Smart City Digital Twin–Enabled Energy Management: Toward Real-Time Urban Building Energy Benchmarking

Abigail Francisco, S.M.ASCE[1]; Neda Mohammadi, Ph.D., A.M.ASCE[2]; and John E. Taylor, Ph.D., M.ASCE[3]

Abstract: To meet energy-reduction goals, cities are challenged with assessing building energy performance and prioritizing efficiency upgrades across existing buildings. Although current top-down building energy benchmarking approaches are useful for identifying overall efficient and poor performers across a portfolio of buildings at a city scale, they are limited in their ability to provide actionable insights regarding efficiency opportunities. Concurrently, advances in smart metering data analytics combined with new data streams available via smart metering infrastructure present the opportunity to incorporate previously undetectable temporal fluctuations into top-down building benchmarking analyses. This paper leveraged smart meter electricity data to develop daily building energy benchmarks segmented by strategic periods to quantify their variation from conventional, annual energy benchmarking strategies and investigate how such metrics can lead to near real-time energy management. The periods considered include occupied periods during the school year, unoccupied periods during the school year, occupied periods during the summer, unoccupied periods during the summer, and peak summer demand periods. Results showed that temporally segmented building energy benchmarks are distinct from a building's overall benchmark. This demonstrates that a building's overall benchmark masks periods in which a building is over- or underperforming during the day, week, or month; thus, temporally segmented energy benchmarks can provide a more specific and accurate measure for building efficiency. We discussed how these findings establish the foundation for digital twin–enabled urban energy management platforms by enabling identification of building retrofit strategies and near-real-time efficiency in the context of the performance of an entire building portfolio. Temporally segmented energy benchmarking measures generated from smart meter data streams are a critical step for integrating smart meter analytics with building energy benchmarking techniques, and for conducting smarter energy management across a large geographic scale of buildings. **DOI: 10.1061/(ASCE)ME.1943-5479.0000741.**

Author keywords: Building energy benchmarking; Commercial buildings; Community energy-efficiency; Digital twin; Energy management; Smart city.

Introduction

In cities, prominent challenges such as urbanization and rising greenhouse gas emissions have sparked efforts to make cities smarter (Pierce and Andersson 2017). Because buildings account for the majority of energy consumption in cities, and because of their high potential for energy conservation through retrofits or operational improvements (IPCC 2007), they have become a key focus for smart city initiatives (Baxter et al. 2011). At the intersection of smart cities and building energy efficiency lies the opportunity for real-time intelligent planning and urban energy management (Hastak and Koo 2016). Smart city digital twins, a recent endeavor to create a digital replica of city infrastructure linked to real-time city data, are envisioned to improve city monitoring, control, and decision making through enhanced visualization and interaction with city data (Mohammadi and Taylor 2017). Smart city digital twins are intended to capture and incorporate urban complexities across time and space through streamed data; given the increasing availability of building performance data at urban scales (BuildSmart DC 2017), they can be a promising platform for building portfolio performance assessment and urban energy management (i.e., digital twin–enabled energy management).

Concurrently, policies aimed at transitioning cities to more sustainable, energy-efficient urban areas are increasing. As of March 2019, over 90 cities, ten counties, and two states in the US have committed to consuming energy entirely from renewable energy sources by no later than 2050 (Sierra Club 2019). Other policies such as building benchmarking ordinance requirements are requiring public release of whole-building energy consumption and production data for individual buildings at community and city scales (e.g., Building Rating 2019; BuildSmart DC 2017). Harnessing the potential of such data, made available through large investments in smart infrastructure, is critical to fulfill greenhouse gas emission reduction commitments (Zuo et al. 2013) and to strive toward digital twin–enabled smart city energy management.

Traditional ways of accomplishing building portfolio assessments across large scales include building energy benchmarking, which typically is conducted on an annual basis (Borgstein et al. 2016). However, such metrics do little to inform specific efficiency opportunities to target or support real-time management of energy efficiency. The availability of smart meter data across a community

[1]Ph.D. Student, School of Civil and Environmental Engineering, Georgia Institute of Technology, Mason Bldg., 790 Atlantic Dr., Atlanta, GA 30332. Email: afrancisco8@gatech.edu

[2]City Infrastructure Analytics Director, School of Civil and Environmental Engineering, Georgia Institute of Technology, Mason Bldg., 790 Atlantic Dr., Atlanta, GA 30332. Email: nedam@gatech.edu

[3]Frederick Law Olmsted Professor, School of Civil and Environmental Engineering, Georgia Institute of Technology, Mason Bldg., 790 Atlantic Dr., Atlanta, GA 30332 (corresponding author). ORCID: https://orcid.org/0000-0002-8949-3248. Email: jet@gatech.edu

Note. This manuscript was submitted on April 3, 2019; approved on August 26, 2019; published online on November 30, 2019. Discussion period open until April 30, 2020; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Management in Engineering***, © ASCE, ISSN 0742-597X.

of buildings can enable the construction of benchmarks developed at finer temporal scales and across specific time segments, which we define as temporally segmented building energy benchmarks. Because many buildings require different levels of energy consumption based on the time of day or week, temporally segmented building energy benchmarks have potential to provide a more accurate measure for building efficiency across a group of buildings (Francisco et al. 2018; Roth and Jain 2018). This paper leveraged smart meter data for a community of buildings to benchmark energy consumption in different periods and determined that a digital-twin-enabled energy management platform that constitutes temporally segmented building energy benchmarks can help detect previously undiscoverable insights and identify more specific time-driven strategies for near real-time urban energy management.

Background

Energy Performance Assessments in Existing Buildings

Energy performance assessments commonly are used to assess the energy performance of existing buildings, and can be divided into two categories: building energy classification and energy diagnosis (Wang et al. 2012). Building energy classification is applied across a group of buildings and adopts methodologies that standardize each building's annual energy consumption based on its characteristics, enabling comparison of building energy efficiency levels between buildings with different characteristics (e.g., floor area and space use) (Pérez-Lombard et al. 2009). One of the most commonly used standardization processes within energy classification is building energy benchmarking, which is a top-down approach applying statistics or machine learning algorithms (Buck and Young 2007; Kavousian and Rajagopal 2014; Zhang et al. 2015) that classify whole-building energy efficiency levels on an annual basis (Borgstein et al. 2016; Kinney and Piette 2002). Such macrolevel metrics are generally easily understood and aim to connect building owners, policymakers, and the public with the information necessary to identify poorly performing buildings and motivate stakeholders to implement energy efficiency improvements (Wang et al. 2012). In general, these approaches require little data on building technology (Zhao et al. 2017) or the physical characteristics of a building (Li et al. 2014; Hong et al. 2012), which is advantageous in large-scale studies across a community or city in which the availability of such data in most existing buildings is limited. Although benchmarking techniques help identify overall building performance, their findings are not specific enough to guide operational or physical improvement recommendations for a building (Borgstein et al. 2016). To identify the root causes of building inefficiencies, energy diagnosis methodologies often are required (Borgstein et al. 2016).

