived from satellite based imaging, combined with an empirical atmospheric model using a geographic information system (GIS) framework [45,46]. The NOAA and GIS weather data files include: temperature, dew point, wind speed, and global horizontal irradiance. Our prior study demonstrated the strong correlation among these datasets [18], and specifically for this study the weather datasets correlate such that RN, TX: $r = 0.995$ and for SJ, CA: $r = 0.887$. Both NOAA and GIS weather datasets are used interchangeably in this analysis on the premise that future analyses may employ either datasets depending on building location or data availability. Richardson, TX is in a Cfa Köppen Geiger climate zone and exhibits a temperate climate with mild winters, hot summers (average maximum temperature of 77.7 °F (25.4 °C) and average minimum temperature of 57.3 °F (14.1 °C)), and skies that are typically sunny to partly cloudy. In contrast, San Jose, CA is in a Csb Köppen Geiger climate zone and has a Mediterranean climate with mild winters and dry, warm summers (average maximum temperature of 69.5 °F (20.8 °C) and average minimum temperature of 51.8 °F (11.0 °C)). The datastreams from all three data sources were assembled into time-series R dataframes for each building, with predictors (independent variables) and responses (dependent variables) as columns, and the observations stored as rows, indexed by their 15 min timestamp, as described in [18].

Table 1
Building characteristics of all six buildings. Includes location, size, purpose, climate, and electric HVAC [18].

Building	Location	Size (ft^2)	Purpose	Climate	Electric HVAC
#1	RN, TX	226,000	Office/Lab	Cfa	Heat
#2	RN, TX	168,000	Office/Lab	Cfa	Heat/Cool
#3	RN, TX	244,000	Office/Lab	Cfa	Heat/Cool
#4	SJ, CA	109,000	Office	Csb	Cool
#5	SJ, CA	115,000	Office	Csb	Cool
#6	SJ, CA	168,000	Office	Csb	Cool

3. Statistical methods of classical time series decomposition of building energy consumption

Time series decomposition is a statistical method which can uncover underlying trends that may be useful for prediction, for determining typical behaviors in a dataset, or for developing an initial understanding of a building's performance. Classical decomposition uses the method of moving averages to examine trend cycles and "seasonal" or periodic behaviors (which should not to be confused with the four typical meteorological seasons of a year) [32,47]. The *trend* cycles capture the long term behavior of the data, while *seasonal* aspects are revealed from the periodic (i.e. daily) fluctuation of the data. Unexplained events and uncertainty must also be assessed to fully represent the behavior of the data as described by an error term or *random* component. Mathematically the decomposition is described as:

$$X_t = f(M_t, S_t, Z_t) \tag{1}$$

where X_t is the observed time series data, M_t is the *trend* component, S_t is the *seasonal* component, and Z_t is the error term, or *random* component. The additive model:

$$X_t = M_t + S_t + Z_t \tag{2}$$

is used here to capture the seasonal variation changes through successive periods [47]. The variables M_t, S_t, and Z_t are developed

through moving averages defined by the observed period of the data. Decomposition requires an input of period (P = one day) to first calculate the trend component, M_t, using a $2 \times P$-moving average if P is even and a P-moving average if P is odd. The detrended series, D_t is calculated as

$$D_t = X_t - M_t \tag{3}$$

Second, the detrended series is averaged over the given period to compute the seasonal component, S_t. To complete the decomposition, the remainder, Z_t, is calculated by

$$Z_t = X_t - M_t - S_t. \tag{4}$$

These three result datasets represent a decomposition of the original time series dataset. The seasonal component produces information on daily operations, the trend indicates the underlying movement of electricity consumption, and the random component may pinpoint specific periods in the data displaying uncharacteristic operation.

3.1. Classical time series decomposition for building daily operation signatures

Classical time series decomposition is performed on two years of building data using a period of 24 h (i.e. P = 96 points for 15 min interval time series datasets). A period of one day allows for the seasonal, or periodic, component to reflect the average electricity consumption for a given day. Fig. 1 demonstrates the entire decomposition method performed on datasets of weekdays in July and August (2012–2014), including temperature and electricity data from Building 1. The basic components of a classical time series decomposition are shown: observed, trend, seasonal, and random. The observed data can be "restored" or replicated from these components by summing the trend, seasonal, and random components. First, the temperature data is decomposed to highlight the characteristics of the seasonal, trend, and random. The seasonal component captures the diurnal temperature fluctuation throughout each day and repeats this exact seasonal component throughout the decomposition. The trend component captures the overall magnitude and progression of temperature, while the random component displays any deviations in the observed data from the trend component and seasonal component data in aggregate. With regards to electricity data, a similar decomposition applies. The trend component represents the magnitude of the consumption, 340 kWh. Graphically, the trend component appears to move substantially, but has less than a 2% variability over the week interval. The random component exhibits many jagged features as it encompasses events unexplained by the daily periodicity, which is most likely unscheduled human interaction with plug load, lab experiments, or other non-HVAC equipment operation. Finally, the seasonal component is seen repeated 5 times over, again due to the daily periodicity, which repeats each week for the work days of Monday through Friday. This repeated seasonal component yields the most interesting insights into the typical daily operation of a building and is analyzed in more detail in Section 4. Note that both seasonal and random components have negative values as a result of the additive decomposition and are relative to the day-long moving average. The values of both random and seasonal components are instructive for magnitude comparisons and are not a negative electricity usage.

3.2. Case based subsetting

Building operation depends on a variety of different predictors and associated responses. The method of subsetting, retrieving only parts of datasets which are of particular interest, allows for the division of the data into various groups based on defined criteria (e.g. day of week, exterior temperature) and provides an opportunity to compare various characteristics among the criteria. For example, choosing only electricity data which occurred on a Monday is a subset of the entire energy dataset from index 1 to n, where n is the number of days in the electricity dataset. Therefore, the subset of data E_{Monday} is approximately 1/7 of the total dataset and can be interpreted mathematically as:

$$E_{Monday} \subset E_{total} = E_{j=1}^{n} \tag{5}$$

As noted earlier, the trend component of a classical time series is computed using a $2 \times P$- or P- moving average, therefore the dataset must remain continuous for a correctly computed classical time series decomposition. However, most subsetted datasets are no longer continuous (although monotonically increasing in time) such as the above example considering only Mondays. To navigate this continuity condition the trend of the entire dataset and the detrended dataset are computed before subsetting. The detrended dataset can then be subsetted into a semi-continuous form where full periods, in this case days, are kept in intact to ensure a correct seasonal component is calculated. The seasonal component is subsequently calculated by averaging each timepoint (i.e. each 15 min interval in this study) and the random is computed by subtracting the associated trend and seasonal from the observed dataset. Again, verifying each period of the dataset is kept intact following subsetting is a necessary condition and therefore any subsetting smaller than one period (e.g. hourly variable subsetting criteria) would lead to an incorrect decomposition.

3.3. Day of the week

Subsetting the datasets into each day of the week provides specific additional insight that one might not appreciate otherwise. As previously mentioned, the continuity of the energy data is compromised when applying subsetting techniques. For example, a subset of only Mondays do not represent a continuous dataset in a day-of-the-week analysis and the corresponding baseline/night operation might vary. To account for these variations when comparing each day of the week against one another, each seasonal component is shifted so that the first timepoint, in this case midnight, starts at 0 kWh. This allows for clearer comparisons among the subsetted data for given days. All six buildings are analyzed by day of the week and presented in the results section.

