9-10-2021

A Case Study to Quantify Variability in Building Load Profiles

ANDREW PARKER

SAM MOAYEDI

KEVIN JAMES

Dongming Peng

Mahmoud Alahmad

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received July 17, 2021, accepted August 29, 2021, date of publication September 10, 2021, date of current version September 21, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3112103

A Case Study to Quantify Variability in Building Load Profiles

ANDREW PARKER[1], SAM MOAYEDI[2], KEVIN JAMES[2], DONGMING PENG[3], AND MAHMOUD A. ALAHMAD[1,2], (Senior Member, IEEE)

[1]National Renewable Energy Laboratory, Golden, CO 80401, USA
[2]Durham School of Architectural Engineering and Construction, University of Nebraska–Lincoln, Omaha, NE 68182, USA
[3]Department of Electrical and Computer Engineering, University of Nebraska–Lincoln, Lincoln, NE 68588, USA

Corresponding author: Mahmoud A. Alahmad (malahmad2@unl.edu)

This work was supported 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 DE-AC36-08GO28308, and in part by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Building Technologies Office.

ABSTRACT Recent technology development and penetration of advanced metering infrastructure (AMI), advanced building control systems, and the internet-of-things (IoT) in the built environment are providing detailed information on building operation, performance, and user's comfort and behavior. Building owners can obtain a wide range of energy consumption details at various levels of time granularity to augment their decisions as they manage the building operation and interact with the grid. AMI data are providing a new level of detail and visibility that may enhance building services and assets in the smart grid domain and make buildings inch closer to becoming a grid-interactive energy efficient buildings (GEB). While utility-installed AMI typically records energy consumption at a 15, 30, or 60-minute resolution, building-owner-installed metering can record energy consumption at one-minute or sub-minute time scales, providing information about how much the energy consumption varies from one sub-minute to the other (i.e. variability) at a finer time resolutions than typically available from AMI. This paper examines one-minute building load profile data sets and presents a framework to study, define, extract, quantify and analyze variability in buildings' load profiles. The discussion of variability and its analysis is based on a case study of an actual sub-minute time-resolution data set, collected in 2019, for two buildings in a Midwest state in the USA. The result shows that for the case studies, the level of variability in an end-use category is not simply proportional to its consumption. Furthermore, distinct and predictable daily variability patterns emerge in end-use load categories. This information is useful for a host of applications including prediction, forecasting, and modeling.

INDEX TERMS Load profile, building energy modeling, discrete wavelet transform, empirical mode decomposition, variability analysis, variational mode decomposition.

I. INTRODUCTION

As the number of electrical loads in the built environment continues to grow and the penetration of distributed energy resources (DERs) continues to increase, their voltage, frequency and power consumption/generation fluctuations in real time must be considered for thorough analysis. Specifically, accurate power flow studies in the electrical grid must consider these fluctuations to guarantee electrical system flexibility, reliability, and operation without compromising system performance. By increasing the utilization of DERs, electrical system operation, control, economy, and planning studies become critical to meet the target DER penetrations [1]. When DER is interconnected to the distribution system, conventional distribution studies and special system impact studies methodologies need to be considered as explained in IEEE Std 1547.7 [1], [2]. Hence, correct modeling of the distribution systems' components is key to conducting accurate studies, which require time-series data of the relevant loads at varying temporal and spatial conditions [1]. IEEE Std 1547.7 (2013) states that "Multiple power flow simulations including time series or quasi-static

The associate editor coordinating the review of this manuscript and approving it for publication was Yeliz Karaca.

simulations may be needed to fully study the impacts of DR [distributed resources] on Area EPS [Electric Power System] voltage'' ([2], page 8). Furthermore, the IEEE standard describes quasi-static simulations: ''Quasi-static simulation refers to a sequence of steady-state power flow conducted at a timestep of no less than 1 second but that can use a time step of up to one hour'' ([2], page 77) [1]. For example, authors in [3] studied the impact of PV systems on electrical systems using quasi-static time-series (QSTS) simulations. The QSTS simulations in this work require the use of model data, historical time-series load data and PV data, however, time-series load data is not always available at the desired temporal resolutions, often limited to a 15-minute, 30-minute, or 1-hour time resolution based on the typical recording frequencies of utility-installed advanced metering infrastructure (AMI). Authors emphasized that such low-resolution data is not sufficient for some of the time dependent QSTS simulations, and access to the high-resolution data is preferred when conducting such studies.

Building load profiles are time-series load data that varies over a sub-minute, a minute, an hour, a day, a week, a month or a year [4], [5]. Recent technology development and penetration of advanced metering infrastructure, advanced building control systems, sensor networks, and the internet-of-things (IoT) in the built environment are providing detailed information on building loads, operation, performance, and user behavior. By measuring load profiles and other sensor related information at a high temporal resolution, precise knowledge and near real-time energy consumption information is achievable. With respect to energy consumption, new metering systems are providing a new level of detail and visibility that may enhance building services and assets in the smart grid domain and making building inch closer to becoming a grid-interactive energy efficient buildings (GEB) [24]. Different metering systems are able to provide different levels of detail, from a single reading of total energy consumption at the whole-building level, to an intermediate level of consumption recording each end-use load category, to fine-grained details recording individual load components. In addition, the time-resolution of the measured data varies from sub-minute to hourly data. These various levels of load profile details provide critical spatio-temporal energy demand information that can be used to inform energy efficiency measures among other applications [6] and can be analyzed in time or spectral domains to provide information about what is actually taking place in a building [7]. Furthermore, these variability characteristics and varying levels of time-series load profile data can provide more information on the impact of customers on electricity demand, which can be used for various power distribution system analysis such as DER planning, tools for short-term load forecasting, synthetic load profile development, and demand-side management. For example, the authors in [8] discussed the impact of variability on the power grid. They evaluated 17 million daily load samples to extract consumption variability at the individual and neighborhood level. They used a one-hour timestamp for their analysis.

Building load profile time series data sets are limited, as most building owners do not share their consumption in an open forum and load profiles at a resolution less than 15-minutes are scarce. To overcome the lack of load profile and load profile variability characteristics data, researchers rely on available aggregate historical data for a given distribution system and use common fixed fractions to assign loads at the transformer/building level. In doing so, the variability in the data is lost [4]. To overcome this limitation, the authors proposed a generative model to generate synthetic profiles. In addition to this method, Building Energy Modeling tools (BEM) are a great resource to develop building load profiles for applications in distribution system analysis, energy efficiency and so forth. Simulation software such as EnergyPlus was developed to provide a means for modeling building energy consumption in detail. EnergyPlus is able to generate modeled load profile data, based on set schedules and parameters, at various timescales, ranging from one-minute to one hour. However, the developed load profiles contain less variability than what actually takes place in real buildings [1]. In other ongoing work by the authors, one application of variability load profiles is utilized to generate synthetic load profiles that resemble real building load profiles [1]. By focusing on variability in this context, the interest is focused on extracting variability from a subset of the limitedly available buildings with high time-resolution measurements and leveraging these detailed profiles to predict variability in similar load profiles with lower time- resolution measurements, to generate synthetic load profiles for a wide range of applications.

Toward this end, this paper builds on ongoing work, focuses on the variability profile, looks at methods to extract this variability from measured load profile data, and investigates the impact of end-use category on building variability. Specifically, the paper presents a method to define variability in building load profile data; methods to extract the variability using decomposition techniques; metrics to quantify the extracted variability; and tools to analyze variability in both full buildings and the end-use load categories that comprise the total data. Figure 1 shows the framework used in this research. It is noteworthy to mention that the methodology presented here is a first step in understanding the typical variability found in measured building load profiles in order to leverage these characteristics to generate more realistic or accurate synthetic load profiles from models or measured load profiles with lower time resolution.

The major contribution and novelty of this work focuses on the development of a framework methodology that can be applied to modeled and/or measured time series data. Specifically, the contributions are:

1. Developing a definition for variability in high-resolution time series data. This definition does not apply to building load profile data only, but can be used for any time series data in different fields.

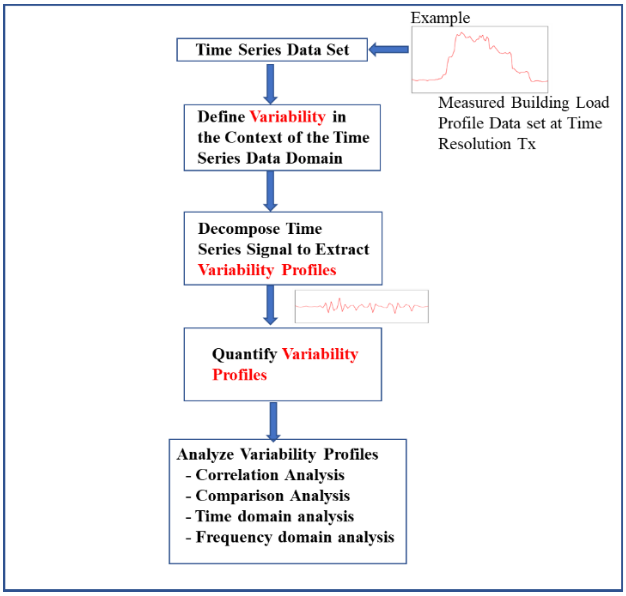


FIGURE 1. Overall framework of the methodology in this work.

2. Extracting various levels of variability profiles using well established decomposition techniques. The level of variability to extract is application- specific and will differ for different time series data domains. While most published work on decomposition discards the variability profile and focuses on the low-resolution signal, this work focuses on analyzing the variability profile instead.
3. Quantifying the extracted variability and developing metrics specific to building load profile data to:
   a. Measure and contrast variability profiles in similar time series data sets.
   b. Measure the impact of top-down and bottom-up monitoring by contrasting aggregated building load profile data with intermediate end-use load category load profile data.
   c. Identify relationships between consumption percentages and variability levels.