In contrast to building energy benchmarking, energy diagnosis methodologies use bottom-up approaches to help identify where energy inefficiencies exist in a building. Common methodologies for diagnosis include energy simulations and engineering calculations (Borgstein et al. 2016; Burman et al. 2014). Most commercially available energy modeling software requires building geometries, physical properties, thermal zones, and operational inputs to simulate building energy consumption (Burman et al. 2014). Additionally, engineering calculation methodologies, often accomplished through energy audits, use building system specifications and operational schedules to predict energy consumption and efficiency levels of separate building systems (Burman et al. 2014). Diagnosis methodologies provide results that are easily translated into specific steps to improve building energy performance, and thus overcome shortcomings of the energy benchmarking methodologies mentioned above. However, they are limited for the following reasons. First, although research has made substantial progress in the reconciliation of estimated energy consumption from bottom-up methodologies and measured consumption, their accuracy is not yet consistently reliable, and the reconciliation process is time-intensive, computationally expensive, and requires building science expertise (Wang et al. 2012). Second, bottom-up approaches rely on extensive and accurate data collection of building systems on site or through detailed review of building plans (Borgstein et al. 2016). Because municipalities often are constrained by budgets and manage diverse buildings, the extensive time and experience required to comprehensively assess citywide building performance using bottom-up methodologies is not practical.

Given the advantages and disadvantages of energy benchmarking and diagnosis approaches, community- or city-level building energy assessments could be improved by providing more actionable and system-specific efficiency indicators (i.e., the results of energy diagnosis) while still applying methodologies that do not require extensive data collection, time, or expertise (i.e., the methodologies of energy benchmarking). One promising research avenue aiming to develop more actionable results while leveraging the scalability of benchmarking methods is applying smart meter data to examine building energy benchmarks segmented by time. As real-time energy management tools become integral to energy efficiency decision making (Ramachandra et al. 2018; Kitchin et al. 2015), energy benchmarks segmented by time have immense potential to further the efficacy of digital twin–enabled energy management platforms if they can provide more insights, value, and frequent feedback compared to their annual counterparts.

Temporal Dimension of Building Energy Performance

Although building characteristic data (e.g., wall insulation levels, HVAC equipment types, number of appliances) are challenging to collect at scale, energy data recording electricity use at granular levels (meter readings less than once per hour) is becoming increasingly accessible through advanced smart metering infrastructure (EIA 2018). The availability of smart meter data has spurred a variety of research assessing how these data, combined with statistical or machine learning algorithms, can support a wide range of applications including energy load analysis, forecasting, and management (Wang et al. 2018). Broadly defined as smart meter data analytics, such research enables near real-time analyses of energy use (Wang et al. 2018), which previously was not feasible using traditional energy meters. Previous literature has used smart meter data extensively across various applications to further real-time analytics. For example, these data were used to detect anomalous consumers in real-time using decision trees and support vector machine (SVM) classifiers (Jindal et al. 2016), to understand energy behaviors of commercial building occupants using a *k*-nearest neighbors classifier (Rafsanjani et al. 2018), and to optimize real-time energy pricing using innovative clustering techniques (Joseph and Erakkath Abdu 2018). In general, smart meter data analytics applications have focused on consumer segmentation, prediction, and demand response applications. Few studies have examined the potential of smart meter data analytics applied to building benchmarking and performance applications, particularly for real-time use across larger scales of buildings.

A few recent studies have integrated smart meter data into building benchmark analyses (Francisco et al. 2018; Grolinger et al. 2018; Roth and Jain 2018). The general approach involves leveraging the granularity offered by smart meter data to segment

J. Manage. Eng., 2020, 36(2): 04019045

building efficiency levels by different periods to help enlighten more-specific areas for efficiency improvement, which can be applied across a large scale of buildings. This approach stands in contrast to traditional building energy benchmarks, which are calculated on an annual basis and offer little insight into how to improve building performance. Grolinger et al. (2018) calculated building energy benchmarks during periods in which events occurred in two sports arenas to identify the most and least efficient events. The purpose of these benchmarks was to help identify efficiency opportunities and where to prioritize efficiency efforts across the two sports arenas. In a comparable analysis, Roth and Jain (2018) expanded the scope of buildings included to assess 500 K-12 school buildings. They benchmarked energy use segmented by operational and nonoperational periods and suggested potential targeted areas for efficiency opportunities based on visual analysis of these metrics. The results of both of these studies suggested that differences exist between temporally segmented building energy benchmarks, and the likely potential for these to enlighten efficiency opportunities across a group of buildings. However, these analyses were based on the assumption that temporally segmented benchmarks differ systematically from their annual counterparts, whereas the magnitude, distribution, and statistical significance of the differences between temporally segmented and annual benchmarks have yet to be examined, which may compromise the validity of such an approach.

With research only beginning to examine the development and utility of temporally segmented energy benchmarking, this paper contributes to this work by investigating the research question of whether temporally segmented building energy benchmarks differ from their nontemporally segmented counterparts. Furthermore, if this is the case, how can temporally segmented building energy benchmarks be used to improve energy management across a portfolio of buildings? To evaluate these questions, we used smart meter electricity data across a community of buildings to develop daily benchmarks that are segmented by strategic periods and statistically evaluated their deviations from a control group, their nontemporally segmented counterparts. Following this analysis, we discuss in detail the practical implications of these findings, particularly through a smart city digital twin lens. We explain how these results establish new capacities for digital twin–enabled energy management through several examples that show the potential of temporally segmented energy benchmarks to better inform energy-efficiency decision making by enabling (1) identification and prioritization of specific retrofit strategies, and (2) near-real-time building energy management. Both implications propel the development of digital twin–enabled energy management systems, which require informative metrics to be successful in supporting building operators or portfolio owners with energy decisions.

Methods

The Georgia Institute of Technology (GT) university campus was selected as a test bed to quantify building energy-efficiency scores' temporal variation within a community. University campuses have diverse and dynamic operations, consisting of a combination of offices, laboratories, recreation, health, food, retail, and classroom facilities that are comparable to the operations of a small town or community (Klein-Banai and Theis 2011). The data scope of this analysis covered building-level electricity consumption for 38 buildings on the GT campus. Although the buildings in this sample have diverse functions, they all have heating and cooling provided by a district water loop. Buildings with individual electric cooling or heating systems were eliminated from the sample to avoid bias in efficiency scores due to unequal end uses. Average power was recorded in 15-min increments, and the experimental data set ranged from September 26, 2015, to September 25, 2016. The following subsections describe the methods used to segment building energy consumption by period, compute daily energy benchmarks for each building, and conduct a series of hypotheses tests.

Data Segmentation by Period

To establish building energy benchmarks for each building across each period, electricity use for each building first was segmented by the following periods: occupied periods during the school year (A), unoccupied periods during the school year (B), occupied periods during the summer (C), unoccupied periods during the summer (D), peak summer periods (E), and the total period (Total). Although additional periods could be selected and assessed, the periods listed were selected based on their alignment with operational shifts that buildings commonly undergo, and are supported by the literature in having a high potential to enlighten efficiency opportunities. For example, studies have found that substantial energy waste occurs in buildings during unoccupied periods due to misaligned operational schedules (Gul and Patidar 2015; Masoso and Grobler 2010). Knowing this, it likely is helpful to differentiate between efficiency levels during occupied and unoccupied states in targeting efficiency opportunities. Building operations also can shift seasonally. One study found that annual energy-efficiency scores for university buildings were skewed due to significant operational shifts during the summer months, and recommended separating summer months from annual efficiency scores to uncover actual efficiency levels (Tu 2015). Furthermore, energy consumption during summer peak demand periods is a pressing concern for facility managers due to utility peak demand charges (Neufeld 1987). Specific retrofit opportunities exist to reduce energy demand during summer peak periods, such as improving air conditioner efficiency (Yarbrough et al. 2015).