3.4. Heating- and cooling-season operations

The building data can be further subsetted into heating and cooling seasons for additional insight. In this paper heating seasons are defined as the months of December–February, while cooling season includes months June-August. Additionally, temperature constraints are included in the subsetting criteria to ensure outliers are omitted, such as unusually warm heating-season days. The criteria is set using a full day mean exterior temperature and the constraints for the temperature ranges are set differently for each climate. Eq. (6) defines a full day mean exterior temperature with $N = 96$.

$$T_{mean} = \frac{1}{N} \sum_{j=1}^{N} T_j \tag{6}$$

Where the set of T_{mean} temperatures can be defined as:

$$(T_{mean})_{k=1}^{n} \tag{7}$$

where n is the total number of days in the dataset. Fig. 2 shows the large discrepancy between the two climates and the need for individualized temperature constraints. Richardson shows high cooling-season temperatures, as well as colder heating-season temperatures, while San Jose displays a mild climate and a much

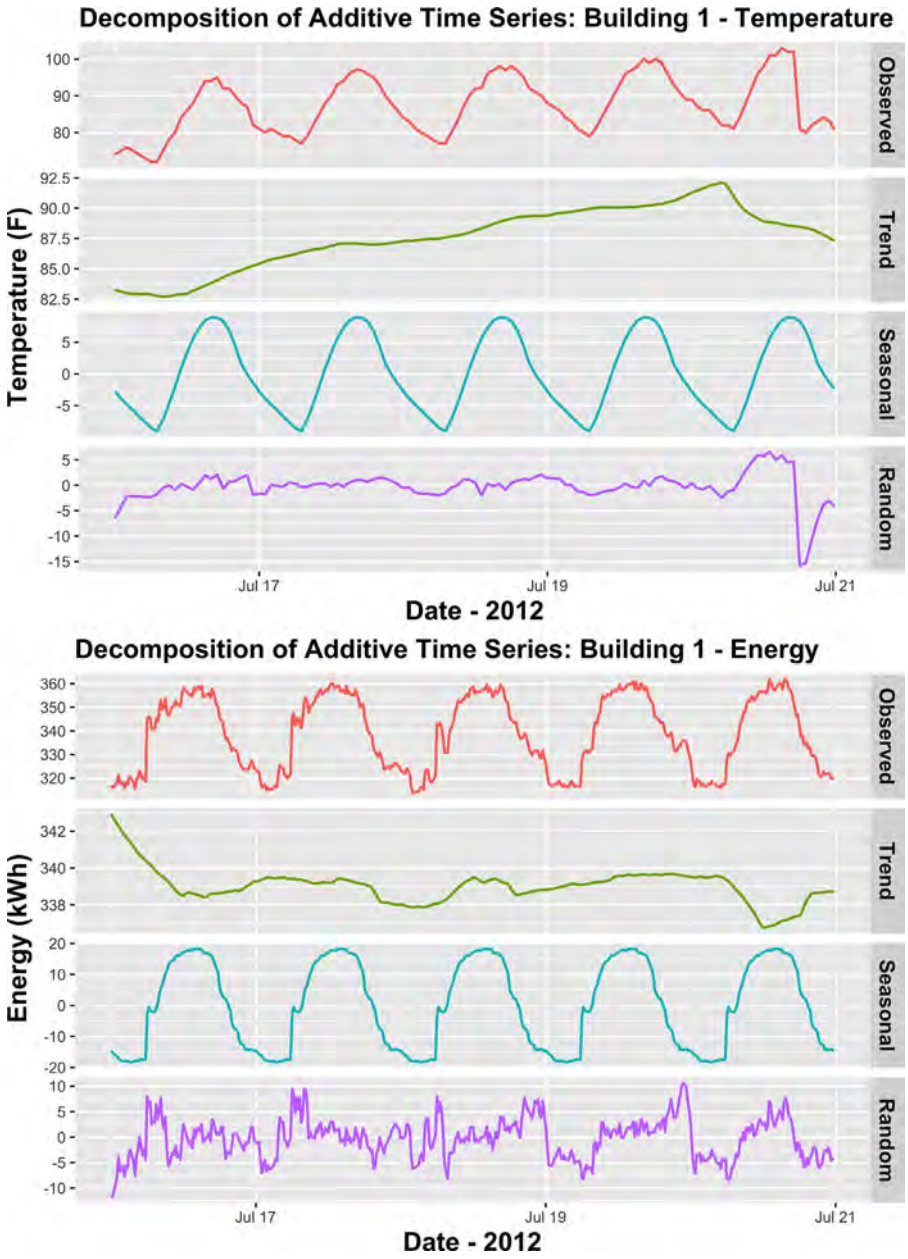

Fig. 1. Time series decomposition of temperature (top) and energy (bottom) for Building 1 broken down into observed, trend, seasonal, and random components.

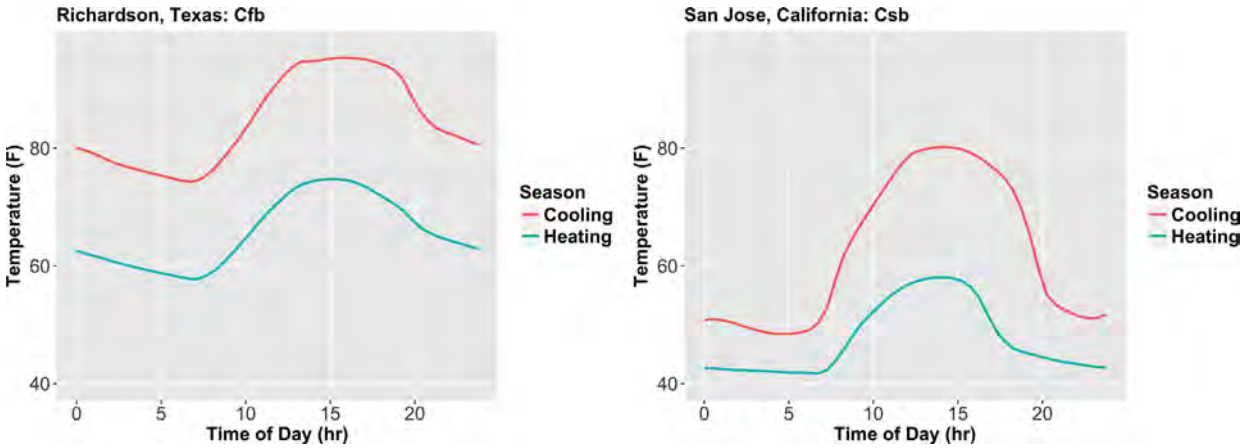

Fig. 2. Cooling and heating season climate temperature differences - Richardson, TX and San Jose, CA.

smaller difference between cooling- and heating-season temperatures. In an effort to generalize an approach accounting for these different climates, the average and standard deviation of the mean day temperatures on each climate are used to determine heating and cooling seasons. In particular we observe that the temperatures above the mean plus one third of the standard deviation,

$$T_{Cooling} = T_{mean,k} < \mu((T_{mean})_{k=1}^{n}) + \sigma((T_{mean})_{k=1}^{n})/3, \quad (8)$$

necessitate cooling operations and temperatures below the mean minus one third of a standard deviation,

$$T_{Heating} = T_{mean,k} < \mu((T_{mean})_{k=1}^{n}) - \sigma((T_{mean})_{k=1}^{n})/3, \quad (9)$$

require heating, where μ is the mean of all T_{mean} and σ is the standard deviation of all T_{mean}. Additionally, when considering normally distributed weather data, this metric ensures both cooling and heating seasons retain approximately 37% of the data(270 days) for further time series analysis. This approach results in the following ranges: Richardson cooling $>$ 84.6 °F (29.2 °C) and heating $<$ 65.7 °F (18.7 °C); San Jose cooling $>$ 62.4 °F (16.9 °C) and heating $<$ 47.5 °F (8.6 °C). Again, the determination of using one third of the standard deviation is based upon observations, where the temperature ranges make physical sense and give distinctly differing building operations shown later in this study. This also ensures the requirement of sufficent days for time series analysis. A more robust method for determining heating and cooling seasons could prove useful as more climates are considered.

3.5. Solar thermal load in cold heating-season operation

Solar irradiance is known to impact the thermal loads on buildings [11,48] and the impact of irradiance can be determined by additionally subsetting the assembled building datasets by the mean daily solar irradiance of sunny versus cloudy days. During cold heating-season temperatures, buildings' heating loads are lessened during high solar irradiance days compared to low solar irradiance conditions. To apply a time series decomposition to the dataset, both temperature and solar irradiance constraints were identified such that at least 10 days pass each criteria set and create a large distinction between sunny and cloudy days. The constraints were set for this analysis such that a cold heating-season day was redefined as having a mean temperature less than 48 °F (8.9 °C). This constraint is necessary to ensure that the cold days represent days that entirely rely on the heating system; on a warmer mean temperature day, such as 57 °F, with extremely sunny conditions, the building may not require heating HVAC and may actually use cooling HVAC instead. A sunny day is defined for these two locations as having a mean solar irradiance above 210 W/m^2 and a cloudy day as less than 70 W/m^2. The mean solar irradiance is calculated considering a full day data, including night time values (i.e. 0 W/m^2). Unlike the other building operational signatures, the signatures are computed using the seasonal component plus the mean of the trend values to account for any baseload differences between cloudy and sunny operations. The impact of solar irradiance is only investigated on Building 1 and is presented in the results.