II. VARIABILITY IN BUILDING LOAD PROFILES

A. DEFINING VARIABILITY

Energy measurements and predictions only tell us the average power demand over some time frame. Depending on the time frame used, this average might significantly overpredict/underpredict actual energy demand in shorter time frames. For example, Figure 2 shows that the daily averages are clearly not a good prediction of hourly energy. A single, constant prediction vastly underpredicts peak demand, while overpredicting overnight demand. This daily measurement/prediction would be very wrong at almost all times, because the measurement has averaged two very different modes of building operation in terms of energy consumption.

The same is true for the one-minute and one-hour time scales. The one-hour averages may not be accurate predictions of the energy at the one-minute interval as shown in Figure 3. Specifically, when using one-hour averages, accuracy is lost from the peak demand assumptions, and information is lost about how much the energy varies over shorter time frames.

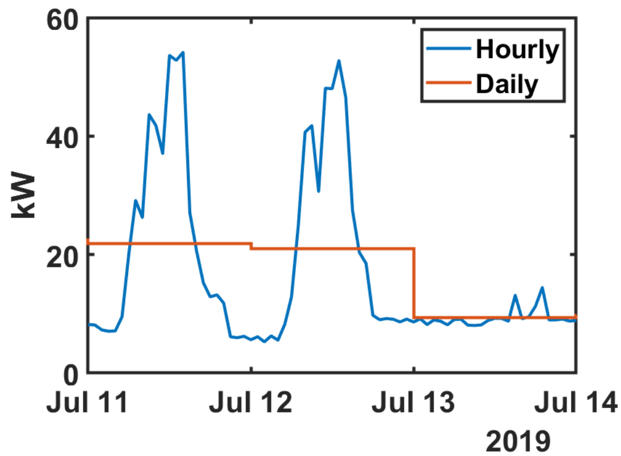


FIGURE 2. Hourly and aggregated daily load profile measurements for an actual building on July 11-14, 2019.

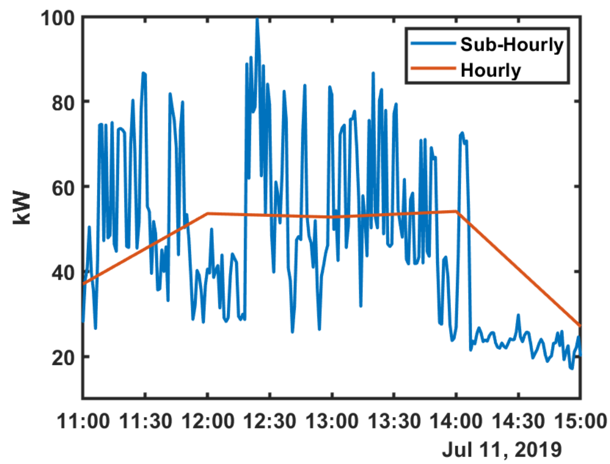


FIGURE 3. One-minute measurements and aggregated hourly measurements for an actual building on July 11, 2019.

Based on Figures 2 and 3, the exact definition of building load profile variability is thus flexible, and depends on the information desired and the information available. The desired information seeks to identify what timescale is needed in order to conduct meaningful analysis of a given building energy load profile, whereas the available information is constrained by the actual time scale of measured or modeled data for the building(s) under consideration. A load profile's variability over a certain timescale is the information that fills this gap. This variability may be affected by equipment, occupancy behavior, building size, or any number of factors. Thus, the load profile variability is a function of the building under consideration and the measured/desired timescale (time resolution). This can be described by equation 1:

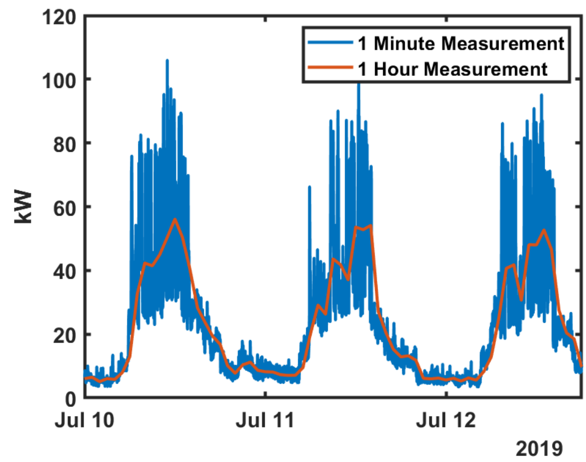
$$\text{Load Profile Variability} = f(\text{Building characteristics}, \text{measured timescale}, \text{desired timescale}) \quad (1)$$

Furthermore, because of this flexibility, a general definition for load profile variability can be expressed as a quantifiable quantity that is bounded (i.e., with confidence), or represented as a distribution, quantifying the range of possible energy values that exist at times between two different time-resolutions. In another way, variability can be described as the difference between a given measured, modeled or predicted load profile for a given building and a lower-resolution version of that same load profile. The lower-resolution version of this load profile will be referred to as the 'base load'. Figure 4 illustrates this definition; the blue line is the desired timescale, the high-resolution load profile (measured at the one-minute time scale), and the red line is the base load, the lower resolution version of this load profile (measured/averaged at the one-hour time scale). The difference between these two lines is the variability load profile that is missing in the one-hour time-resolution scale data and present in the one-minute time-resolution scale data. While the discussion is focusing on measured building load profiles, the same analysis/conclusions apply to modeled, predicted and/or synthesized load profiles in buildings, as well as any similar time series data sets.

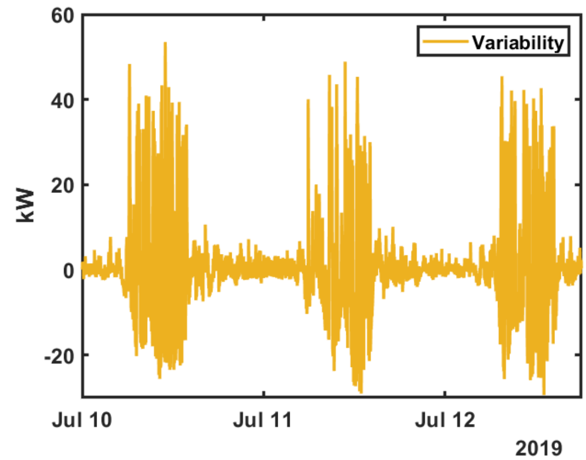
It is assumed that the lower-resolution load profile signals are accurate at the timescales they represent - whether these are measurements or predictions using hourly, daily or any other time-resolution. Therefore, one of the criteria in defining and quantifying variability signals is for the variability profile to have an average energy of zero or near zero. The variability profiles are a measurement of how much the high-resolution values went above and below the base load (the low-resolution values). For building load profiles, measurements can take place at the sub-minute resolution, up to hourly or longer, with the more common utility-installed AMI measurements taking place at 15-minute, 30-minute and one-hour intervals. For modeled building load profiles, the resolution is not explicitly defined, but rather reflects the resolution at which the model is designed to be accurate -most commonly no shorter than hourly. The analysis that follows will thus focus on a base load at the one-hour time-resolution, and a high-resolution version of that load profile, the measured load profile, at the one-minute resolution. With this definition, the following section discusses the possible methods to extract the difference between the measured load profile and its base load, i.e. the variability profile.

B. EXTRACTING VARIABILITY FROM BUILDINGS' LOAD PROFILES

Among the variety of methods to decompose time series signals to extract and study variability are Discrete



a) 1 Minute (red) and 1 Hour (blue) load profiles



b) Variability Profile

FIGURE 4. A visual representation of variability (yellow profile) as the difference between the 1-minute and 1-hour measurements (the difference between the blue and red load profiles).

Wavelet Transform (DWT), Hilbert Vibration Decomposition (HVD) [9], Wavelet packet decomposition (WPD) [10], Empirical Mode Decomposition (EMD), and Variational Mode Decomposition (VMD). Among the mentioned methods, DWT, EMD, and VMD are widely used in non-stationary time series analysis. The following sub-section provide a brief discussion on the EMD and VMD, then a detailed description of DWT. DWT is selected because of its bounded frequency range capability, a desirable feature for variability profile signals, and an essential part in the framework to analyze load profile variability at different time-resolutions.

1) EMPIRICAL MODE DECOMPOSITION (EMD)

EMD is a time-frequency method that is widely used in non-stationary time series data analysis. EMD decompose the data into "fluctuations" and "trend" signals, decomposing a time series signal into intrinsic mode functions (IMFs) and a residual (fast and slow oscillations, respectively) in a recursive process without utilizing filters or any priori basis system [11]. The original time series signal $s(t)$ can be represented by summation of the EMD decomposed signals (IMFs and residual) as shown in equation (2) [12]:

$$s\left(t\right)=\sum_{k=1}^{K}c_{k}\left(t\right)+r_{K}\left(t\right) \tag{2}$$

where $c_k(t)$ represents the IMFs at iteration k and $r_K(t)$ is the final residual. By applying the Hilbert transform on IMFs, frequency information can be extracted [13]. It should be noted that the decomposed signals are extremely dependent on finding maxima/minima techniques, interpolation procedure, and stopping criteria [14]. Among most of the time-frequency decomposition methods, EMD has a lower resolution [6]. Due to lack of theoretical foundation, it is difficult to compare the EMD method with other time-frequency methods. Mode mixing is considered another weakness of the EMD method [6].