Table 1 outlines the times and days included in each period specified previously, which encompasses a 1-year period in total. The start of the 1-year period was selected strategically to minimize the number of data gaps in the data set. In determining occupancy states, measured occupancy levels were not available for every building in the sample, and building occupancy states were assumed to be occupied between 8 a.m. and 8 p.m. each weekday. These hours reflect when building entrances on campus typically are open/unlocked (~8 a.m.) and when they start to require key

Table 1. Segmented period details

Period	Days	Times	Total number of days
Occupied during school year (A)	9/26/15–5/7/16, 8/21/16–9/25/16	8 a.m.–8 p.m. (M–F)	174
Unoccupied during school year (B)	9/26/15–5/7/16, 8/21/16–9/25/16	8 p.m.–8 a.m. (M–F)	174
Occupied during summer (C)	5/8/16–8/20/16	8 a.m.–8 p.m. (M–F)	75
Unoccupied during summer (D)	5/8/16–8/20/16	8 p.m.–8 a.m. (M–F)	75
Peak summer (E)	9/26/15–9/30/15, 6/01/16–9/25/16	2 p.m.–7 p.m. (M–F)	86
Total	9/26/15–9/25/16	12 a.m.–11:59 p.m.	365

access (~8 p.m.). Hours after 8 p.m. and before 8 a.m. were assumed to have low to no occupancy, and are referred to as the unoccupied state in this paper. Weekends were excluded from the analysis because occupancy states during weekends are less consistent, and therefore estimates would be less reliable. Additionally, the 2-week period encompassing the winter break (between December 19, 2015, and January 3, 2016) was removed from the analysis due to unknown occupancy states. The seasonal shifts were divided between the school year and summer, as defined by the GT academic calendar for the 2015–2016 school year. The summer peak demand time range was based on when Georgia Power, the power supplier of the campus, charged customers peak billing demand rates during the summer.

Next, the energy data for each period were aggregated from 15-min interval data to daily energy use values, referred to in this paper as the temporally segmented data, in order to perform energy benchmarking on a daily basis. To aggregate energy data to the daily level, for each day the average of the 15-min interval energy data was calculated. This procedure was followed for Periods A, B, C, D, and Total. Because peak period charges are based on the maximum kilowatt use in a 30-min period, the data during the peak summer period (E) were averaged using a 30-min running average, and the maximum value for the day was selected. The total period energy data represented the average electricity use across all 24 h of each day, and served as the control. After following energy benchmarking procedures, detailed subsequently, energy efficiency scores for each building across each period were generated. Temporally segmented and total period building energy daily efficiency scores were compared for each building, meaning that 38 comparisons were conducted for each period. The hypotheses tested were as follows:

- Hypothesis A (1–38): A building's daily efficiency scores during the total period and its daily efficiency scores during occupied periods in the school year are not the same.
- Hypothesis B (1–38): A building's daily efficiency scores during the total period and its efficiency scores during unoccupied periods in the school year are not the same.
- Hypothesis C (1–38): A building's daily efficiency scores during the total period and its efficiency scores during occupied periods in the summer are not the same.
- Hypothesis D (1–38): A building's daily efficiency scores during the total period and its efficiency scores during unoccupied periods in the summer are not the same.
- Hypothesis E (1–38): A building's daily efficiency scores during the total period and its efficiency scores during summer peak demand periods are not the same.

Efficiency Score Development

The purpose of building benchmarking is to generate building efficiency scores that enable more-accurate comparisons of energy efficiency between buildings. Early benchmarking methods created simple ratio metrics, such as energy use per area or energy use per occupant, which normalized energy use based on a single building characteristic (Pérez-Lombard et al. 2009). Subsequently, more-complex approaches such as statistical or machine learning methods were introduced, which attempted to normalize energy use based on more than one building characteristic. Using such approaches, an efficiency score is generated by adjusting a building's energy use to account for multiple building characteristics simultaneously. This study adopted the regression-based methodology developed by Chung et al. (2006). Such an approach is similar to methodologies commonly used by industry benchmarking applications such as the Energy Star score (Borgstein et al. 2016; Shrestha and Prajakta 2013). The following paragraphs detail the steps taken to apply this methodology to generate building efficiency scores for each day within each period.

First, the dependent and independent variables for the benchmark model were defined. The independent variables represented explanatory variables of energy consumption, which were used to normalize energy use across different buildings (Chung et al. 2006). These variables encompassed building characteristics that were outside the control of the building operators or occupants, to normalize building energy use based on the building features that are unlikely or infeasible to change. For example, different building space types (e.g., lab and office spaces) have different energy requirements for operation (Park et al. 2016). Normalizing building energy use by the area of different space types can enable more-accurate comparison of the energy use of buildings with different functions. Other inflexible building characteristics (i.e., characteristics that cannot be easily changed by management or occupants) were collected from a publicly accessible database from the university's Capital Planning and Space Management group. All explanatory variables and the dependent variable used in the model are outlined in Table 2. Space type areas were converted to building use ratios (BURs), ranging between 0 and 1, by dividing the area of the space type by the total building floor area (Park et al. 2016). In addition, the dependent variable for each period was divided by building floor area to generate energy use intensities (EUIs), which are the primary unit for energy benchmarking analyses because building floor area is highly correlated with energy use (Sharp 1995). Similar to the approaches of previous benchmarking studies (Buck and Young 2007; Chung 2012; Park et al. 2016), floor area also was included as an independent variable. Figures with the distributions of the explanatory variables are provided in the Supplemental Data (Fig. S1).

Next, temporally segmented daily EUIs for each building were normalized by the identified explanatory factors using a multivariate linear-regression approach, similar to that used by Chung et al. (2006). The first step of the normalization process was to create a model to quantify the relationship between the building explanatory factors and the EUIs. A regression model was created for each day within each period. For example, 174 models were created for Period A (Table 1). Several transformations were made to the independent and dependent factors. To account for skewed distribution characteristics, EUI and floor area model inputs were

Table 2. Explanatory and dependent variable descriptions

Variable	Feature
Independent variables	Floor area (ft^2)
	Building age (years)
	Years since renovation (years)
	Number of floors (floors)
	Percent renovated (%)
	BUR: laboratory (%)
	BUR: office (%)
	BUR: mechanical (%)
	BUR: general (%)
	BUR: circulation (%)
	BUR: service (%)
	BUR: supply (%)
	BUR: classroom (%)
	BUR: study (%)
	BUR: special (%)
Dependent variable	Daily average EUI (Total period, A–E periods) ($kWh/ft^2/day$)

Note: BUR = building use ratio.

log-transformed. The explanatory variables in Table 2 were rescaled to have a mean of 0 and standard deviation of 1 in order to aid the interpretation of the regression coefficient results. The rescaled explanatory variables served as regression model inputs $(x_1, \ldots, x_p)$. For each day within each period, the regression model took the following form:

$$\mathrm{EUI} = a + b_1 x_1^* + \cdots + b_k x_k^* + \varepsilon \quad (1)$$

where a = intercept; ε = error term; $x_1^*, \ldots, x_k^*$ = significant explanatory variables (where $k \leq p$); and $b_1, \ldots, b_k$ = model coefficients. Forward selection was applied to identify the significant explanatory values included in the model, with the Akaike information criterion providing the basis for variable selection. Based on the regression model results for a particular day, the daily EUI for a building can be normalized as follows:

$$\mathrm{EUI}_{\mathrm{norm}} = \mathrm{EUI}_0 - \mathrm{EUI} + a \quad (2)$$

where EUI_0 = measured EUI of the building for that day; EUI = predicted EUI based on the model coefficients from Eq. (1); a = model intercept in Eq. (1); and $\mathrm{EUI}_{\mathrm{norm}}$ = building's normalized EUI. Eq. (2) is equivalent to calculating the residual for a building $(\mathrm{EUI}_0 - \mathrm{EUI})$ and adding this to the model intercept. The model intercept a represents the EUI for the average building in the data set, because the explanatory variables are rescaled with a mean of 0. To illustrate the normalized EUI calculation, if the expected EUI for a building (EUI) is less than the actual EUI (EUI_0), this difference is added to the average EUI across the buildings (a), leading to a higher normalized EUI ($\mathrm{EUI}_{\mathrm{norm}}$). A regression model was created for each day within each period, and thus the explanatory variables, coefficients, and intercept values in Eqs. (1) and (2) changed from day to day across periods.