3.6. Error quantification of classical time series decomposition

The error of each building time series datasets is assessed for a normally distributed 95% confidence interval. In a classical time series analysis the random component, Z, of the decomposition captures the error in the analysis. For each decomposition, the error, μ_Z, is computed by

$$\mu_Z = 1.96 \times \sigma_Z/\sqrt{n}, \quad (10)$$

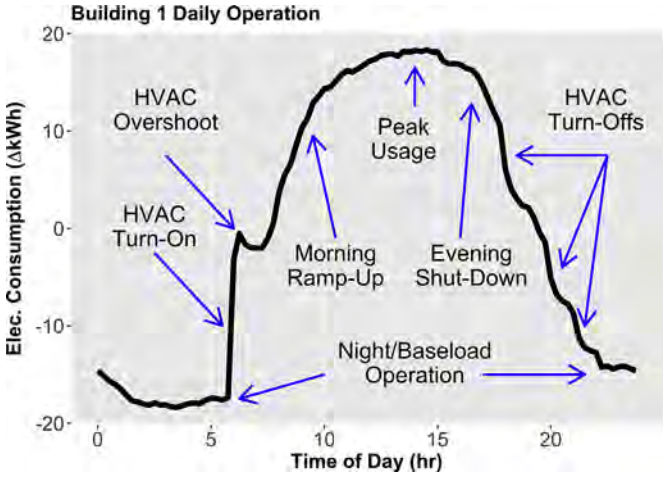

Fig. 3. Building 1 daily operation with operational characteristics identified, determined from two years of data using only July/August weekdays.

where 1.96 is the normal distribution 95% confidence interval coefficient, n is the number of days in the analysis, and σ_Z is the standard deviation of the random component of energy consumption. The standard deviation of the random component is calculated by

$$\sigma_Z = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Z_i - \bar{Z})}. \quad (11)$$

This error is computed for both the heating- and cooling-season operations and for the solar thermal load analyses, both of which have the least amount of days after subsetting.

4. Results and discussion

4.1. Classical time series decomposition of building energy consumption

Using classical time series decomposition, the average electricity consumption can be analyzed to illustrate a building's typical daily operation, or daily operational signature. This can be accomplished using at least one or preferably two or more full years of data. As an example, Fig. 3 shows one period, or one day, of the seasonal component of electricity data in Fig. 1. This average electricity consumption shown in Fig. 3 is a building's daily operational signature. The data analyzed corresponds only to summer weekdays gathered from over two years for Building 1, with electricity consumption along the y-axis and the hour of the day along the x-axis. By examining this representative curve, one can gain specific insight into typical usage characteristics and operational tendencies. As shown, there is an apparent spike in usage at approximately 6 a.m., indicating a scheduled HVAC event (i.e. precooling) by the building management system, followed by a sharp decrease, showing the tendency for the HVAC units to overshoot demand and drop in usage. The building then undergoes a gradual increase in use from growing occupancy (plug load, lighting, and HVAC) peaking at about 1 p.m. Note that during the summer in Texas, temperature will typically rise significantly throughout the morning and peak in the early afternoon, leading to a corresponding increased demand on air conditioning throughout that period (unless the building exhibits a significant thermal mass). The gradual fall from approximately 3 p.m. onward results from reaching an approximate steady state or even reduced HVAC operation, as well as from decreasing occupancy. Finally, three more significant signature drops occur at 5 p.m., 7 p.m., and 8 p.m. indicating more scheduled HVAC events, or "turn-off" times. These

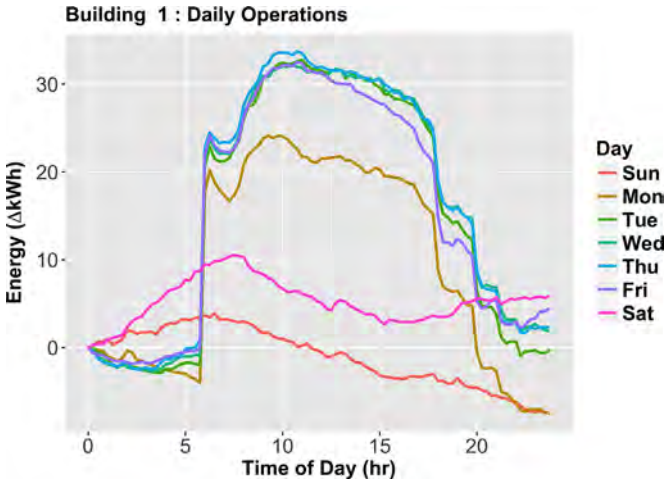

Fig. 4. Building 1 daily operational signatures for various days of the week using all two years of data.

conclusions were verified with the building energy manager. The classical time series decomposition approach was used to further examine building operational characteristics by subsetting the data into: day of the week, heating/cooling season, and high/low solar load datasets.

4.2. Day of the week

Fig. 4 presents the daily operational signature of Building 1, in kWh versus time of day, after using time series decomposition of two years of time series data and subsetting criteria for day of the week. The plot shows the obvious difference in the weekend and weekday lines, as well as the differences between the weekdays, particularly Monday and Friday. The decreased consumption on Fridays, compared to Tuesday-Thursday, in this office/lab can be attributed to reduced occupancy. Monday, on the other hand, is more difficult to analyze. The early morning hours of Monday show decreasing energy consumption - a characteristic seen in the later hours of Sunday. This decrease in early morning consumption shifts the entire Monday curve down throughout the rest of the day, misleading one to believe Monday consumes significantly less energy than other days of the week. If the timepoint for 0 kWh reference were moved to 6 a.m. one would see a much similar usage between Monday and the other weekdays (although choosing an alternative reference, such as 6 a.m., enhances this particular insight, choosing 12 a.m. was determined to give clearer and more consistent results among multiple buildings and subsetting analyses). However, the gap between Monday and the rest of the weekdays is slightly larger than the early morning difference at 6 a.m. This additional difference may be attributed to many overnight systems (e.g. large equipment, desktop computers) not yet in full operation after being shutdown for the weekend or a decreased occupancy of employees. These findings were confirmed with the building energy manager.

Fig. 5 shows the analysis of all six buildings to allow for further comparison and insight into energy consumption throughout the day as a function of the day of the week. Building 2 shows highest consumption on Tuesday, lowest weekday consumption on Friday (when referencing 4:45 a.m. as 0 kWh), and lowest weekend consumption on Sunday. Distinct spikes in energy usage can be identified visually in all buildings except Building 5, indicating scheduled building systems. Building 2 shows a distinct “turn-on” period at 4:45–5:30 a.m. and two distinct “turn-off” periods at 5:45 p.m. and 1 11 p.m. Similar turn-on/off times were identified for the other buildings (excluding Building 4). These spikes are later found to differ in magnitude between heating/cooling season operations discussed later in this paper, revealing them to be HVAC related. Most interesting is the weekend operational signature compared to the weekday operational signature. Buildings 1, 3, 4, and 6 display various turn-on and turn-off events during the weekdays, but not the weekends. In contrast, Building 2 does show similar HVAC turn-on and turn-off events even on weekends, although at a potentially different magnitude than on weekdays (while Building 5 does not show any turn-on/off events.) In other words, Building 2 operates its HVAC systems on weekends at a level that may not be required, leading to energy waste and a potential opportunity to save energy.

Since this analysis can also find the change in electricity usage at the distinct turn-on or turn-off times, it can reveal the potential size of an HVAC unit and its overall electricity consumption. Using this approach for Building 2, a statistically significant value of 30 kWh was consistently displayed over 15 min intervals, corresponding to a 120 kW unit. Considering the constant operation observed in the curve from 4:45 a.m. until 10:45 p.m. this piece of equipment consumes approximately 2160 kWh a day. Through the reduction of this HVAC event on the weekends, the building would save approximately 225 MWh, or about $27,000 annually (considering a $0.12/kWh rate [49]). Furthermore, rescheduling all of the turn-on/off events, weekday and weekend, to a 12 h schedule, such as 7:00 a.m to 7:00 p.m., would result in another 258 MWh, or $31,000 annually. The building energy manager was alerted to these potential savings in Building 2 and has adjusted operational characteristics to capture them.