Both DWT and EMD decompose the data into "fluctuations" and "trend" signals, but DWT does decomposition through predetermined scales, whereas scales are adaptive (data-driven) in the EMD method [15].

2) VARIATIONAL MODE DECOMPOSITION (VMD)

VMD is a data-adaptive algorithm that attempts to decompose a signal without utilizing a fixed analysis or function into discrete number of modes, which are called IMFs, where each IMF mode is band-limited in the spectral domain [16]–[18]. VMD determines the modes concurrently by identifying the signal peaks in frequency domain. In [14], authors stated that VMD is superior to EMD when separating closely spaced frequencies. VMD decomposes the input signal $s(t)$ into K sets of narrowband IMFs, as shown in equation (3) [19].

$$s\left(t\right)=\sum_{k=1}^{K}u_{k}\left(t\right) \tag{3}$$

where each IMF mode, $u_k\left(t\right)$, can be represented by a frequency and amplitude modulated signal, as shown in equation (4) [19].

$$u_{k}\left(t\right)=A_{k}(\mathrm{t})\cos(\phi_{k}(t)) \tag{4}$$

where $A_k(\mathrm{t})$ and $\phi_k(t)$ are the mode's envelope and phase, respectively. The Modes' envelope is positive and varying slowly, which has an instantaneous frequency of $\phi_k'(t)$ around the central frequency f_k of the mode. The detailed VMD algorithms and calculating steps are available in the MathWorks website [14].

3) DISCRETE WAVELET TRANSFORM (DWT)

Wavelet analysis has been considered as a tool for time series analysis in different research domains, including power system quality analysis [20], signal processing [21], image processing of medical imaging [22], and hydrological time series data analysis [23]. Wavelet analysis provides localized details in time and frequency and can detect changes and provide tools to extract these features. During the decomposition process, wavelet functions provide an approximation signal and a detailed signal based on predetermined scales. These signals are then utilized depending on the objectives of the research analysis [1]. Specifically, time series data can be decomposed into a series of approximation and detail signals (in the time and frequency domain) via DWT process to extract features at different timescales. The original signal is fed into low-pass and high-pass filters, resulting in an approximation signal and a details signal or coefficients, respectively. A signal's trend information (base load signal) is the approximation (low frequency) signal, whereas the detail signal is a high frequency signal that carries the abrupt changes in the original signal, or the variability profile (see Figure 5). Localizing key features of the signal at different timescales is a property of the applied mother wavelet [1].

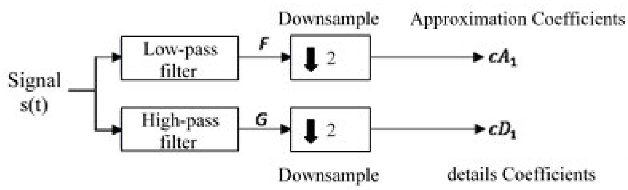


FIGURE 5. **1-level DWT process.**

To reconstruct the original signal $s(t)$, all the details and last approximation level are added together as shown in equation (5).

$$s\left(t\right)=\sum_{k=1}^{K}d_{k}\left(t\right)+a_{K}\left(t\right) \tag{5}$$

where d_k is the details signal at level k and a_K is approximation signal at final level K.

In this research project, the focus is on capturing different detail signals (variability profiles), because they contain the sudden changes that take place when the building and/or the occupant(s) interact with the electrical system to start/stop/alter the energy consumption behavior. The variability profile signal that is extracted is dependent not only on the underlying data, but on the parameters chosen for the DWT process, including level of decomposition and choice of wavelet function. The level of decomposition depends on the type of analysis and intended use case, and on the number of data points available in the measured time series load profile data set. Each decomposition level extracts a fixed frequency range of variability. For example, for a measured data set with a 15-minute time resolution, the first level contains the 15- to 30-minute high-resolution variability, whereas the second level contains the 30- to 60-minute variability. This process is shown in Figure 6 and described in more detail in the following section.

As a final validation on DWT decomposition to extract variability load profiles, the correlation coefficient of the individual variability profile signals (details) are measured

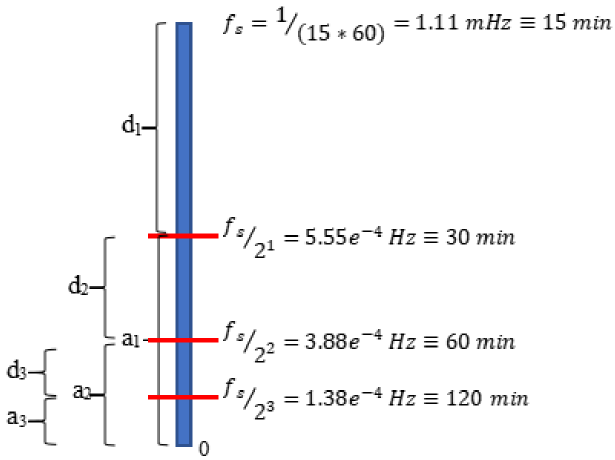


FIGURE 6. DWT - 3 levels of decomposition process.

with respect to the original signal to show the ability of DWT in isolating each level of variability based on a fixed frequency band. The correlation coefficient of two random variables is a measure of their linear dependence and it is shown in equation (6).

$$\rho\,(A, B) = \frac{cov(A, B)}{\sigma_A \sigma_B} \tag{6}$$

where A and B are two random variables or signals, σ_A and σ_B are standard deviation (SD) of A and SD of B respectively.

Results of the correlation coefficient between variability details at different levels for the building load profile are shown in Table 1. As shown in the Table 1, there is no correlation between the levels, which implies the ability of DWT in isolating each level based on the fixed frequencies.

TABLE 1. Correlation coefficients of details signals from level 1 to level 3 and original signal.

Correlation coefficients	Level #1	Level #2	Level #3
Level #1	1.000	0.000	0.000
Level #2	0.000	1.000	0.000
Level #3	0.000	0.000	1.000

C. DEMONSTRATING VARIABILITY EXTRACTION USING A CASE STUDY

1) CASE STUDY DESCRIPTION

In this study, one-minute time-resolution measured load profile data for a cafeteria building in a Midwest state in the USA, for the full year 2019, is analyzed using DWT to extract and analyze load variability features. The data set consists of measured load profiles for the full building consumption and for the end-use load profile categories that make-up the full building consumption: lighting, cooking, mechanical and miscellaneous electrical loads (Mels). The cooking load is fed from two separately monitored electrical panels, identified as Cooking1 and Cooking2. In general, time resolution energy measurements at the one-minute time- resolution is averaged and provided as the reading at that specific timestamp. However, the data that is obtained for this case study is based on single sampled power demand measurements taken at each designated one-minute timestamp. This measurement approach has advantages and disadvantages. The disadvantage comes from the fact that it is a single measurement of power and not an average measurement of the energy over the full minute. However, using such single sampled data can be advantageous in that it represents more extreme variability than available in one-minute energy data. Furthermore, the full building measurement used in the analysis is the summation of each of the end-use load categories at each timestamp. This is used for consistency in the analysis of the full load (the aggregate data) and each of the end-use load category data. As will be seen by the analysis of the results in the remainder of the paper, any conclusions inferred from the results is applicable to the building under consideration and not a generalization. The focus of paper and its analysis is to provide a framework that is applicable to any time series data sets. In this context, the paper presents this data to illustrate variability load profiles and to provide a methodology/approach for those who are interested in similar work.

a: LOOKING AT AGGREGATED FULL BUILDING LOAD PROFILE DATA

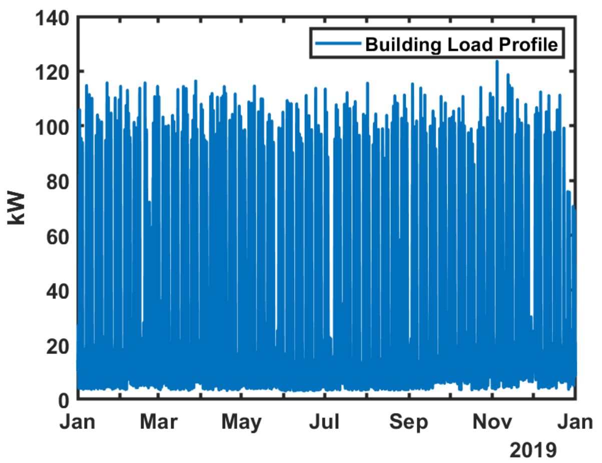
Data pre-processing is performed before applying time-frequency analysis by isolating null data and replacing them using interpolation. With a one-minute time-resolution, a total of five sets of 525,600 data points, each set representing one end-use load category over a full year, are available for this analysis. Of this total, 1524 data points (0.2%) had null values and were replaced. By adding each of the end-use category data points at each timestamp, the full building consumption is aggregated from the end-use data. Figure 7 shows the aggregated electricity consumption load profile for the entire year, the load duration curve for the year, and an example load profile for one day on June 4, 2019.

b: LOOKING AT END-USE LOAD CATEGORY PROFILE DATA

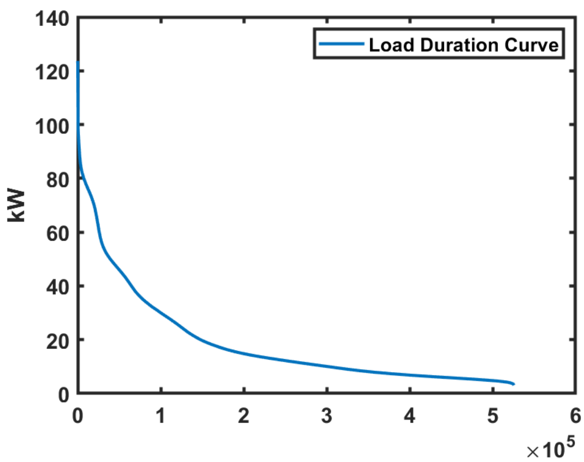
In addition to the aggregated building load profile, one-minute time resolution load profiles for the five end-use load categories are also available and used in the variability analysis. Figure 8 shows similar plots to Figure 7, for each end use category. Table 2 provides several summary statistics for the case study building, including the peak, average and minimum load for the aggregate load and for each of the end-use load categories independently.