After normalizing each building's EUI, the $\mathrm{EUI}_{\mathrm{norm}}$ values were rescaled between 0 and 1 for each day. These represented the efficiency scores, where 0 was the least efficient and 1 was the most efficient. The result was a distribution of daily efficiency scores for each building within each period.

To compare the efficiency score distributions and test the hypotheses, the Wilcoxon signed-rank test was used, which is a nonparametric test that accepts ordinal data, such as energy efficiency scores. The Wilcoxon signed-rank test is a paired test, and thus a building's efficiency score during the total period was compared with the temporally segmented period for the same day. Because the total period had more days than the temporally segmented periods, efficiency scores were included in the statistical analysis only if they both were computed on the same day. Each hypothesis was tested for each building, introducing the multiple comparison problem, which increases the probability of committing a Type I error (Bretz et al. 2010). Therefore, adjusted p-values were computed to control for familywise errors using the Holm procedure. The Holm procedure is a more powerful modification of the Bonferroni correction, and was used because the Bonferroni often is regarded as too conservative when a large number of tests are conducted (Holm 1979). Confidence intervals of 95% or greater indicated statistically distinct efficiency score medians. Efficiency score development and statistical analysis were calculated using R version 3.5.1.

Results

The regression model coefficients for each day were used to normalize building daily EUI and to generate energy efficiency scores. To test Hypothesis A1, the distribution of daily efficiency scores for Building 1 during occupied hours in the school year was compared with the distribution of daily efficiency scores for the same building during the total period using the Wilcoxon signed-rank test (i.e., a paired difference test). This was repeated for all 38 buildings in Hypothesis A. Next, this process was repeated for the remaining temporally segmented periods. Adjusted p-values, to account for the multiple comparison problem, are displayed in Table 3. At the 95% confidence interval, there was enough evidence to reject the null hypothesis for 34 buildings in Hypothesis A, 32 buildings in Hypothesis B, 31 buildings in Hypothesis C, 30 buildings in Hypothesis D, and 32 buildings in Hypothesis E. Column Total significant cases in Table 3 lists the total number of times the null hypotheses were rejected for each building. Almost 75% of the buildings ($n = 28$) had a statistically distinct distribution of efficiency scores in four or more of the examined periods compared with the total period. All of the buildings had a statistically significant difference for at least one period. The p-values equivalent to 1 were the result of applying the Holm procedure, which adjusted the p-values to be more conservative (in terms of Type I errors) in order to account for the multiple hypothesis problem.

Fig. 1 summarizes the magnitude of absolute differences between total and temporally segmented efficiency scores for all statistically significant buildings. Among the significant buildings, the mean absolute difference between the total and temporally segmented efficiency scores during Period A was 0.079, whereas the maximum absolute difference was 0.483. This implies that across the buildings with statistically significant efficiency scores differences, daily efficiency scores during occupied periods in the school year differed by an average of 7.9% from the total efficiency score, and by as much as 48.3%. These magnitudes changed by period (Fig. 1). These differences are in comparison with the total period, which is representative of the average efficiency score between all periods. Thus, computing the differences between efficiency scores of two temporally segmented periods many times resulted in even larger differences.

Discussion

This study developed building energy benchmarking scores segmented by strategic periods and statistically assessed how they vary from conventional, total benchmarking scores. Daily scores that vary systematically from total scores can enable more real-time and more-informative predictions to guide operational decision making. Leveraging individual building smart meter data across a portfolio of buildings enabled the development of daily, temporally segmented benchmarking scores, and the results of the statistical analysis showed that temporally segmented building energy benchmarks were significantly distinct from their total counterparts for the vast majority of buildings in the sample (between 30 and 34 of the 38 buildings in the portfolio). Previous studies pointed out that conventional building energy benchmarks are limited in their ability to help target areas for efficiency improvement (Borgstein et al. 2016). Although recent work has leveraged smart meter data to explore the potential of temporally segmented energy benchmarks in gaining more-specific efficiency insights (Francisco et al. 2018; Grolinger et al. 2018; Roth and Jain 2018), the deviations between this novel technique and their conventional counterparts had not previously been statistically assessed. This study contributes to emerging work examining the temporal dimensions of energy benchmarking (Francisco et al. 2018; Grolinger et al. 2018; Roth and Jain 2018) by assessing the statistical significance and magnitude of differences between temporally segmented and conventional benchmarks for a community of buildings. In addition, this

Table 3. Adjusted p-values for Hypotheses A, B, C, D, and E

Building	A	B	C	D	E	Total significant cases
1	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	5
2	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	5
3	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	0.002[b]	5
4	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	5
5	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	5
6	<0.001[a]	0.005[b]	0.048[c]	<0.001[a]	<0.001[a]	5
7	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	5
8	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	5
9	0.021[c]	<0.001[a]	<0.001[a]	<0.001[a]	0.03[c]	5
10	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	5
11	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	5
12	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	5
13	<0.001[a]	0.034[c]	<0.001[a]	<0.001[a]	<0.001[a]	5
14	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	5
15	<0.001[a]	<0.001[a]	<0.001[a]	0.004[b]	<0.001[a]	5
16	<0.001[a]	<0.001[a]	0.006[b]	<0.001[a]	0.026[c]	5
17	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	5
18	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	5
19	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	5
20	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	5
21	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]	5
22	0.005[b]	0.005[b]	<0.001[a]	<0.001[a]	0.728	4
23	1	<0.001[a]	<0.001[a]	0.001[b]	<0.001[a]	4
24	<0.001[a]	0.005[b]	<0.001[a]	1	<0.001[a]	4
25	<0.001[a]	<0.001[a]	<0.001[a]	0.085	<0.001[a]	4
26	<0.001[a]	1	<0.001[a]	<0.001[a]	<0.001[a]	4
27	1	<0.001[a]	<0.001[a]	<0.001[a]	0.002[b]	4
28	<0.001[a]	<0.001[a]	0.959	<0.001[a]	<0.001[a]	4
29	0.001[b]	0.309	0.959	0.001[b]	<0.001[a]	3
30	<0.001[a]	0.473	0.003[b]	<0.001[a]	0.728	3
31	<0.001[a]	0.003[b]	0.009[b]	1	0.728	3
32	<0.001[a]	<0.001[a]	0.312	1	0.011[c]	3
33	0.006[b]	<0.001[a]	0.959	<0.001[a]	1	3
34	<0.001[a]	0.004[b]	0.006[b]	1	0.292	3
35	<0.001[a]	0.075	<0.001[a]	1	<0.001[a]	3
36	0.687	<0.001[a]	0.445	<0.001[a]	1	2
37	<0.001[a]	0.063	0.443	1	<0.001[a]	2
38	0.773	1	0.765	0.264	0.03[c]	1
Total significant cases	34	32	31	30	32	

[a]$p < 0.001$.
[b]$p < 0.01$.
[c]$p < 0.05$.