Building 3 shows similar characteristics as Building 1 and 2, including scheduled HVAC events and lower weekend over weekday usage. The most interesting characteristic in this particular building is the consumption spike at 8:00 a.m. for a brief period on Tuesdays only. This is an uncharacteristic building consumption feature and may require the attention from the building manager. The event does not affect consumption significantly, but could pose a power demand issue or cycling fatigue of a piece of equipment considering the sharp increase and then decrease of the event.

Building 4 begins an analysis of the San Jose buildings, which operate in a milder climate and consist of office space only, compared to Richardson’s office and lab combined buildings. As shown in Fig. 5, Building 4 shows a much smoother operational signature with small scheduled HVAC events at 4:00 a.m. and 9:15 p.m., along with dips at 7:30 a.m. and increases at 7:00 p.m. These dips and increases were verified by the building manager to be the result of daylight sensing exterior lighting which turns on at sunset and off at sunrise. Specifically, exterior lighting events occur at differing times throughout the year, hindering the seasonal component and must be identified and disaggregated in the analyses for future study.

Building 5 displays an extremely smooth operational signature throughout each day of the week, indicating efficient building operation. Although, high usage into unoccupied hours of the day (as provided by the building manager) indicates a lack of unoccupied HVAC set point changes. Further, the absence of turn on/off does not indicate clearly whether the weekend days are operating at different interior temperature set points. In the other buildings, the turn-on/off events clearly displayed interior temperature set point changes and the absence of them during weekends showed the weekends were run in an unoccupied set point state. The results of Building 5 do not show this set point distinction and the building may only be operating at one interior set point. The addition of unoccupied set point states would drastically alter the consumption and save significant amounts of energy in nighttime/baseload and weekend operation.

The daily operational signature of Building 6 is unlike any of the other buildings in shape, but does show common turn-on/off fea-

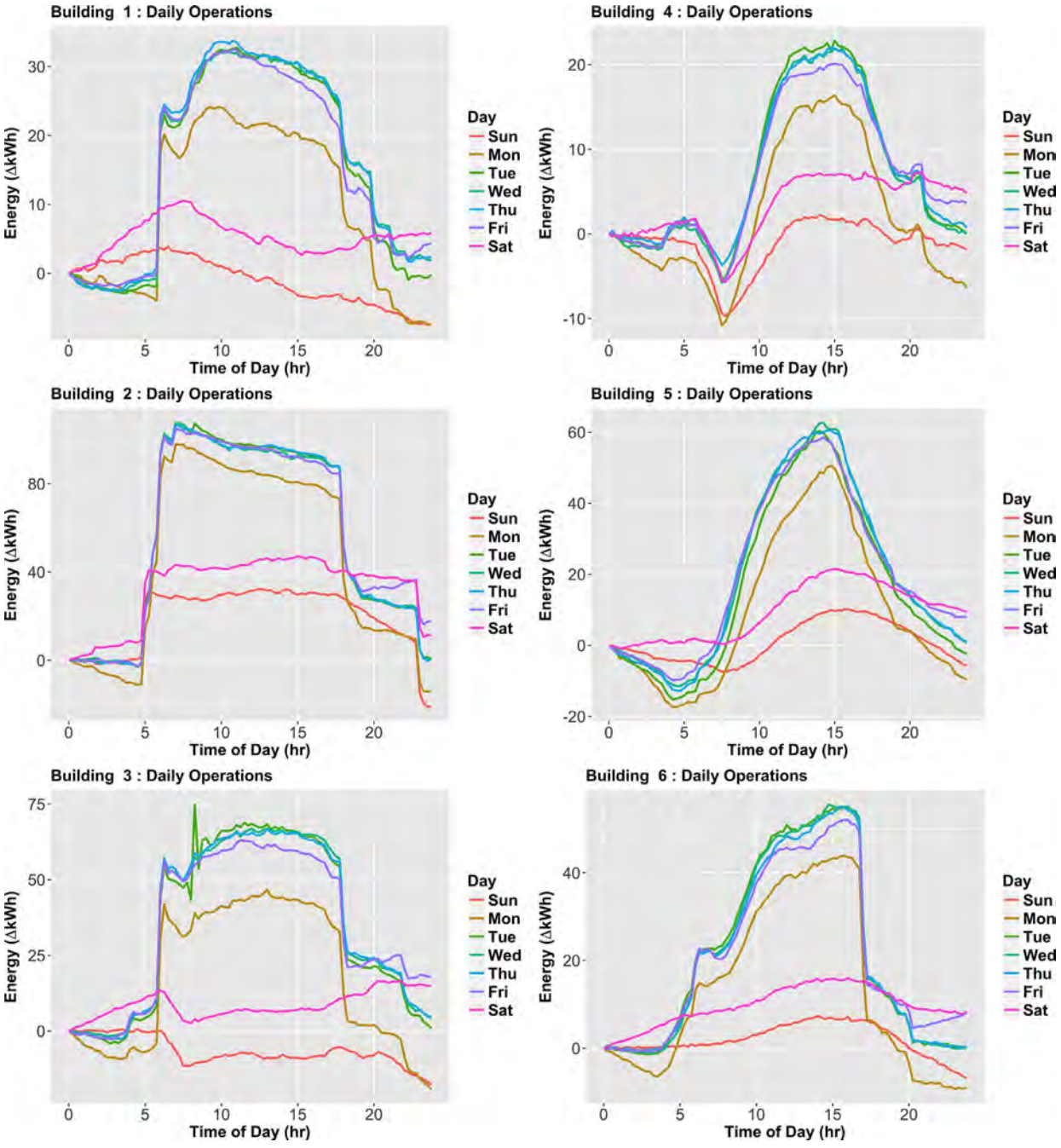

Fig. 5. Daily operational signatures for six buildings identifying operational characteristics for each day of the week using all two years of data.

tures. Gradual HVAC turn-on events are observed between 4:30–6:00 a.m., while a large turn-off is seen at 5:00 p.m., followed by a small turn-off at 8:00 p.m. The magnitude differences between the turn-on event and the large turn-off indicate that only some HVAC units are used during the morning for pre-conditioning, while other HVAC units turn on when needed during the morning ramp up. Then in the evening hours, most of these units turn off due to scheduling, followed by the final HVAC units at the smaller and later turn-off event. Additionally, it can easily be seen that weekend operations consume significantly less, revealing the change of set points for unoccupied states.

Considering the multiple occurrences of early start-up and late-shutdowns, a scheduling retrofit to a 12 h workday, such as 7:00 a.m.–7:00 p.m. is summarized in energy and monetary savings in Table 2 for all of the buildings. Of course such a schedule may not be suitable for all buildings—these results are shown as an example of how alternative scheduling can be assessed. These calculations were determined directly from the building energy data and not derived through any simulations or models and amount to over 700 MWh and $92,000 in savings per year for all six buildings. Although these savings are small compared to the total energy cost of the buildings, the cost to identify and perform a scheduling change would be on the order of $1,000. Therefore, the resulting ROI would be over 90, a return very attractive to a building manager.

Table 2
Energy and monetary savings for all six buildings if rescheduled major systems to a 7 a.m.–7 p.m. schedule. (Assumed electricity rates: RN, TX: $0.12/kWh and SJ, CA: $0.20/kWh [49]).

Building	Energy savings (*MWh/yr.*)	Monetary savings ($/*yr.*)
Building 1	40	4,800
Building 2	483	57,960
Building 3	125	15,000
Building 4	27	5,400
Building 5		
Building 6	47	9,400

Table 3
Error quantification for each building time series heating/cooling operation analysis.

Building	Heating +/− kWh	Cooling +/− kWh
1	0.98	1.98
2	5.51	2.57
3	2.06	1.86
4	2.31	1.27
5	2.25	1.90
6	1.72	1.08

4.3. Heating- and cooling-season operations

All six buildings are analyzed by heating and cooling seasons while also showing day of the week differences. By subsetting into heating and cooling seasons, as well as day of the week, for Buildings 1, 2, and 3 using classical time series decomposition results in findings displayed in Fig. 6. Fig. 7 continues the cooling/heating-season daily operational signature analysis for Buildings 4–6 located in San Jose, CA and Table 3 reports the error for each building classical time series decomposition subsetted by heating and cooling season.