2) VARIABILITY EXTRACTION

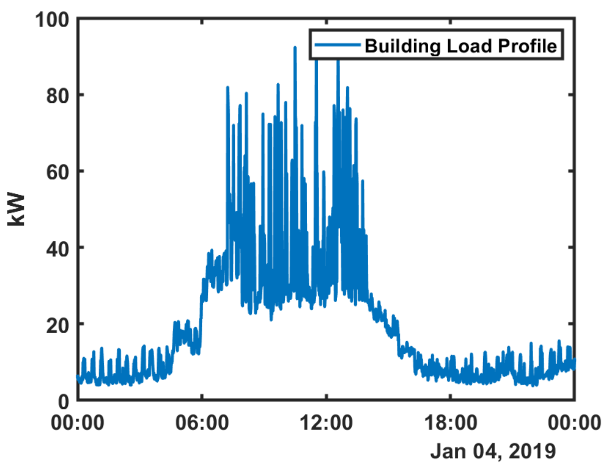
As discussed previously, variability can be described as the difference between a given measured, modeled or predicted load profile for a given building and a lower-resolution version of that same load profile (its base load). The choice



a) Full year load profile

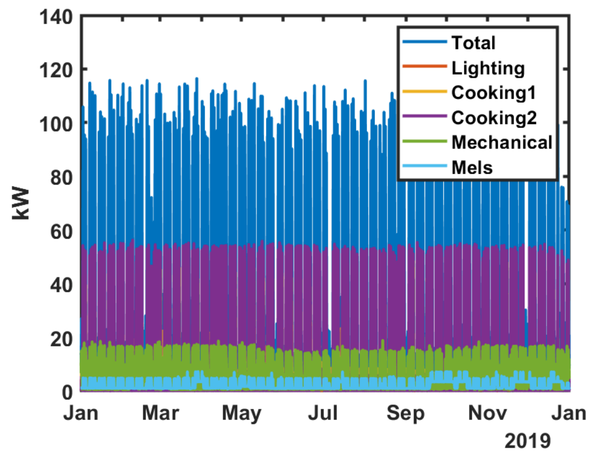


b) Load duration curve

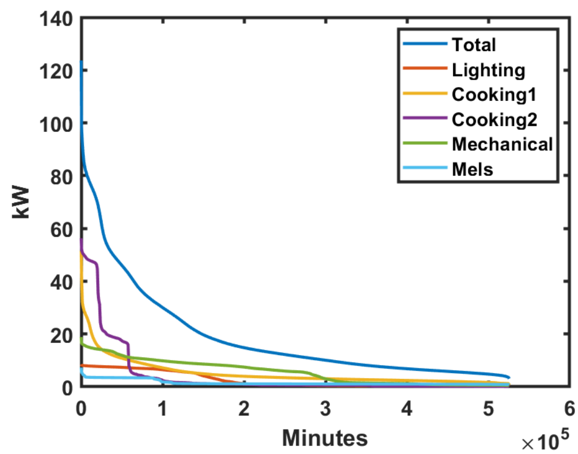


c) Sample daily load profile

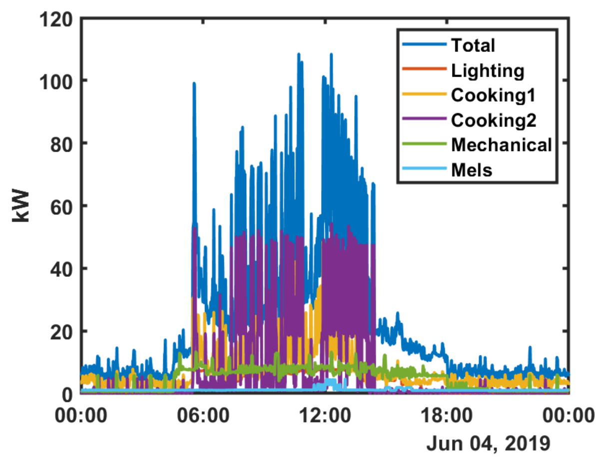
FIGURE 7. Full building load profile plots for 2019.



a) Electricity usage for the total building and all end-use load categories in 2019.



b) Load duration curves for the total building and for all end-use load categories in 2019.



c) Sample daily load profiles on June 4, 2019.

FIGURE 8. Load profile for Total electricity usage and for end-use load categories for the building in 2019.

of the desired level of variability to extract depends on the desired/considered base load and the time resolution of the measured building load profile. Furthermore, depending on the given time-resolution, various base loads and associated variability profiles can be extracted and quantified. With the application of DWT - Daubechies-4 (db4) wavelet function

TABLE 2. Summary statistics for the Case study load profiles.

		Cafe Building Load Data (in kW)		
		Peak Load	Average Load	Minimum Load
End-Use Load Category	Aggregate Load	123.6	18.8	3.0
	Lighting Load	8.5	2.4	0.0
	Cooking1 Load	51.5	5.1	0.7
	Cooking2 Load	56.3	4.1	0.0
	Mechanical Load	18.9	5.6	0.4
	Mels Load	7.3	1.5	0.7

to the case study under considerations, 19 distinct/singular frequency bands with associated base loads and variability profiles are available for selection depending on the purpose and application as shown in Table 3 [1]. In this case study, the time resolution range is based on the measured time resolution (one-minute scale) and the desired base load (lower time scale version at the one-hour scale). Since DWT isolates ranges based on the power of two, with a one-minute scale, the upper bound of the range for this one-hour scale is 64-minutes and not exactly 60-minutes.

TABLE 3. Variability content at each extracted DWT level.

High Frequency Signal Information Content at each Time –Resolution (min) and Decomposition Level			
DWT Levels	Load Profile Time –Resolution Range	DWT Levels	Load Profile Time – Resolution Range
1	1 - 2 min	11	1024 - 2048 min
2	2 - 4 min	12	2048 - 4096 min
3	4 - 8 min	13	4096 - 8192 min
4	8 - 16 min	14	8192 - 16384 min
5	16 - 32 min	15	16384 - 32768 min
6	32 - 64 min	16	32768 - 65536 min
7	64 - 128 min'	17	65536 - 131072 min
8	128 - 256 min	18	131072 - 262144 min
9	256 - 512 min	19	262144 - 524288 min
10	512 - 1024 min		

For the case study, singular and cumulative load variability profiles at the 6th level are considered when using the measured one-minute time-resolution data and the hourly base load. This level is the 32-64-minute range, indicating that all high frequency content that is below the 64-minute range is considered variability. To illustrate how the variability load profile changes with scaled levels, consider the results when using a DWT level 1 decomposition. In this level, the high frequency variations below 2-minutes are considered part of the variability profile and any remaining high frequency content above the 2-minute range is considered part of the base load. The result of applying a DWT level 1 decomposition on the building load profile is shown in Figure 9 for one day. In this Figure, the original measured building load profile is shown as the blue line, the base load is shown in red, and the extracted variability profile in yellow. This variability profile is the difference between the red line and the blue line (the difference between the original measured data at the one-minute time-resolution and the base load at the 2-minute). As can be seen in the Figure, the base load still contains high frequency variations.

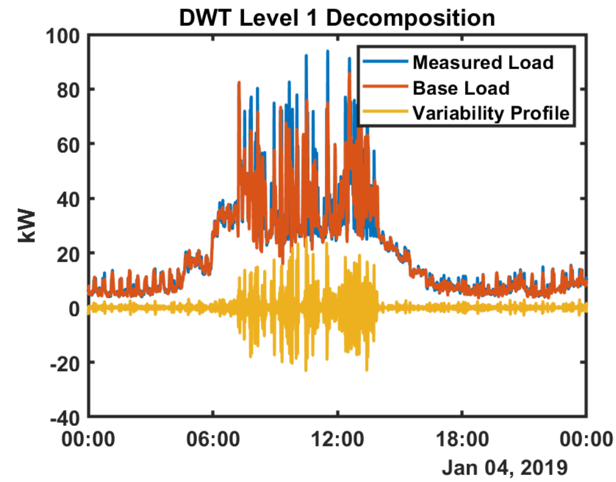


FIGURE 9. Result of applying a DWT level 1 decomposition on a sample daily load profile.

Now consider the results if a DWT level 4 decomposition takes place. As can be seen in Figure 10, as the level of decomposition increases, the base load profile (approximation signal) is a much smoother version of the original signal, i.e. more variability is extracted in the process (the yellow line). Specifically, the high frequency variations below 16-minutes are considered part of the variability profile, and any remaining content above the 16- minute range is considered part of the base load. Note that the level 4 variability consists of the cumulative level 1 to level 4 decomposition details, as discussed above (see equation 5). Again, sub-hourly variation is still present in the base load.