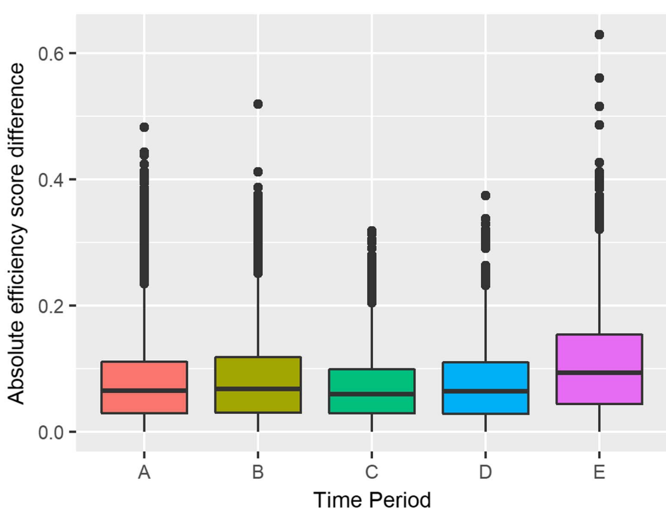

Fig. 1. Absolute difference between total period efficiency scores and segmented period efficiency scores for statistically significant buildings.

work furthers smart meter analytics research by documenting how this area can be integrated with and applied to building energy benchmarking methods. This is a critical step to understand deviations in building performance throughout the day and year relative to a community of buildings, which importantly has several practical implications for urban energy management.

The results of daily and temporally segmented benchmarks detect performance variations across time and have the potential to support with targeting, prioritizing, and managing individual building efficiency opportunities across a large geographic scale of buildings. Digital twin–enabled energy management platforms are envisioned to stream energy data sources (e.g., smart meter data) and are conceptually a promising platform to support cities with building portfolio performance assessments and urban energy management (Mohammadi and Taylor 2017). The construction of dynamic and real-time metrics that transform smart meter data into useful information is an integral element for digital twin-energy management to be successful. A smart city digital twin energy management platform built around temporally segmented building energy benchmarks (Fig. 2) offers the potential for managers to identify and prioritize specific retrofit strategies and

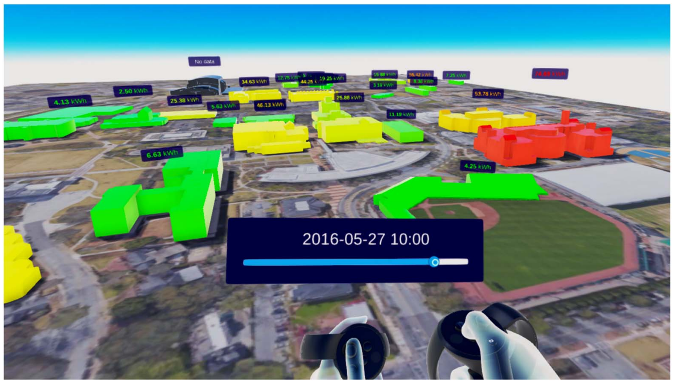

Fig. 2. Digital twin–enabled energy management platform.

detect near-real-time deviations in building efficiency in the context of the performance of the entire building portfolio. The following two sections provide specific examples of how temporally segmented building energy benchmarks facilitate both energy efficiency prioritization and near-real-time decision making.

Prioritization of Specific Retrofit Strategies across Buildings

Temporal fluctuations in efficiency scores can enlighten how to prioritize building efficiency improvements. Fig. 3 illustrates the difference between total and summer peak daily efficiency scores across the temporal state summer peak demand for four buildings. For building 18 the total daily efficiency scores were consistently higher (i.e., more efficient) than the peak summer efficiency scores, as indicated by the "+" signs in the graph. This distinction is necessary when prioritizing efficiency improvements, because specific measures are appropriate for reducing energy use during summer peak periods, such as increasing air conditioner efficiency or peak-load shifting (Koomey and Brown 2002). If a building manager considered only the total efficiency score for Building 18, this building would appear to be more efficient than its actual performance during summer peak demand hours. Specifically, the total efficiency scores masked the peak efficiency scores with differences up to 34.1%. This building's efficiency score rank during summer peak demand periods can help building managers decide whether to invest in peak demand reduction improvements with this building.

On the other hand, different trends in summer peak demand efficiency scores were found in Buildings 20, 6, and 8. Building 20 had summer peak demand efficiency scores that often were more efficient than the total efficiency score. This indicates that this building is more efficient during summer peak demand periods compared with the total period, and likely should have lower prioritization for resources allocated to reduce peak demand. Alternatively, Building 8 transitioned from predominantly positive scores to predominantly negative scores, whereas Building 6 had more-mixed variations. These trends show that energy performance during peak periods relative to total periods changed sharply from day to day or week to week. In buildings in which these changes were more extreme (e.g., indicated by darker shades), operational measures such as reviewing the building management system or other automated controls should be investigated to determine if the programmed operations still reflect the actual building conditions.

Near-Real-Time Energy Management across Buildings

The previous example considered deviations between the total and temporally segmented efficiency scores to compare efficiency opportunities between buildings. Fig. 4 presents the raw efficiency

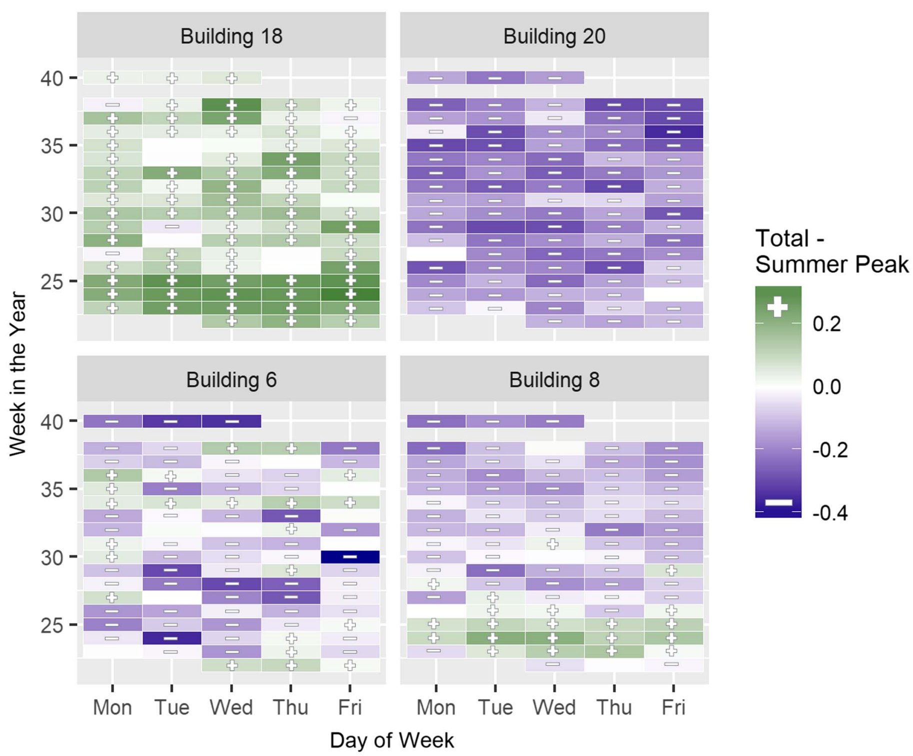

Fig. 3. Difference in daily efficiency scores for summer peak periods.