Subsetting by cooling and heating season, Fig. 6 reveals unique insights into these seasonal operational signatures. First, all heating- and cooling-season daily operational signatures show drastically different shapes: cooling-season operations are smooth, rounded features and heating-season operations exhibit sharp features. These features are most likely due to three components of the building's thermal load: exterior temperature, solar irradiance, and occupancy effects (heat given off by people and equipment). In the heating season, the morning presents the coldest time of the day which requires the most heating, followed then by an increase in temperature, solar irradiance, and occupancy effects throughout the day. As these thermal loads increase, the necessary HVAC load is decreased. The opposite is true for the cooling season, the morning presents cooler temperatures resulting in decreased HVAC loads and then as thermal loads increase the HVAC load also increases. Exterior temperature, solar irradiance, and occupancy effects gradually increase as the day progresses and then gradually decrease in the afternoon to evening hours, which requires additional HVAC loads to cool the building. Peaks in the cooling-season months for all buildings occur at approximately 1:00-3:00pm, while heating-season peaks occur at 6 a.m. As expected, more energy is consumed during warmer times of the day (i.e. mid-afternoon) in cooling-season operation, and similarly less HVAC is used during the mid-afternoon hours during the heating season, also due in part to occupancy induced thermal load. The heating season peaks in the early morning because that is both when exterior temperatures are typically low and there is a corresponding increase in the thermostat setpoint. Even more, a second small peak occurs at 8:00 a.m. during the heating season, when morning arrivals likely lead to higher infiltration rates. Cooling-season operational signatures gradually approach these peaks (1:00 p.m.–3:00 p.m.), while heating-season operation shows a sudden jump in energy usage followed by a constant linear decline in usage throughout the rest of the day. This behavior strongly supports that these events are due to HVAC. This conclusion was verified with the building manager.

Building 2, unlike Buildings 1 and 3, shows a much more constant usage throughout the daytime hours in cooling-season months, but a more variable load during weekend operation. This is a result of Building 2 using scheduled HVAC units during the weekend unlike Buildings 1 and 3. For the same reason mentioned above, the exterior temperature and solar irradiance increase during day time hours and require additional HVAC usage, while the heating-season weekend for Building 2 slightly decreases throughout the days due to exterior temperature and solar irradiance. Buildings 1 and 3 do not have this behavior due to unoccupied set points maintained during weekend days.

Building 4 shows a difference in shape, in Fig. 7, having less blocky features in its cooling-season operational signatures, yet also more defined turn-on and turn-off HVAC events in its heating-season operations. The difference in shape is due to the use of electric HVAC in the cooling season and natural gas heating in the heating season. Therefore, the heating-season graph is much more jagged as occupancy load tendencies and the corresponding ventilation component of the building are dominant rather than electric heating needs. In contrast, the cooling-season operation is smooth as electric air conditioning masks the random nature of occupancy loads. Further analysis of the heating-season operation then allows for a better understanding of the pure occupancy load of a building. Here one sees there is a large difference between Friday's heating load and other weekdays consumption due to decreased occupancy on Fridays.

Building 5 shows very similar operational signatures in both heating and cooling seasons, except for an additional evening turn-on that may be necessary due to colder temperatures heading into night hours. This additional turn-on time provides a second indication that this building does not employ a comparably significant set point change in the building evening hours during the week (or a complete lack of set point change as described earlier). Considering that the building maintains one set point throughout all 24 weekday hours, it is necessary that a heating system must turn on during the late evening colder outdoor temperatures to maintain the set point. However, the other five buildings in this analysis do exhibit set point changes for unoccupied evening hours. By lowering their set points (e.g. 72 °F–66 °F), the thermal mass of the buildings prevent, on average, the need for the HVAC units to turn on before the set points return to occupied levels the next morning. Therefore, this evening turn-on event further indicates a lack of a weekday unoccupied set point change in Building 5. One additional finding is the relatively increased consumption on Fridays during heating operations. This indicates that decreased occupancy in Building 5 requires a significant amount more in heating load to offset the loss of employees who emit heat. Further, this shows the impact of an employee on HVAC is more significant than what that employee consumes for work purposes in Building 5. Finally, the other two San Jose buildings show increased consumption among heating operation Fridays when compared to cooling Fridays.

Building 5 also shows a similar behavior to that of Building 4, with jagged, occupancy-influenced operation during the heating season. In contrast to Buildings 4 and 5, Building 6 displays a much smoother heating-season operational signature than cooling season, ramping constantly through the morning in the heating season. This difference can be attributed to two reasons. First, the plug load and ventilation loads of Building 6 are much less variable than the other two buildings which is most likely due to the larger size and capacity of this particular building; greater numbers of

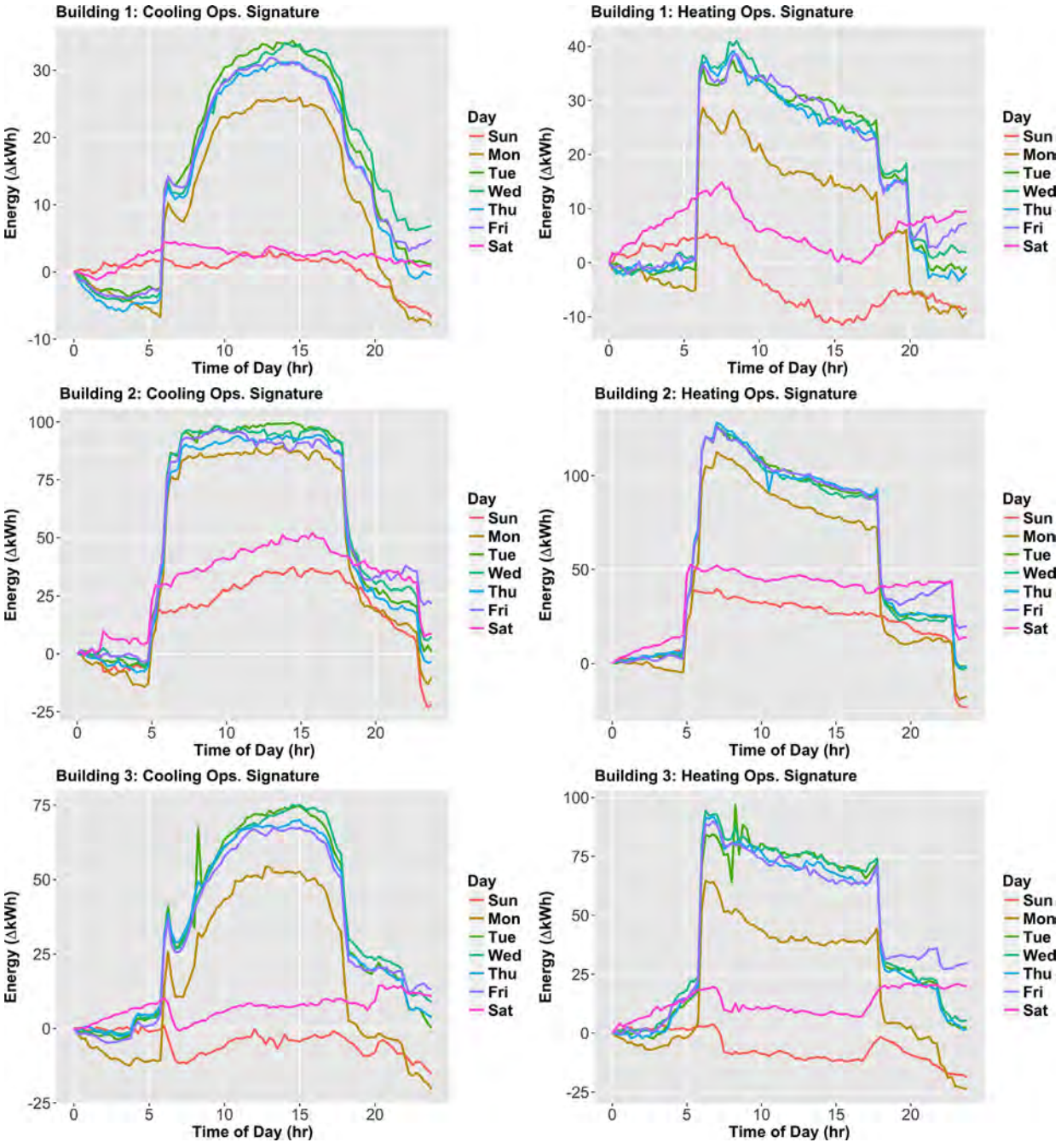

Fig. 6. Cooling- and heating-season daily operational signatures for Richardson, Texas buildings.

employees results in decreased impact from variations in occupancy. Second, the building does not engage all of its HVAC units in start up operation (as noted earlier) and the differences in times that the units turn on results in a more variable time series decomposition. The building also shows a differing late turn-off time from cooling to heating seasons, turning off at 8 p.m. in the cooling season and 7 p.m. in the heating season, which may be important to consider if changes in HVAC scheduling are opportunities for efficiency enhancements. Overall, each building shows a much higher magnitude of electricity consumption during the cooling-season months than what might be expected during the heating season given the corresponding winter temperatures in San Jose, CA. The lower than expected magnitude is an indicator that a non-electric heating system may be in use; this was verified by the building manager.