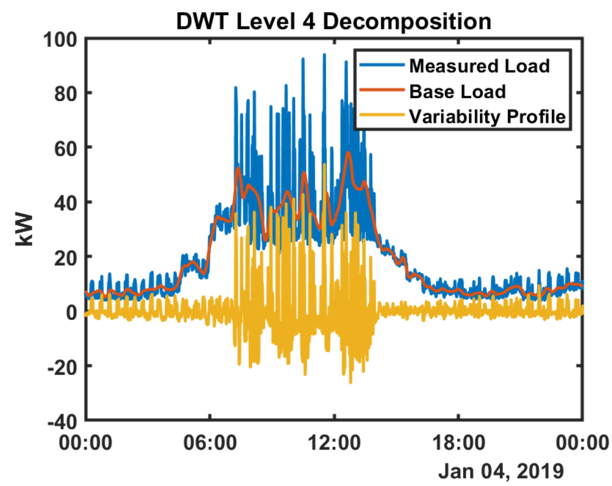


FIGURE 10. Result of applying a DWT level 4 decomposition on a sample daily load profile.

Finally, consider the results when applying a DWT level 6 decomposition. The results are shown in Figure 11 for the

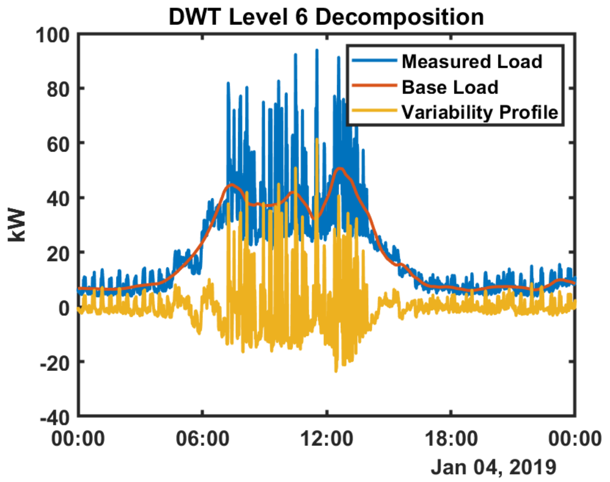


FIGURE 11. **Result of applying a DWT level 6 decomposition on a sample daily load profile.**

same day. Again, with a level 6 decomposition, the high frequency variations below 64-minutes is considered part of the variability profile and any remaining high frequency content above the 64- minute range is considered part of the base load.

The same process is applied to each of the end-use load categories. The lighting load is presented in Figure 12 for a sample day in blue. The base load and variability profile using a level 6 DWT decomposition are shown in red and yellow. As the figure shows, the lighting end-use load category contains a low level of variability on this day, when compared with the Cooking2 and Mechanical end-use load categories shown in Figure 13a and 13b.

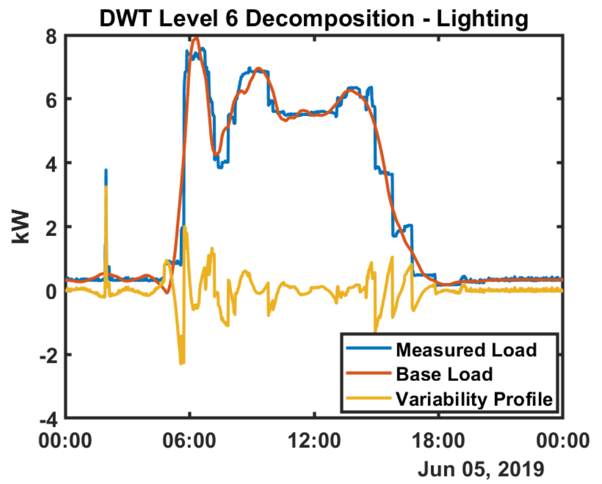
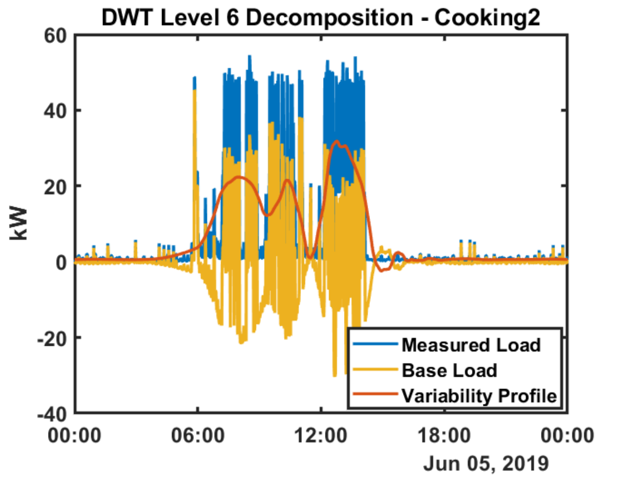


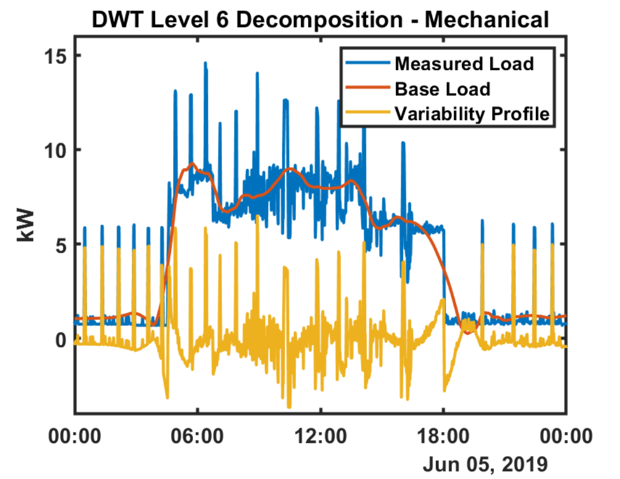
FIGURE 12. **Sample daily Lighting category profile, with extracted base load and variability.**

III. QUANTIFYING BUILDING LOAD PROFILE VARIABILITY

The previous section explored tools to extract load variability profiles and define variability based on well-defined frequency bounds. This section will present a metric to



a) Result of applying a DWT level 6 decomposition on a sample daily Cooking2 load profile.



b) Result of applying a DWT level 6 decomposition on a sample daily Mechanical load profile.

FIGURE 13. **Sample Daily end-use load category profiles, with extracted base loads and variability.**

quantify the extracted variability profile and provide analysis on the variability results. Figure 14 shows the extracted time series variability profile signal for the total building for a full year using a DWT level 6 decomposition. In addition, Figure 15 shows the same results, but for the total building and for each end-use load category for a three-day period, to focus on daily behavior. With a DWT level 6 decomposition, it is implied that all the high frequency content below 64 minutes (i.e. the cumulative high frequency signal between 1 minute and 64 minutes) is considered variability.

In addition to visualizing the variability as a time series signal, it is useful to examine the statistical distribution of variability over a given time frame. For this purpose, the sign of the variability is ignored, focusing only on the magnitude of variability above or below the base load. Figure 16 shows the cumulative distribution of the absolute variability for the full building and each end-use category. Similar to a load

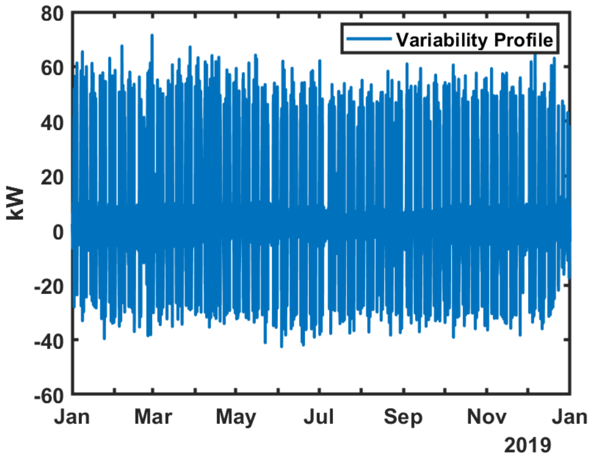


FIGURE 14. Full year Variability profile for the total building load profile.

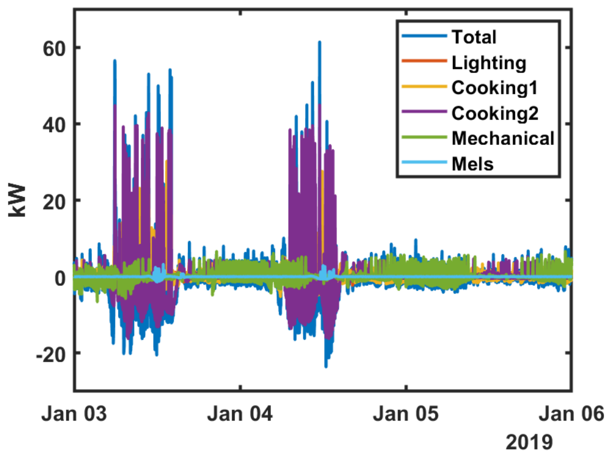


FIGURE 15. Variability profiles for the building's end-use load categories over 3 days.

TABLE 4. Summary statistics of absolute variability for a full year.

		Statistics of Absolute Variability (in kW)			
		Peak	Mean	Median	RMS
End-Use Load Category	Aggregate Load	71.6	4.30	1.70	8
	Lighting Load	7.3	0.2	0.1	0.5
	Cooking1 Load	35.5	1.4	0.8	2.8
	Cooking2 Load	50.7	3.2	0.5	7.2
	Mechanical Load	9	1.2	0.7	1.7
	Mels Load	3.6	0.10	0	0.3

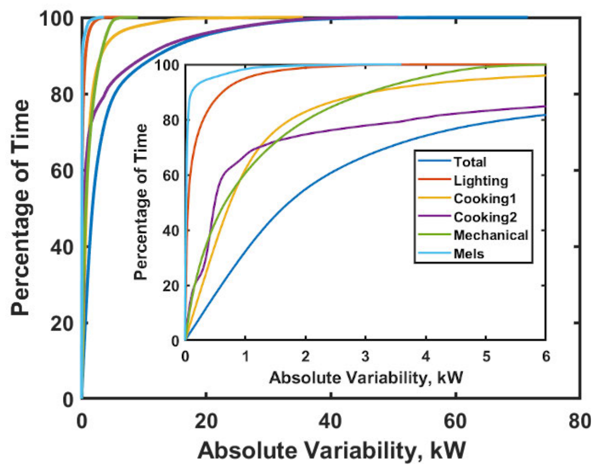


FIGURE 16. Cumulative distribution of absolute variability for the full building and all end-use load categories, over a full year. The inlaid plot zooms in on the leftmost portion of the distribution.

duration curve, this represents the probability (or percentage of time) that the absolute variability falls below a given kW value. Table 4 provides several summary statistics for the distribution of variability over the full year.