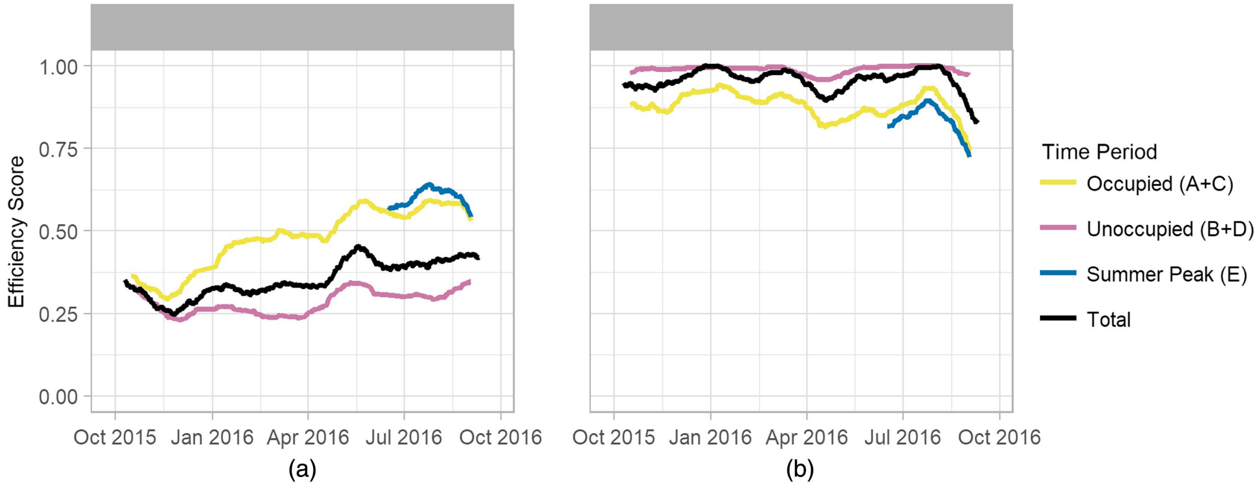

Fig. 4. 30-day moving average of raw efficiency scores for (a) Building 14; and (b) Building 4.

scores across periods to show how building performance changed over time within buildings. This can help generate insights with several uses, including identifying buildings with sudden changes in performance and buildings with consistently low levels of performance, and helping to demonstrate how temporally segmented benchmarks can support more real-time energy management. Buildings with sudden changes in performance may simply require a review and update to operational controls, whereas buildings with continuously low performance may require investment in more-capital-intensive upgrades. Fig. 4 shows lines for only three segmented periods because the occupied and unoccupied lines contain both the school year and summer periods.

Fig. 4 shows the 30-day moving average of the raw efficiency scores for two buildings across the year. For Building 14, the occupied and unoccupied efficiency scores were similar to the total score for the first 2 months. In late December, the building's efficiency score during occupied periods improved, whereas the efficiency score during unoccupied periods decreased. This gap remained throughout the remainder of the year. This relatively quick change in efficiency scores can indicate that an operational shift has occurred in the building, causing it to perform worse during unoccupied hours. Previous studies highlighted how misalignment of building automation systems (Gul and Patidar 2015) and poor occupant behaviors (Masoso and Grobler 2010) can reduce building efficiency during unoccupied hours. Efficiency scores during unoccupied periods can highlight which buildings to prioritize efforts to review automation schedules and implement behavior change campaigns.

On the other hand, Building 4 performed consistently very well during unoccupied periods, with efficiency scores above 0.90. The building's occupied efficiency scores were about 10% lower than the unoccupied scores for the first half of the year, and this difference increased in the last 2 months of the year. The summer peak demand scores followed the same trends as the occupied periods during the summer, which was expected because their specified times during the day were very similar. For this building, this could indicate that improvements aimed at decreasing energy use during occupied hours, particularly summer peaks, may be most appropriate. In contrast to unoccupied-period efficiency measures, types of efficiency measures to address occupied consumption include more-capital-intensive efforts, such as retrofitting light fixtures and air conditioners and installing demand-controlled ventilation (Koomey and Brown 2002).

External Validation

The distribution of the regression models' fit was computed by period. The means of the adjusted R-squared values for each segmented period ranged between 0.72 and 0.80. These values are consistent with the fit of regression models in other regression benchmarking studies (Buck and Young 2007; Chung et al. 2006; Xuchao et al. 2010). Fig. S2 provides density plots showing the distribution of the models' fit for each period.

The regression coefficients for each daily EUI model resulted in a distribution of coefficients for each explanatory variable. Statistically significant explanatory variables (p-values < 0.05) indicate key drivers of energy consumption across the group of buildings. Table 4 lists the frequency with each explanatory variable was significant. Of the 15 explanatory variables, 14 variables significantly impacted the daily EUI for at least 1 day. Although all the variables were included in energy benchmark calculations, the following external validation discussion focuses on the first seven variables in Table 4, which were frequently significant (significant for >75% of the models) for at least three of the periods. The other variables were less frequently significant in the models; building age was significant for 22%–57% of the models, depending on the period, and the remaining seven variables were consistently insignificant (significant for <25% of the models) across all periods.

The majority of the directions of the significant coefficients aligned with previous studies and expectations. Fig. S3 shows density plots visualizing the regression coefficient distributions for all explanatory variables. Buildings with a higher percentage of area dedicated to laboratory or mechanical room space had higher EUIs. Such spaces have energy-intensive equipment, such as lab ventilation hoods in laboratories or IT equipment in mechanical rooms. Higher percentages of office space also were associated with higher EUIs, which aligns with findings from previous studies (Park et al. 2016). Circulation space (e.g., hallways and lobbies) was associated with lower EUIs, which makes sense because these spaces typically have less-consistent occupancy loads and more open space without energy-intensive equipment. In addition, buildings that had not been renovated recently had higher EUIs, which can be attributed to having older, less-efficient equipment such as lighting, plug loads, and mechanical systems.

For some coefficients, the relationship was not consistently supported by the literature. Floor area was positively associated with EUI, which aligns with some studies (Park et al. 2016) and

Table 4. Frequency of regression coefficient statistical significance (%)

Variable	Total ($n = 365$)	A ($n = 174$)	B ($n = 174$)	C ($n = 75$)	D ($n = 75$)	E ($n = 86$)
BUR: service	100	100	100	100	100	100
Floor area	99	98	99	100	100	100
BUR: laboratory	94	98	96	93	91	95
BUR: circulation	91	67	91	72	100	24
BUR: office	88	93	84	95	91	90
BUR: mechanical	85	89	84	97	88	100
Years since renovation	83	62	83	99	87	65
Building age	39	22	39	57	52	33
BUR: classroom	21	9	15	23	24	8
BUR: supply	18	6	16	12	32	10
BUR: general	11	3	4	12	21	9
BUR: study	9	2	2	13	19	5
BUR: special	1	10	1	0	0	14
Number of floors	1	2	1	0	0	0
Percent renovated	0	0	0	0	0	0

Note: n = number of days in period.

contradicts others (Chung et al. 2006). Other coefficients have yet to be examined in the context of energy benchmarking. Interestingly, service space (e.g., bathroom and janitorial areas) was positively associated with EUI. To our knowledge, previous energy benchmarking studies have not documented this variable's association with energy use. Possible drivers of energy consumption in service spaces include hot water energy use, ventilation loads, and cleaning equipment.

Limitations and Future Directions

Several limitations exist for this study, prompting avenues for future research. Common across building energy benchmarking studies, it is challenging to determine how well the benchmarking indicators agree with the actual efficiency levels of the buildings, particularly when such benchmarks are developed across large scales of buildings. In a similar vein, it is difficult to determine if the efficiency recommendations informed by temporally segmented benchmarks are the optimal efficiency improvements for the building. Future research will dig deeper into this by investigating the effect of real retrofits or operational changes on daily benchmarking results. This analysis will compare the computed daily benchmark with an associated operational or capital change to examine whether the daily benchmark trends follow the expected pattern based on the particular retrofit install or operational change made.