4.4. Solar thermal load in cold heating-season operation

Fig. 8 shows Building 1 daily operational signatures with varying daily solar irradiance on the left ($\mu = \pm 2.82$ kWh for sunny, $\mu = \pm 2.99$ kWh for cloudy) and the associated average solar irradiance (W/m^2) vs. time of day. Mean trend values are 408 ± 18 kWh and 415 ± 34 kWh for sunny and cloudy days, respectively (all values listed in Table 4).

To consider the impact of the solar thermal load on heating-season behavior, consider Fig. 8 which shows that during cloudy

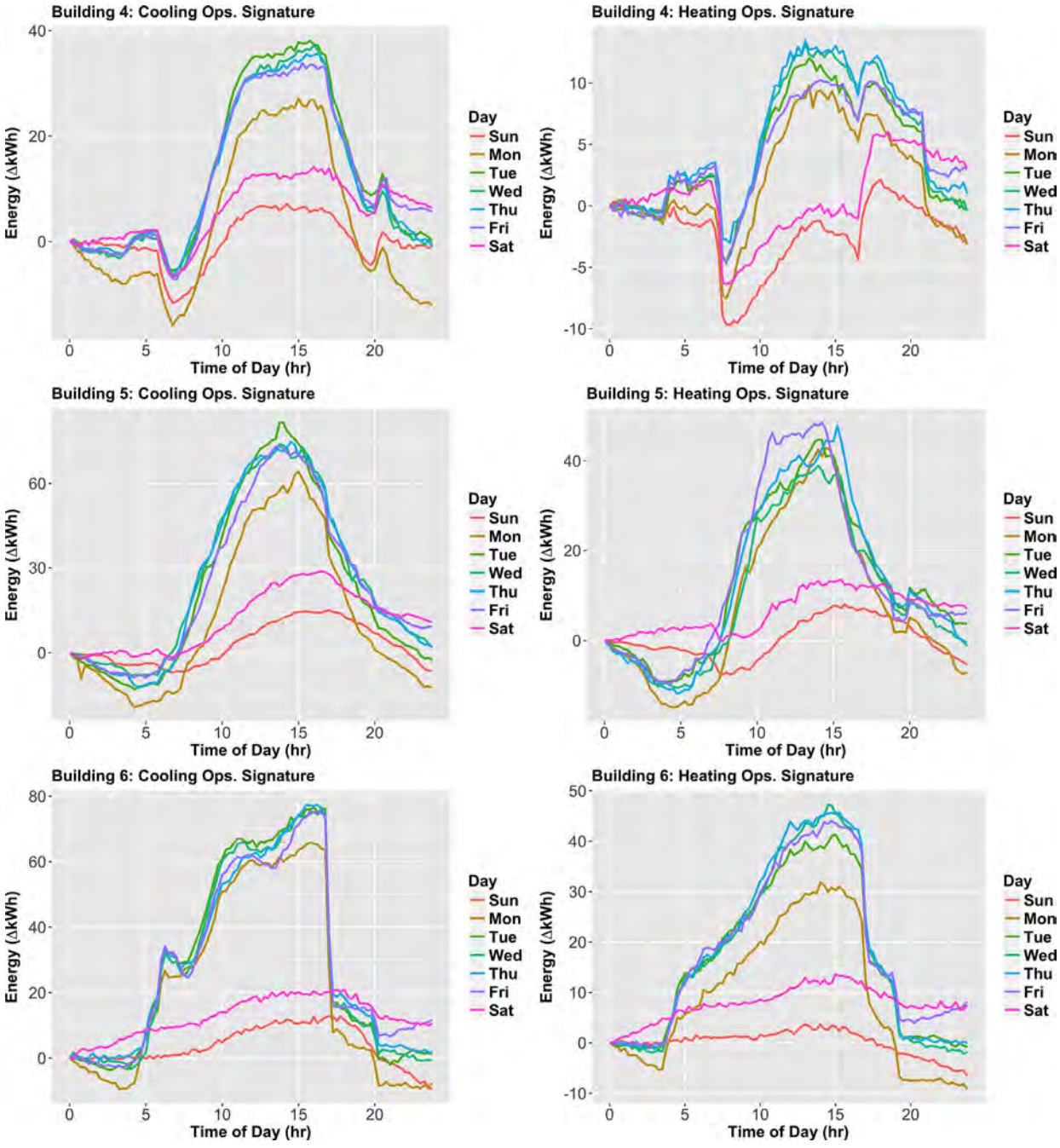

Fig. 7. Cooling- and heating-season daily operational signatures for San Jose, California buildings.

Table 4
Solar irradiance analysis values: sunny/cloudy, mean irradiance (W/m^2), daily operation error (kWh), and trend value (kWh).

Cold operation (< 47 °F)	Sunny operation	Cloudy operation
Mean irradiance	$> 210 W/m^2$	$< 70 W/m^2$
Daily operation error	±2.82 kWh	±2.99 kWh
Trend value	408 ± 18 kWh	415 ± 34 kWh

days, much more overall energy is required to maintain the set point. At the same time on sunny days, in a heating season, one can observe a reduction in consumption throughout the day as the solar irradiance assists by contributing to heating. Computing this difference between the signatures show cloudy days consume 675 kWh ($81) more than equally cold sunny days. Also, the lower morning usage for cloudy days and higher morning load for sunny days shows the effect of night cloud cover maintaining higher temperatures as compared to clear nights. Overall, these observations give insights into a building's solar-thermal load, which may point to opportunities for energy savings by using window shading accordingly during sunny days or a more optimized HVAC scheduling that accounts for predicted solar insolation patterns.

4.5. Error of the classical time series decomposition results

Now we discuss the error of the building classical time series decompositions presented in the heating- and cooling-season and

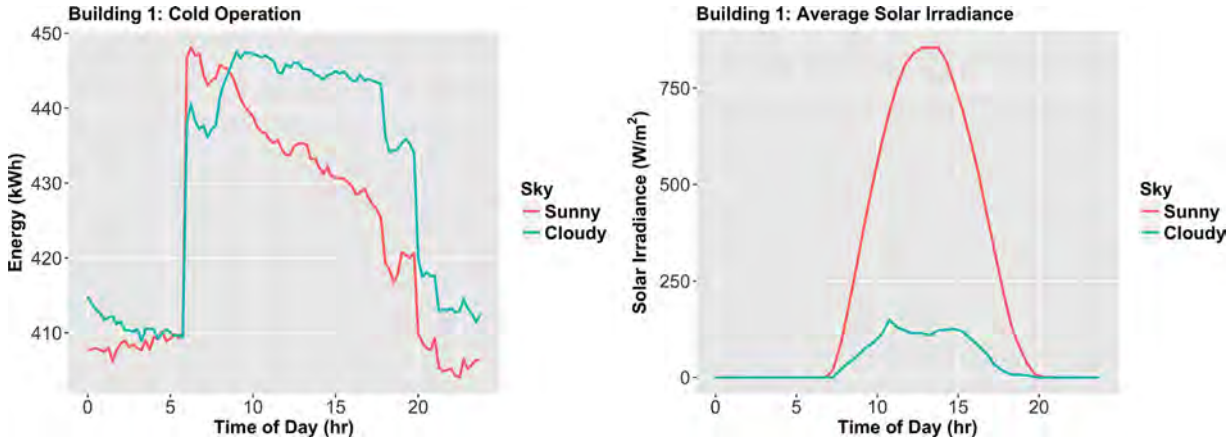

Fig. 8. Operation signatures of cold heating season days in cloudy and sunny weather for Building 1 (left) ($\mu = \pm 2.82$ kWh for sunny, $\mu = \pm 2.99$ kWh for cloudy) and average solar irradiance for sunny and cloudy days (right).

solar irradiance analyses. Errors in the electricity values are calculated using the random component of the decomposition. In most cases this error is less than 5 kWh and the mean error among all heating- and cooling-season results was 2.11 kWh. Considering the magnitude of building operational signatures are on the order of 100 kWh gives confidence that the insights of each signature are not the result of noise in the data. Additionally, those buildings with higher errors also show relatively higher magnitudes in operation. Therefore, the relative error is still low at $< 2.5\%$. These analyses are computed using two years of data and result in n values of approximately 30 days after subsetting. If the analyses were computed on longer time series datasets, such as 4 years of data, errors may be further minimized. However, one must be cautious is analyzing larger time intervals as buildings are dynamical systems that may undergo significant operational changes or degrade in performance over time. Care must first be taken by analyzing the data in smaller segments to determine if the any major change has occured before analyzing full datasets.