It is evident from Figure 16 that the variability in the building and all end-use categories is highly skewed - a small percentage of times exhibit variability significantly higher than the average. This is also reflected in the high ratio of mean to median variability, as well as the ratio of root-mean-square to mean.

In order to quantitatively compare variability over different time frames, end use categories, and buildings, it is useful to summarize the distribution of variability with one or more of the metrics above. For the remaining analysis in this paper, the Root-Mean-Square Variability (RMSV) is chosen, formally defined in Equation 7, where, d_i is the variability profile at timestamp i, and N is the number of timestamps.

$$RMSV = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(d_i)^2} \quad (7)$$

The RMSV can be loosely interpreted as the average sign-independent deviation of the actual load profile from the base load, over a given period of time. It differs from the actual absolute mean variability in that it is more sensitive to outliers, and thus somewhat more representative of skewed distributions such as those in Figure 16.

While Table 4 presents statistics for the full year of data, the RMSV can be calculated over any time frame to compare variability at various scales. Figure 17 shows a 4-day sample aggregate building load profile along with the extracted variability and the associated daily RMSV levels. In this example the RMSV provides a useful summary of the higher variability of the weekdays compared to that of the weekend.

IV. ANALYZING BUILDING LOAD PROFILE VARIABILITY

A. ANALYSIS OF YEARLY VARIABILITY

Figure 18 shows the yearly RMSV for the total building load profile and for each of the end-use load categories. As can be seen from the Figure, each end-use load category has different level of variability. The RMSV for the entire building

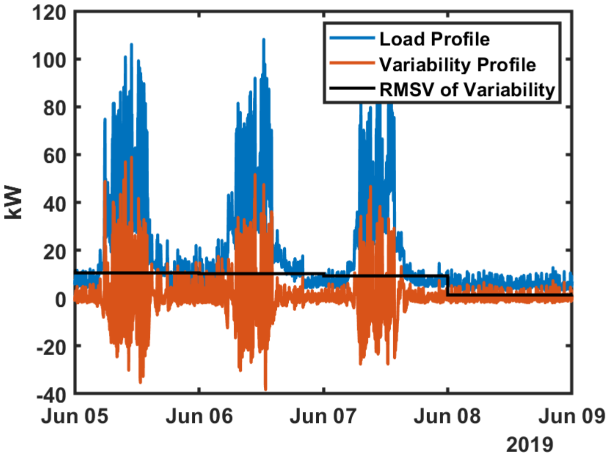


FIGURE 17. Full building load profile and variability for 4 days, with RMSV.

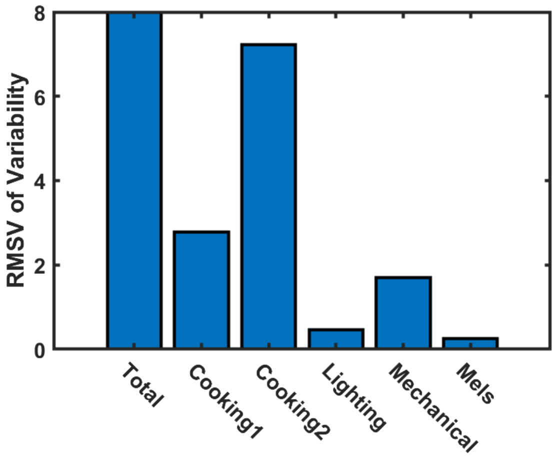


FIGURE 18. Yearly RMSV for the total building and its end-use load categories.

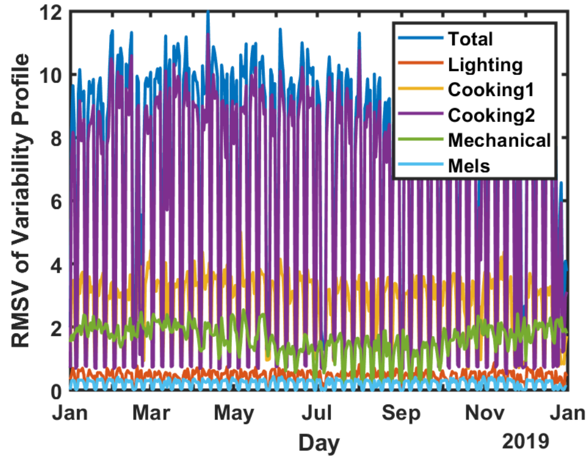


FIGURE 19. Daily RMSV for the full building and each end-use category throughout the year.

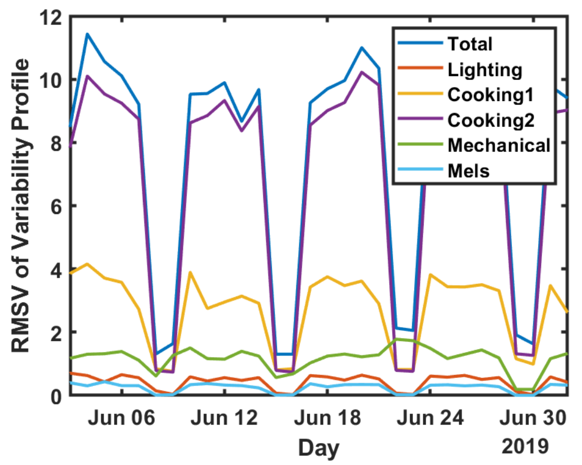


FIGURE 20. A zoomed-in view of Daily RMSV in Figure 10 for one month.

load profile is 8 kW whereas for Cooking2 it is 7.2 kW. Cooking2 has a higher variability than Cooking1 and higher than all other end-use load categories. Also, the variability is not additive, meaning that the summation of each variability in the end-use categories does not equal to the variability in the aggregate building load variability profile.

B. ANALYSIS OF VARIABILITY PROFILE CHANGES THROUGHOUT THE YEAR?

In addition to the RMSV for the full year, daily RMSV for the building total and for each end-use load category load profiles are analyzed. Figure 19 shows the full year, daily RMSV for the total load and for each end use for a DWT level 6 decomposition. Figure 20 shows a zoomed-in view of Figure 19 for one month. In this Figure, daily patterns of low and high variability profiles emerge.

Figure 21 shows a plot of the cumulative distribution for the daily RMSV for the total building load profile and for each of the end-use load categories. The flat portions of the distribution indicate gaps in observed daily variability. Approximately 30% of the days in this year are weekends or other

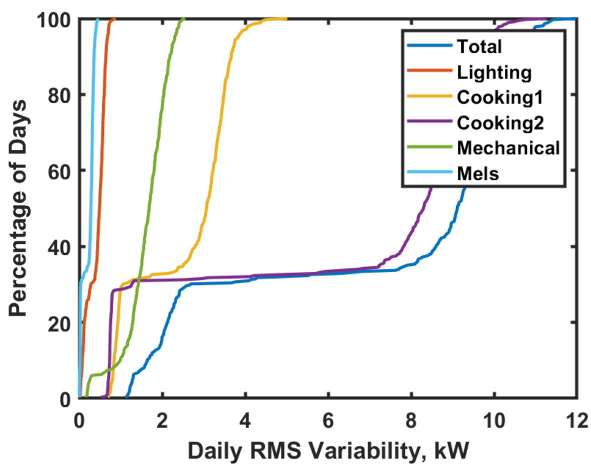


FIGURE 21. Cumulative Distribution of Daily RMSV over a full year, for each end-use category.

low-operation days, with RMSVs ranging from 0 to 2.5 kW for the full building. The remaining 70% of days exhibit much higher RMSVs, between 8 and 12 kW. Almost no days have an RMSV between 2.5 and 8 kW. The variability of the Cooking2 end-use category closely follows that of the full building, and all end-use categories except the mechanical systems exhibit the same discontinuity in weekend-weekday behavior. The discontinuity in the mechanical system variability is instead at approximately 8% of days, indicating that the variability in this category falls within a narrow, continuous range independent of whether the building is occupied, except for a smaller fraction of days where the equipment exhibits a steep drop in variability. Table 5 provides several summary statistics of daily RMSV over the full year.

TABLE 5. **Summary statistics of daily RMSV over a full year.**

		Statistics of Daily RMSV (in kW)		
		Peak	Average	Minimum
End-Use Load Category	Aggregate Load	12	7.1	1.1
	Lighting Load	0.9	0.4	0
	Cooking1 Load	5	2.5	0.7
	Cooking2 Load	11.3	6.2	0.5
	Mechanical Load	2.6	1.6	0.2
	Mels Load	0.4	0.2	0

C. RELATIONSHIP BETWEEN VARIABILITY AND ENERGY CONSUMPTION

The RMSV is an absolute measure of variability, in the same units as the load profile itself (in this case kW), as opposed to a relative measure of variability (such as a percentage increase or decrease). Knowing that Cooking2 has variability with an average yearly RMSV of 6.2 kW for instance, does not on its own convey whether that variability is 'large' or 'small'. To fully compare variability across different time frames, end-use categories or buildings, it is important to establish an appropriate normalized metric as well.