In addition, regression-based benchmarking techniques, including the methodology developed in this study, assume that the regression residuals reflect only building inefficiencies, whereas in reality they contain statistical noise, measurement error, and unexplained factors (Chung 2011). However, other benchmarking methodologies have other inherent limitations, such as large sensitivities to outliers and loss of physical meaning (Borgstein et al. 2016). The aim of this study was to apply existing benchmarking techniques to assess deviations between energy benchmarks during different periods, and we opted to apply regression-based techniques due to their high interpretability (i.e., the results have physical meaning as it relates to the building) and common adoption in industry applications, such as the Energy Star score. Future studies could apply other benchmarking techniques to assess the consistency of the results across different methodologies. This is particularly relevant for benchmarking analyses performed at different scales, such as daily, weekly, or monthly, to assess the robustness and sensitivity of different techniques. Furthermore, building occupancy states were estimated, because measured data were not available for all buildings. Thus, estimated occupancy states may not reflect the actual occupancy levels in the buildings. Incorporating data that contain or represent a proxy for actual occupancy within each building into the benchmarking models will enhance the accuracy of the models.

In its current state, our digital twin–enabled energy management system (Fig. 2) demonstrates a proof of concept for the platform. Future work will more deeply integrate smart meter data, temporally segmented energy benchmarks, and other resource data, such as gas, heating, cooling, and water consumption. This effort also will involve user-interface updates and testing with specific user groups (e.g., facility managers) to assess the utility and future direction of the platform.

Conclusion

Approaches for assessing building energy performance diverge in their ability to handle large-scale analyses while still providing specific, actionable findings. Energy benchmarking methodologies can be applied across a large number of buildings; however, they provide narrow insights and are limited in their ability to identify specific areas for efficiency improvement (Borgstein et al. 2016). Conversely, energy diagnosis methodologies provide more actionable energy conservation measures, but they are most appropriate at a single-building level and require extensive and accurate data collection (Borgstein et al. 2016). This paper expanded recent top-down, data-driven approaches to building performance assessments (Francisco et al. 2018; Grolinger et al. 2018; Roth and Jain 2018) by creating temporally segmented, daily building energy benchmarks and evaluating their statistical deviations from conventional energy benchmarks.

Our findings demonstrate that across all of the buildings in the sample, temporally segmented energy efficiency scores were statistically distinct from efficiency scores during the total period for at least one period. For the vast majority of buildings, such scores were statistically distinct for at least four of the five periods. This indicates that although a building may rank as efficient overall, it is not necessarily efficient during certain periods, and that a building that is not efficient overall may in fact be efficient during certain periods. Thus, total efficiency scores mask underlying periods of inefficiencies or efficient performance. In addition, we established that efficiency scores fluctuate not only between periods, but also within periods. This understanding is crucial for near-real-time operational decision making and management. Fluctuations in energy efficiency throughout the year indicate whether a building is consistently performing well, is consistently underperforming, or if a sudden change in performance has occurred. This is a crucial distinction that can support decision makers in developing strategies for whether to investigate operational procedure modifications or opportunities for more-capital-intensive investments.

Overall, these results expand the usability and accuracy of traditional building energy benchmarking approaches. Temporally segmented daily efficiency metrics integrated into digital twin–enabled energy management platforms can transform approaches to energy management across a portfolio of buildings. This is of critical importance as cities are working under limited budgets to make substantial reductions in building energy consumption and strive toward smarter operations. Temporally segmented building energy benchmarks give new insights using building benchmarking techniques to enable more systematic, real-time, and accurate management of city-scale building energy consumption and help urban areas reach low-carbon energy goals.

Data Availability Statement

Building energy consumption data used during the study were provided by a third party. Direct requests for these materials may be made to the provider as indicated in the Acknowledgments. R programming code and files generated during the study are available from the corresponding author by request.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant Nos. DGE-1650044 and CPS-1837021. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors thank Jessica Rose and others at the Georgia Tech Facilities Management Office for data access and other input.

Supplemental Data

Figs. S1–S3 are available online in the ASCE Library (www.ascelibrary.org).

References

Baxter, R., N. Hastings, A. Law, and E. J. Glass. 2011. "A theory of smart cities." In *Proc., 55th Annual Meeting of the Int. Society for the Systems Sciences*, 1–15. Pocklington, York: ISSS.

Borgstein, E. H., R. Lamberts, and J. L. M. Hensen. 2016. "Evaluating energy performance in non-domestic buildings: A review." *Energy Build.* 128 (Sep): 734–755. https://doi.org/10.1016/j.enbuild.2016.07.018.

Bretz, F., T. Hothorn, and P. Westfall. 2010. *Multiple comparisons using R*. Boca Raton, FL: Taylor & Francis.

Buck, J., and D. Young. 2007. "The potential for energy efficiency gains in the Canadian commercial building sector: A stochastic frontier study." *Energy* 32 (9): 1769–1780. https://doi.org/10.1016/j.energy.2006.11.008.

Building Rating. 2019. "Sharing transparency for a more efficient future." Accessed January 30, 2019. https://www.buildingrating.org/.

BuildSmart DC. 2017. "Buildings." Accessed June 15, 2017. http://www.buildsmartdc.com/.

Burman, E., S. M. Hong, G. Paterson, J. Kimpian, and D. Mumovic. 2014. "A comparative study of benchmarking approaches for non-domestic buildings: Part 2—Bottom-up approach." *Int. J. Sustainable Built Environ.* 3 (2): 247–261. https://doi.org/10.1016/j.ijsbe.2014.12.001.

Chung, W. 2011. "Review of building energy-use performance benchmarking methodologies." *Appl. Energy* 88 (5): 1470–1479. https://doi.org/10.1016/j.apenergy.2010.11.022.

Chung, W. 2012. "Using the fuzzy linear regression method to benchmark the energy efficiency of commercial buildings." *Appl. Energy* 95 (July): 45–49. https://doi.org/10.1016/j.apenergy.2012.01.061.

Chung, W., Y. V. Hui, and Y. M. Lam. 2006. "Benchmarking the energy efficiency of commercial buildings." *Appl. Energy* 83 (1): 1–14. https://doi.org/10.1016/j.apenergy.2004.11.003.

EIA (Energy Information Administration). 2018. "How many smart meters are installed in the United States, and who has them?" Accessed June 27, 2019. https://www.eia.gov/tools/faqs/faq.php?id=108&t=3.

Francisco, A., N. Mohammadi, and J. E. Taylor. 2018. "Evaluating temporal shifts in city scale building energy benchmarks." In *Proc., Construction Research Congress 2018*, 450–460. Reston, VA: ASCE.

Grolinger, K., H. F. ElYamany, W. A. Higashino, M. A. M. Capretz, and L. Seewald. 2018. "Energy slices: Benchmarking with time slicing." *Energy Effic.* 11 (2): 521–538. https://doi.org/10.1007/s12053-017-9582-8.

Gul, M. S., and S. Patidar. 2015. "Understanding the energy consumption and occupancy of a multi-purpose academic building." *Energy Build.* 87 (Jan): 155–165. https://doi.org/10.1016/j.enbuild.2014.11.027.

Hastak, M., and C. Koo. 2016. "Theory of an intelligent planning unit for the complex built environment." *J. Manage. Eng.* 33 (3): 1–18. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000486.