5. Conclusions

The building operational signatures presented here provide a novel method to capture significant insights about each building into a few curves. For this study, we demonstrated the ability to uncover distinct building operational signatures through classical decomposition and subsetting of days of the week, heating and cooling seasons, and solar irradiance. Classical decomposition quickly showed large operational differences between office/lab buildings (i.e. between RN,TX labs and SJ, CA offices) and produced overall HVAC scheduling tendencies. Through a simple rescheduling of these HVAC systems across all six buildings suggest savings of over 700 MWh and $92,000 per year at virtually no retrofitting cost to the building owner. Even though these savings are small compared to overall energy costs, these types of low-cost/high-reward actions are attractive to building managers implementing lean management practices or those who have insufficient retrofit budgets.

The days of the week operational analysis brought insight upon occupancy levels throughout the work week. For all buildings, maximum occupancy was determined to occur on Tuesday, while minimum occupancy was observed on Mondays or Fridays. The day of the week analysis was also able to uncover a malfunctioning system occurring strictly on Tuesdays in Building 3 and suggested a lack of unoccupied set point changes in Building 5. The heating- and cooling-season analysis further supported this latter finding, as it displayed HVAC operation during the late evening hours, unlike the other five buildings. The analysis also pointed to significant operation differences as the RN, TX buildings shifted peak usage from 1 to 3 p.m. in cooling season, to 6 a.m. in heating season. This building also exhibited usage increases at 8 a.m. in heating season due to infiltration of cold air as employees arrived for work.

Finally, solar irradiance impacts upon Building 1 were investigated. Through analysis of cold heating-season days with both sunny and cloudy conditions, Building 1 was found to use an additional 675 kWh of electricity ($81) during cold cloudy days when compared to a cold sunny day. This result shows the potential of including solar irradiance within models and provides a method of analysis to better understand the complexities of solar-thermal loads in buildings.

Currently, this analysis is limited by two years of time series data. When subsetting datasets, the data under analysis becomes smaller with each new restriction. Therefore, a limited number of subsetting criteria may be used before errors become large. In this study, the analysis of solar irradiance on cold heating seasons were subject to the most subsetting constraints, yet still reported a significantly low mean error of ± 2.90 kWh. Future work on larger datasets may improve results as long as these analyses ensure significant changes to the building have not occurred over that larger time span. If longer building datasets are properly accounted for, additional subsetting analyses and insights may be found (e.g. improved solar irradiance effect on consumption).

Classical time series decomposition of 15 min building electricity consumption data provides quick and useful insights on how buildings operate. Through classical decomposition, a quite extensive understanding of a building's operation can be easily achieved. Time series decomposition can be further refined for specific cases, such as weather, day of week, month, solar irradiance, etc., provided the data is available. Due to the ease of classical time series decomposition analysis, these results can be obtained virtually and instantaneously and provide quick insights to evaluate the need for or complement a full building energy audit. Time series operational signatures indicate early start-up/shut-down times for building systems (HVAC), overall occupancy trends of employees, how HVAC units respond to particular weather conditions, and other insights useful to a building manager that have been previously overlooked. Building operational signatures could also be used as a means for analyzing a large number of building datasets to quickly pinpoint buildings of certain characteristics, specific occupancy ten-

dencies, and inefficiencies among thousands. This method of identifying particular building characteristics further opens the door to building energy disaggregation, arguably the most pivotal barrier to building energy efficiency improvements [50]. Finally, since this analysis can be automated to run upon ingestion of whole building electricity data with very few additional details about the building, along with a near-instantaneous report on potential energy savings and building operational characteristics, it provides a unique alternative to other more cumbersome approaches.

6. Funding

This research was performed in the SDLE Research Center, established with Ohio Third Frontier funding under award Tech 11-060, Tech 12-004, and the Great Lakes Energy Institute, both at Case Western Reserve University. The information, data, or work presented herein was also funded in part by the Advanced Research Projects Agency-Energy (ARPA-E), U.S. Department of Energy, under Award Number DE-AR-0000668.