While variability can be compared to, or normalized by, several different metrics, one intuitive and informative possibility is the consumption of the underlying load. Specifically, the average consumption over the same period of time that the RMSV is calculated over. Formally, this Normalized Root Mean Square Variability (NRMSV) can be defined as

$$NRMSV = \left(\sqrt{\frac{1}{N}\sum_{i=1}^{N}(d_i)^2}\right) \Big/ \frac{1}{N}\sum_{i=1}^{N} s_i \tag{8}$$

where d_i and s_i represent the raw variability and the original time series signal respectively, at timestep i. Note that equation 8 applies to the aggregate load profile as well as to each of the end-use load categories.

Figures 22 and 23 show the average annual consumption of the Cafe building and each of its end-use categories, both in absolute terms (kW) and as a percentage of the total load. The percentage results show that the mechanical end-use load category has the highest percentage of consumption at almost 30%, followed by Cooking1, Cooking2, lighting and lastly Mels end-use load category with 8% consumption. The average consumption results show that the aggregate load has an average of 18.8 kW, whereas the mechanical end-use load category has an average of about 5.6 kW.

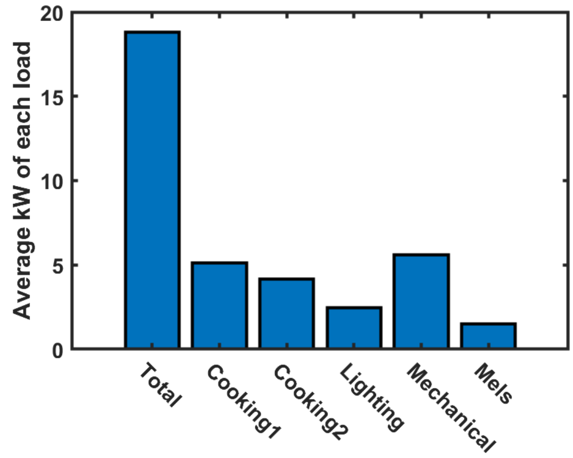


FIGURE 22. **The full year average daily consumption for the total building and each end-use category (Cafe).**

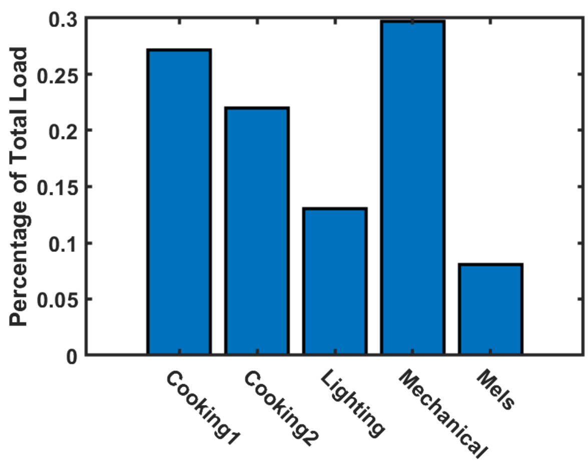


FIGURE 23. **The full year consumption percentage for each end-use category (Cafe).**

Once the average for each load is determined, the normalized RMSV (NRMSV) is calculating according to equation 8, and the results are shown in Figure 24. Comparing the consumption percentage and associated level of variability, in this building, a few important observations can be made. First, the mechanical end-use category has the largest consumption percentage, but the third highest variability (from Figure 18), indicating that the largest consuming end-use load category is not necessarily the main contributor to variability in the building. The Cooking2 end-use category has the third largest consumption percentage, but has the highest variability. These results are of course specific to this particular building and set of end-use categories, but indicate that when comparing disparate systems, the variability of a system may not be simply proportional to the consumption of that system.

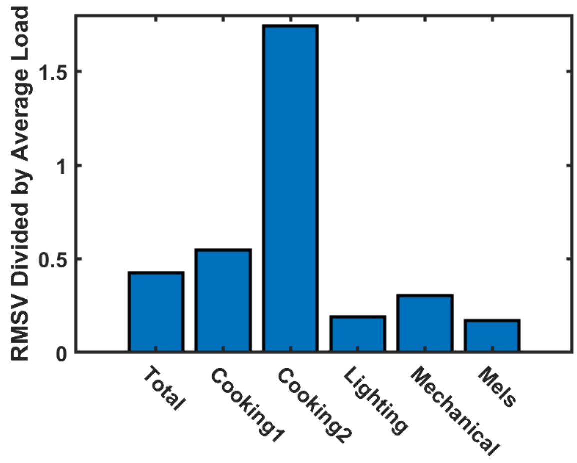


FIGURE 24. **Normalized RMSV for total building and each end-use category (Cafe).**

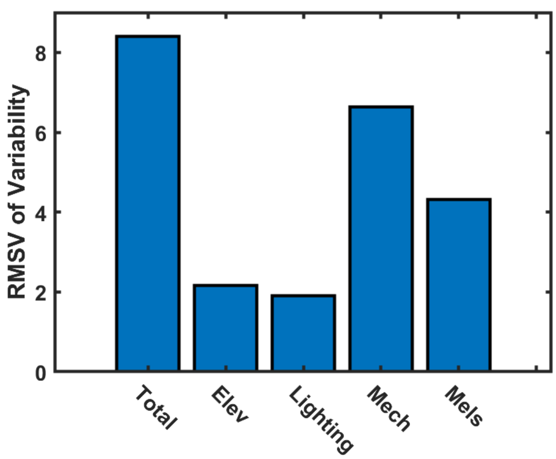


FIGURE 25. **Yearly RMSV for the total building and its end-use load categories (Office).**

This is also evident in analyzing another building, an office building in the same location with same type of time-resolution and data samples. Figure 25 shows the raw RMSV for the end use categories. In this Figure, the mechanical system has the highest absolute variability.

Similar to the process applied earlier, the yearly average and percentage of consumption for each end-use category is plotted in Figures 26 and 27.

The Mels end-use category has the highest consumption, and the Elevator end-use category the lowest by far, accounting for less than 2 percent of the building's yearly consumption. The normalized RMSV is shown in Figure 28. Here it is clear that when considering variability relative to consumption, the elevators have by far the largest NRMSV. This is somewhat intuitive, as while the elevators may not consume much power on average compared to other systems, their consumption is constantly and rapidly changing in an unpredictable manner. Thus the elevators' contribution to the building's variability (in Figure 25) is far higher than might be expected from consumption alone.

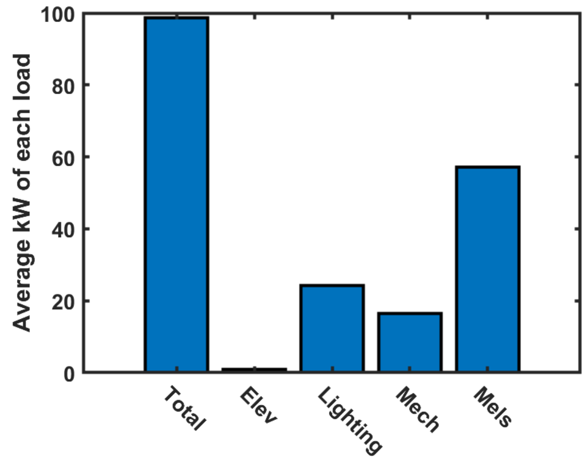


FIGURE 26. **The full year average daily consumption for total building and each end-use category.**

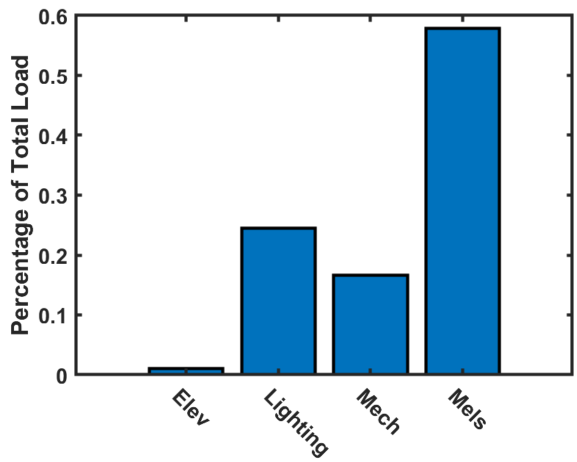


FIGURE 27. **The full year consumption percentage for each end-use category (Office).**

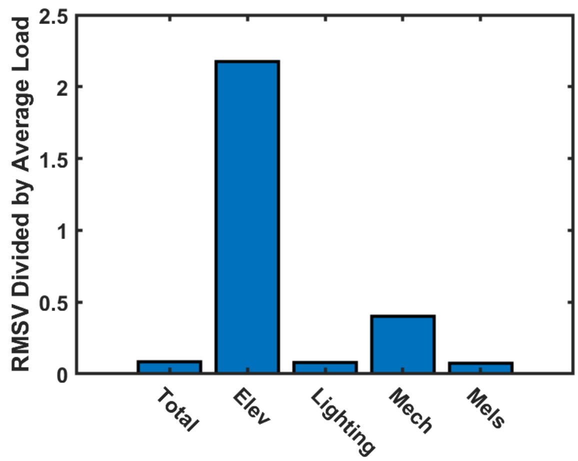


FIGURE 28. **Normalized RMSV for total building and each end-use category.**

V. SUMMARY AND CONCLUSION

This is part of ongoing work to discover typical variability in aggregate building load profiles and in end-use load category profiles for a large set of building types. The presented work herein examines high resolution building energy consumption load profiles and presents a methodology to analyze it and extract useful information that can be used in a host of applications. Specifically, the paper presented an investigation of variability in load profiles and presented detailed study to define, extract, quantify and analyze variability in aggregate building and in end-use load category AMI data. The discussion of variability and its analysis is based on a case study of an actual one-minute time resolution for two buildings in a Midwest state in the USA for a full year of data collected in 2019.