Holm, S. 1979. "A simple sequentially rejective multiple test procedure." *Scand. J. Stat.* 6 (2): 65–70. https://www.jstor.org/stable/4615733?seq=1#metadata_info_tab_contents.

Hong, T., H. Kim, and T. Kwak. 2012. "Energy-saving techniques for reducing CO_2 emissions in elementary schools." *J. Manage. Eng.* 28 (1): 39–50. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000073.

IPCC (Intergovernmental Panel on Climate Change). 2007. *Mitigation of climate change: Contribution of working group III to the fourth assessment report of the intergovernmental panel on climate change*. Cambridge, UK: Cambridge University Press.

Jindal, A., A. Dua, K. Kaur, M. Singh, N. Kumar, and S. Mishra. 2016. "Decision tree and SVM-based data analytics for theft detection in smart grid." *IEEE Trans. Ind. Inf.* 12 (3): 1005–1016. https://doi.org/10.1109/TII.2016.2543145.

Joseph, S., and J. Erakkath Abdu. 2018. "Real-time retail price determination in smart grid from real-time load profiles." *Int. Trans. Electr. Energy Syst.* 28 (3): 1–11. https://doi.org/10.1002/etep.2509.

Kavousian, A., and R. Rajagopal. 2014. "Data-driven benchmarking of building energy efficiency utilizing statistical frontier models." *J. Comput. Civ. Eng.* 28 (Feb): 79–88. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000327.

Kinney, S., and M. A. Piette. 2002. "Development of a California commercial building energy benchmarking database." In *Proc., 2002 ACEEE Summer Study on Energy Efficiency in Buildings*, 109–120. Berkeley, CA: Lawrence Berkeley National Lab.

Kitchin, R., T. P. Lauriault, and G. McArdle. 2015. "Knowing and governing cities through urban indicators, city benchmarking and real-time dashboards." *Reg. Stud. Reg. Sci.* 2 (1): 6–28. https://doi.org/10.1080/21681376.2014.983149.

Klein-Banai, C., and T. L. Theis. 2011. "An urban university's ecological footprint and the effect of climate change." *Ecol. Indic.* 11 (3): 857–860. https://doi.org/10.1016/j.ecolind.2010.11.002.

Koomey, J., and R. E. Brown. 2002. "The role of building technologies in reducing and controlling peak electricity demand. Berkeley, CA: Lawrence Berkeley National Laboratory.

Li, Z., Y. Han, and P. Xu. 2014. "Methods for benchmarking building energy consumption against its past or intended performance: An overview." *Appl. Energy* 124 (Jul): 325–334. https://doi.org/10.1016/j.apenergy.2014.03.020.

Masoso, O. T., and L. J. Grobler. 2010. “The dark side of occupants’ behaviour on building energy use.” *Energy Build.* 42 (2): 173–177. https://doi.org/10.1016/j.enbuild.2009.08.009.

Mohammadi, N., and J. E. Taylor. 2017. “Smart city digital twins.” In *Proc., IEEE Symp. Series on Computational Intelligence (SSCI)*, 1–5. Piscataway, NJ: IEEE.

Neufeld, J. 1987. “Price discrimination and the adoption of the electricity demand charge.” *J. Econ. History* 47 (3): 693–709. https://doi.org/10.1017/S0022050700049068.

Park, H. S., M. Lee, H. Kang, T. Hong, and J. Jeong. 2016. “Development of a new energy benchmark for improving the operational rating system of office buildings using various data-mining techniques.” *Appl. Energy* 173 (Jul): 225–237. https://doi.org/10.1016/j.apenergy.2016.04.035.

Pérez-Lombard, L., J. Ortiz, R. González, and I. R. Maestre. 2009. “A review of benchmarking, rating and labelling concepts within the framework of building energy certification schemes.” *Energy Build.* 41 (3): 272–278. https://doi.org/10.1016/j.enbuild.2008.10.004.

Pierce, P., and B. Andersson. 2017. “Challenges with smart cities initiatives—A municipal decision makers’ perspective.” In *Proc., 50th Hawaii Int. Conf. on System Science (HICSS-50)*, 2804–2813. Honolulu: ScholarSpace.

Rafsanjani, H. N., C. R. Ahn, and J. Chen. 2018. “Linking building energy consumption with occupants’ energy-consuming behaviors in commercial buildings: Non-intrusive occupant load monitoring (NIOLM).” *Energy Build.* 172 (Aug): 317–327. https://doi.org/10.1016/j.enbuild.2018.05.007.

Ramachandra, B., P. Nawathe, J. Monroe, K. Han, Y. Ham, and R. R. Vatsavai. 2018. “Real-time energy audit of built environments: Simultaneous localization and thermal mapping.” *J. Infrastruct. Syst.* 24 (3): 4018013. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000431.

Roth, J., and R. K. Jain. 2018. “Data-driven, multi-metric, and time-varying (DMT) building energy benchmarking using smart meter data.” In *Proc., European Group for Intelligent Computing in Engineering (EG-ICE)*, 568–593. Cham, Switzerland: Springer.

Sharp, T. 1995. “Energy benchmarking in commercial office buildings.” In Vol. 4 of *Proc., ACEEE Summer Study on Energy Efficiency in Buildings*, 321–329. Washington, DC: American Council of an Energy Efficient Economy.

Shrestha, P., and K. Prajakta. 2013. “Factors influencing energy consumption of Energy Star and non–Energy Star homes.” *J. Manage. Eng.* 29 (3): 269–278. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000134.

Sierra Club. 2019. “100% commitments in cities, counties, & states.” Accessed March 31, 2019. https://www.sierraclub.org/ready-for-100/commitments.

Tu, K.-J. 2015. “Establishing the DEA energy management system for individual departments within universities.” *Facilities* 33 (11/12): 716–735. https://doi.org/10.1108/F-11-2014-0092.

Wang, S., C. Yan, and F. Xiao. 2012. “Quantitative energy performance assessment methods for existing buildings.” *Energy Build.* 55 (Dec): 873–888. https://doi.org/10.1016/j.enbuild.2012.08.037.

Wang, Y., Q. Chen, T. Hong, and C. Kang. 2018. “Review of smart meter data analytics: Applications, methodologies, and challenges.” In *Proc., IEEE Transactions on Smart Grid*, 1–24. Piscataway, NJ: IEEE.

Xuchao, W., R. Priyadarsini, and L. Siew Eang. 2010. “Benchmarking energy use and greenhouse gas emissions in Singapore’s hotel industry.” *Energy Policy* 38 (8): 4520–4527. https://doi.org/10.1016/j.enpol.2010.04.006.

Yarbrough, I., Q. Sun, D. C. Reeves, K. Hackman, R. Bennett, and D. S. Henshel. 2015. “Visualizing building energy demand for building peak energy analysis.” *Energy Build.* 91 (Mar): 10–15. https://doi.org/10.1016/j.enbuild.2014.11.052.

Zhang, Y., Z. O’Neill, B. Dong, and G. Augenbroe. 2015. “Comparisons of inverse modeling approaches for predicting building energy performance.” *Build. Environ.* 86 (Apr): 177–190. https://doi.org/10.1016/j.buildenv.2014.12.023.

Zhao, D., A. P. McCoy, J. Du, P. Agee, and Y. Lu. 2017. “Interaction effects of building technology and resident behavior on energy consumption in residential buildings.” *Energy Build.* 134 (Jan): 223–233. https://doi.org/10.1016/j.enbuild.2016.10.049.

Zuo, J., B. Read, S. Pullen, and Q. Shi. 2013. “Carbon-neutral commercial building development.” *J. Manage. Eng.* 29 (1): 95–102. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000127.