References

[1] U.S. Briefing, International Energy Outlook 2013, US Energy Information Administration, 2013. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.369.6789&rep=rep1&type=pdf.
[2] D. Austin, Addressing Market Barriers to Energy Efficiency in Buildings, Congressional Budget Office, 2012. http://cbo.gov/sites/default/files/cbofiles/attachments/AddressingMarketBarriersToEnergyEfficiencyInBuildings_WorkingPaper_2012-10.pdf.
[3] Financing Small Commercial Building Energy Performance Upgrades: Challenges and OpportunitiesNational Institute of Building Sciences, Council on Finance, Insurance and Real Estate. https://c.ymcdn.com/sites/www.nibs.org/resource/resmgr/CC/CFIRE_CommBldgFinance-Final.pdf.
[4] Better Buildings Challenge, United States Department of Energy, Building Technologies Office, 2018. http://energy.gov/eere/efficiency/buildings.
[5] N. Aste, C.D. Pero, Energy retrofit of commercial buildings: case study and applied methodology, Energy Effic. 6 (2) (2013) 407–423, doi:10.1007/s12053-012-9168-4.
[6] B. Shen, L. Price, H. Lu, Energy audit practices in china: national and local experiences and issues, Energy Policy 46 (2012) 346–358, doi:10.1016/j.enpol.2012.03.069.
[7] P.D.G. PhD, A case study of multiple energy audits of the same building: conclusions and recommendations, ASHRAE Trans. 120 (2014). GG1, http://search.proquest.com/openview/b599a9420d997ee048f655f3a12e44b5/1?pq-origsite=gscholar.
[8] S. Chirarattananon, J. Taweekun, A technical review of energy conservation programs for commercial and government buildings in thailand, Energy Convers. Manage. 44 (5) (2003) 743–762, doi:10.1016/S0196-8904(02)00082-1.
[9] J. Harris, J. Anderson, W. Shafron, Investment in energy efficiency: a survey of australian firms, Energy Policy 28 (12) (2000) 867–876, doi:10.1016/S0301-4215(00)00075-6.
[10] D.B. Crawley, J.W. Hand, M. Kummert, B.T. Griffith, Contrasting the capabilities of building energy performance simulation programs, Build. Environ. 43 (4) (2008) 661–673, doi:10.1016/j.buildenv.2006.10.027.
[11] P.G. Loutzenhiser, H. Manz, C. Felsmann, P.A. Strachan, T. Frank, G.M. Maxwell, Empirical validation of models to compute solar irradiance on inclined surfaces for building energy simulation, Sol. Energy 81 (2) (2007) 254–267, doi:10.1016/j.solener.2006.03.009.
[12] M. Mirsadeghi, D. Costola, B. Blocken, J.L.M. Hensen, Review of external convective heat transfer coefficient models in building energy simulation programs: implementation and uncertainty, Appl. Therm. Eng. 56 (1) (2013) 134–151, doi:10.1016/j.app1thermaleng.2013.03.003.
[13] J.D. Rhodes, W.H. Gorman, C.R. Upshaw, M.E. Webber, Using BEopt (EnergyPlus) with energy audits and surveys to predict actual residential energy usage, Energy Build. 86 (2015) 808–816, doi:10.1016/j.enbuild.2014.10.076.
[14] T.A. Reddy, I. Maor, C. Panjapornpon, Calibrating detailed building energy simulation programs with measured data - part 1: General methodology (RP-1051), HVAC&R Res. 13 (2) (2007) 221–241, doi:10.1080/10789669.2007.10390952.
[15] N. Fumo, A review on the basics of building energy estimation, Renew. Sustain. Energy Rev. 31 (2014) 53–60, doi:10.1016/j.rser.2013.11.040.
[16] N. Li, Z. Yang, B. Becerik-Gerber, C. Tang, N. Chen, Why is the reliability of building simulation limited as a tool for evaluating energy conservation measures? Appl. Energy 159 (2015) 196–205, doi:10.1016/j.apenergy.2015.09.001.
[17] H.x. Zhao, F. Magoules, A review on the prediction of building energy consumption, Renew. Sustain. Energy Rev. 16 (6) (2012) 3586–3592, doi:10.1016/j.rser.2012.02.049.
[18] E.M. Pickering, M.A. Hossain, J.P. Mousseau, R.A. Swanson, R.H. French, A.R. Abramson, A cross-sectional study of the temporal evolution of electricity consumption of six commercial buildings, PLoS ONE 12 (10) (2017). E0187129 doi: 10.1371/journal.pone.0187129.
[19] ARPA-e OPEN 2015 Project Selections, U.S. Department of Energy. URL http://arpa-e.energy.gov/sites/default/files/documents/files/OPEN_2015_Project_Descriptions.pdf.
[20] M.F. Fels, PRISM: an introduction, Energy Build. 9 (1) (1986) 5–18, doi:10.1016/0378-7788(86)90003-4.
[21] Hung, A. (2015). Energy Efficiency and Statistical Analysis of Buildings. Master's thesisCase Western Reserve University.
[22] Pickering, E. M. 2016. Edifes 0.4: Scalable Data Analytics for Commercial Building Virtual Energy Audits. Master's thesisCase Western Reserve University. (8 2016).
[23] J.K. Kissock, C. Eger, Understanding Industrial Energy Use Through Sliding Regression Analysis, 2007. https://www.udayton.edu/engineering/centers/industrial_assessment/resources/docs/pdf/UnderstandingIndEnergySlidingRegAnal_ACEEE_2007.pdf.
[24] International Performance Measurement & Verification Protocol: Concepts and Options for Determining Energy and Water Savings, Vol. 1, 2002. http://www.nrel.gov/docs/fy02osti/31505.pdf.
[25] J.S. Haberl, ASHRAE'S Guideline 14-2002 for measurement of energy and demand savings: how to determine what was really saved by the retrofit, in: Proceedings of the Fifth International Conference for Enhanced Building Operations, 2005. http://esl.tamu.edu/docs/terp/2005/esl-ic-05-10-50.pdf.
[26] C. Miller, Z. Nagy, A. Schlueter, Automated daily pattern filtering of measured building performance data, Autom. Constr. 49 (2015) 1–17, doi:10.1016/j.autcon.2014.09.004.
[27] I. Khan, A. Capozzoli, S.P. Corgnati, T. Cerquitelli, Fault detection analysis of building energy consumption using data mining techniques, Energy Procedia 42 (2013) 557–566, doi:10.1016/j.egypro.2013.11.057.
[28] J.S. Chou, A.S. Telaga, Real-time detection of anomalous power consumption, Renew. Sustain. Energy Rev. 33 (2014-05) 400–411, doi:10.1016/j.rser.2014.01.088.
[29] H. Janetzko, F. Stoffel, S. Mittelstaedt, D.A. Keim, Anomaly detection for visual analytics of power consumption data, Comput. Graph. 38 (2014) 27–37, doi:10.1016/j.cag.2013.10.006.
[30] T. Hong, M.A. Piette, Y. Chen, S.H. Lee, S.C. Taylor-Lange, R. Zhang, K. Sun, P. Price, Commercial building energy saver: an energy retrofit analysis toolkit, Appl. Energy 159 (2015) 298–309, doi:10.1016/j.apenergy.2015.09.002.
[31] O.T. Masoso, L.J. Grobler, The dark side of occupants' behaviour on building energy use, Energy Build. 42 (2) (2010) 173–177, doi:10.1016/j.enbuild.2009.08.009.
[32] R.H. Shumway, D.S. Stoffer, Time series analysis and its applications: with r examples, forth ed., Springer, New York, NY, 2017.
[33] R. Barras, D. Ferguson, A spectral analysis of building cycles in britain, Environ. Plann. A 17 (10) (1985) 1369–1391.
[34] L.H. Koopmans, The Spectral Analysis of Time Series, Academic Press, 1995.
[35] S. Beveridge, C. Nelson, A new approach to decomposition of economic time-series into permanent and transitory components with particular attention to measurement of the business-cycle, J. Monet. Econ. 7 (2) (1981) 151–174, doi:10.1016/0304-3932(81)90040-4.
[36] C. Bretherton, C. Smith, J. Wallace, An intercomparison of methods for finding coupled patterns in climate data, J. Clim. 5 (6) (1992) 541–560, doi:10.1175/1520-0442(1992)005⟨0541:AIOMFF⟩2.0.CO;2.
[37] B. Ang, Decomposition methodology in industrial energy demand analysis, Energy 20 (11) (1995-11) 1081–1095, doi:10.1016/0360-5442(95)00068-R.
[38] A.M.M. Masih, R. Masih, Energy consumption, real income and temporal causality: results from a multi-country study based on cointegration and error-correction modelling techniques, Energy Econ. 18 (3) (1996) 165–183, doi:10.1016/0140-9883(96)00009-6.
[39] L. Ghelardoni, A. Ghio, D. Anguita, Energy load forecasting using empirical mode decomposition and support vector regression, IEEE Trans. Smart Grid 4 (1) (2013) 549–556. WOS:000325485600059 doi: 10.1109/TSG.2012.2235089.
[40] A. Tascikaraoglu, B.M. Sanandaji, Short-term residential electric load forecasting: a compressive spatio-temporal approach, Energy Build 111 (2016) 380–392. WOS:000369191100034 10.1016/j.enbuild.2015.11.068.
[41] F. Rubel, M. Kottek, Comments on: the thermal zones of the earth by Wladimir Kppen (1884), Meteorol. Z. 20 (3) (2011) 361–365, doi:10.1127/0941-2948/2011/0258.
[42] F. Rubel, K. Brugger, K. Haslinger, & I. Auer, 2017. The climate of the european alps: shift of very high resolution Kppen-Geiger climate zones 1800–2100Meteorologische Zeitschrift. doi:10.1127/metz/2016/0816.
[43] C. Bryant, N.R. Wheeler, F. Rubel, R.H. French, Koeppen-Geiger Climatic Zones, 2017. (Nov.2017) https://cran.r-project.org/web/packages/kgc/index.html.
[44] Noaa Climatic Data Online - Application Router (2016). URL http://www7.ncdc.noaa.gov/CDO/cdopoemain.cmd?datasetabbv=DS3505&countryabbv=&georegionabbv=&resolution=40.
[45] R. Perez, T. Cebecauer, M. ri, Semi-empirical satellite models, in: Solar Energy Forecasting and Resource Assessment, Academic Press, Boston, 2013, pp. 21–48. http://www.sciencedirect.com/science/article/pii/B9780123971777000024.
[46] T. Cebecauer, M. Suri, V. Rajpaul, C. Richter, Site-adaptation of satellite-based DNI and GHI time series: overview and SolarGIS approach, AIP Conf Proc 1734 (1) (2016) 150002, doi:10.1063/1.4949234.
[47] A.V. Metcalfe, P.S. Cowpertwait, Introductory Time Series with R, Springer, New York, 2009. http://link.springer.com/10.1007/978-0-387-88698-5.

[48] H. Manz, P. Loutzenhiser, T. Frank, P.A. Strachan, R. Bundi, G. Maxwell, Series of experiments for empirical validation of solar gain modeling in building energy simulation codes - experimental setup, test cell characterization, specifications and uncertainty analysis, Build Environ. 41 (12) (2006) 1784–1797, doi:10.1016/j.buildenv.2005.07.020.

[49] Average Energy Prices for the U.S. and Selected Metropolitan Areas. http://www.bls.gov/regions/midwest/data/averageenergyprices_selectedareas_table.htm.

[50] K.C. Armel, A. Gupta, G. Shrimali, A. Albert, Is disaggregation the holy grail of energy efficiency? The case of electricity, Energy Policy 52 (2013) 213–234, doi:10.1016/j.enpol.2012.08.062.