A framework is presented to isolate the variability in a measured signal from the base load. This variability can be quantified and compared across various time frames, end-use categories, or buildings, in both absolute and relative terms. The case study results illustrate that both absolute and relative metrics are important, as the variability in a system is not necessarily proportional to its consumption.

This framework lays the groundwork for future research into building load variability. When the variability is isolated and quantified, it can be both analyzed and modeled with respect to any number of potential inputs, in order to both understand and predict variability in complex systems.

ACKNOWLEDGMENT

The authors would like to thank Monica Neukomm at DOE for her support and vision as manager of the project sponsoring the work.

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.

REFERENCES

[1] A. Parker, K. James, D. Peng, and M. A. Alahmad, "Framework for extracting and characterizing load profile variability based on a comparative study of different wavelet functions," *IEEE Access*, vol. 8, pp. 217483–217498, 2020, doi: 10.1109/ACCESS.2020.3042125.

[2] *IEEE Guide for Conducting Distribution Impact Studies for Distributed Resource Interconnection*, Standard 1547.7-2013, 2014, pp. 1–137, doi: 10.1109/IEEESTD.2014.6748837.

[3] R. J. Broderick, J. E. Quiroz, M. J. Reno, A. Ellis, J. Smith, and R. Dugan. (2013). *Time Series Power Flow Analysis for Distribution Connected PV Generation* Accessed: Dec. 15, 2020. [Online]. Available: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

[4] A. Pinceti, O. Kosut, and L. Sankar, "Data-driven generation of synthetic load datasets preserving spatio-temporal features," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Aug. 2019, pp. 1–5, doi: 10.1109/PESGM40551.2019.8973532.

[5] A. J. R. Reis and A. P. A. D. Silva, "Feature extraction via multiresolution analysis for short-term load forecasting," *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 189–198, Feb. 2005, doi: 10.1109/TPWRS.2004.840380.

[6] C. Bianchi, L. Zhang, D. Goldwasser, A. Parker, and H. Horsey, "Modeling occupancy-driven building loads for large and diversified building stocks through the use of parametric schedules," *Appl. Energy*, vol. 276, Oct. 2020, Art. no. 115470, doi: 10.1016/j.apenergy.2020.115470.

[7] S. M. Alshareef and W. G. Morsi, "Probabilistic models for residential and commercial loads with high time resolution," in *Proc. IEEE Electr. Power Energy Conf. (EPEC)*, Oct. 2019, pp. 1–6, doi: 10.1109/EPEC47565.2019.9074784.

[8] A. Albert, T. Gebru, J. Ku, J. Kwac, J. Leskovec, and R. Rajagopal, "Drivers of variability in energy consumption," in *Proc. ECML-PKDD DARE Workshop Energy Anal.*, Sep. 2013, pp. 1–12.

[9] M. Civera and C. Surace, "A comparative analysis of signal decomposition techniques for structural health monitoring on an experimental benchmark," *Sensors*, vol. 21, no. 5, p. 1825, Mar. 2021, doi: 10.3390/s21051825.

[10] J. Kevric and A. Subasi, "Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system," *Biomed. Signal Process. Control*, vol. 31, pp. 398–406, Jan. 2017, doi: 10.1016/j.bspc.2016.09.007.

[11] A. Zeiler, R. Faltermeier, I. R. Keck, A. M. Tome, C. G. Puntonet, and E. W. Lang, "Empirical mode decomposition-an introduction," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2010, pp. 1–8.

[12] P. Flandrin. *Empirical Mode Decomposition vs. wavelets for the analysis of Scaling Processes 'Wavelets and Fractals' Esneux (B)*. Accessed: Jun. 18, 2020. [Online]. Available: http://www.afaw.ulg.ac.be/scam/talks/flandrin.pdf

[13] M. X. Cohen, *Analyzing Neural Time Series Data: Theory and Practice*. Cambridge, MA, USA: MIT Press, 2014.

[14] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2014.

[15] P. Goncalves, P. Abry, G. Rilling, and P. Flandrin, "Fractal dimension estimation: Empirical mode decomposition versuswavelets," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Apr. 2007, pp. 1153–1156, doi: 10.1109/ICASSP.2007.366889.

[16] P. S. Kumar, L. A. Kumaraswamidhas, and S. K. Laha, "Selecting effective intrinsic mode functions of empirical mode decomposition and variational mode decomposition using dynamic time warping algorithm for rolling element bearing fault diagnosis," *Trans. Inst. Meas. Control*, vol. 41, no. 7, pp. 1923–1932, Aug. 2018, doi: 10.1177/0142331218790788.

[17] W. Liu, S. Cao, and Y. Chen, "Applications of variational mode decomposition in seismic time-frequency analysis," *Geophysics*, vol. 81, no. 5, pp. V365–V378, Jul. 2016, doi: 10.1190/geo2015-0489.1.

[18] S. Lahmiri, "Comparative study of ECG signal denoising by wavelet thresholding in empirical and variational mode decomposition domains," *Healthcare Technol. Lett.*, vol. 1, no. 3, pp. 104–109, Sep. 2014, doi: 10.1049/htl.2014.0073.

[19] MathWorks. (2020). *Variational Mode Decomposition—MATLAB VMD*. Accessed: Jan. 13, 2021. [Online]. Available: https://www.mathworks.com/help/wavelet/ref/vmd.html

[20] M. Nicola, C.-I. Nicola, S. Popescu, D. Sacerdotianu, and M. Duta, "Power quality analysis system based on LabVIEW real-time and reconfigurable FPGA modules using wavelet transform," in *Proc. Int. Conf. Appl. Theor. Electr. (ICATE)*, Oct. 2018, pp. 1–6, doi: 10.1109/ICATE.2018.8551364.

[21] M. V. Goryunova, L. S. Kuleshova, and A. I. Khakimova, "Application of signal analysis for diagnostics," in *Proc. Int. Conf. Ind. Eng., Appl. Manuf. (ICIEAM)*, May 2017, pp. 1–5, doi: 10.1109/ICIEAM.2017.8076487.

[22] T. D. Gadhiya, A. K. Roy, S. K. Mitra, and V. Mall, "Use of discrete wavelet transform method for detection and localization of tampering in a digital medical image," in *Proc. IEEE Region 10 Symp. (TENSYMP)*, Jul. 2017, pp. 1–5, doi: 10.1109/TENCONSpring.2017.8070082.

[23] M. Yang, Y.-F. Sang, C. Liu, and Z. Wang, "Discussion on the choice of decomposition level for wavelet based hydrological time series modeling," *Water*, vol. 8, no. 5, p. 197, May 2016, doi: 10.3390/w8050197.

[24] J. Peppanen, M. Hernandez, J. Deboever, and M. Rylander, "Enhanced load modeling: Leveraging expanded monitoring and metering," EPRI, Palo Alto, CA, USA, Tech. Rep. 3002015283, 2019.

ANDREW PARKER currently works as a Mechanical Engineer with the Commercial Buildings Controls and Analytics Group, National Renewable Energy Laboratory. Since joining the laboratory in 2010, he has focused on making high-efficiency building design commonplace by making existing energy analysis tools more accessible. His past projects include the development of a now-commercialized mobile energy auditing tool, the development of the OpenStudio® energy modeling platform, and the development of various software and analysis tools in support of electric and gas utilities. Recently, he has focused on using OpenStudio to develop a nationwide model of the commercial building stock in the USA, and is particularly interested in the integration of building energy modeling with grid simulation.

SAM MOAYEDI received the B.Sc. degree in electrical engineering from Azad University, in 2009, and the M.Sc. degree in electrical engineering from Wichita State University, in 2014. He is currently pursuing the Ph.D. degree in architectural engineering, focusing on electrical systems in building with the University of Nebraska–Lincoln. His research interests include power distribution system analysis, real-time remote power monitoring in buildings' electrical systems, and building's electrical load profile analysis.

KEVIN JAMES received the bachelor's degree in mechanical and electrical engineering from Iowa State University in 2003, and the master's and Ph.D. degrees in mechanical engineering from the University of Michigan in 2005 and 2009, respectively. As a Student and a Postdoctoral Researcher with the University of Michigan, he published several articles on underwater acoustic modeling in uncertain environments. He is currently a part time employee with Durham School of Architectural Engineering and Construction, University of Nebraska–Lincoln, researching topics including electric vehicle usage and variability in electrical grids.

DONGMING PENG received the Ph.D. degree in computer engineering from Texas A&M University. He has been an Associate Professor with the Electrical and Computer Engineering Department, University of Nebraska–Lincoln. He has published over 100 technical articles in the research areas of digital image processing, multimedia security, wireless sensor networks, cybersecurity, computer architecture, and related topics. He has conducted various research projects in these areas with a special contribution in energy-efficient wireless and secure communications in multimedia sensor networks.

MAHMOUD A. ALAHMAD (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Idaho, Moscow, in 1989, 1991, and 2005, respectively. He is currently an Associate Professor with Durham School of Architectural Engineering and Construction, University of Nebraska–Lincoln. His previous publications include several articles in the area of electrified transportation, power management, and microbattery testing and system design implementation. His industry experience includes electrical distribution system infrastructure planning and design for new and renovated facilities. Since 1996, he has been holding a Professional Engineering Licensure. His research interests include electrified transportation, load profile analysis and real-time power monitoring in the built environment, battery power management, and renewable energy alternatives. He is a Peer Reviewer for several journals and conferences, including the IEEE Transactions on Industrial Electronics.