Applied Energy 276 (2020) 115435

Contents lists available at ScienceDirect

Applied Energy

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

Harnessing smart meter data for a Multitiered Energy Management Performance Indicators (MEMPI) framework: A facility manager informed approach

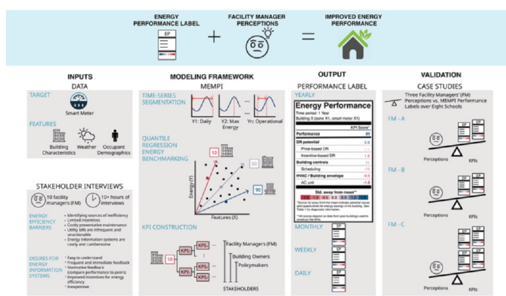
Jonathan Roth, Howard Alexander Brown IV, Rishee K. Jain*

Urban Informatics Lab, Department of Civil & Environmental Engineering, Stanford University, Yang & Yamazaki Energy & Environment (Y2E2) Building, 473 Via Ortega, Room 269B, Stanford, CA 94305, USA

HIGHLIGHTS

- We interview 10 facility managers and consultants to understand how they manage energy.
- Smart meter data is used to benchmark energy use at multiple time scales.
- Key Performance Indicators (KPIs) summarize multitiered energy performance.
- KPIs are compared to perceptions of facility managers through a survey.
- KPIs are used as an energy management system to identify building energy issues.

GRAPHICAL ABSTRACT



ABSTRACT

Energy management information systems (EMIS) play a critical role in providing actionable insights into building operations, timely feedback, and—ultimately—large energy savings. Current EMIS technologies often focus on industrial applications or require large upfront investments and trained operators, therefore greatly limiting its penetration into existing buildings. This paper integrates methods from social, building, and data sciences to understand limitations of current EMIS systems and inform the design of a new Multitiered Energy Management Performance Indicators (MEMPI) framework for characterizing the energy performance of buildings. Specifically, we employed a mixed methods research approach in which we first conduct in-depth qualitative interviews of 10 facility managers and energy consultants. We utilize the insights from our interviews to inform the design of the MEMPI framework, which harnesses highly granular data from already installed advanced metering infrastructure (AMI) (i.e., smart meters). The MEMPI framework employs quantile regression to first benchmark the energy performance of buildings to each other and generate key performance indicators (KPIs). We apply the MEMPI framework to real data from 569 public school buildings in California and measure their energy performance across multiple time scales (e.g., daily, monthly, yearly). Finally, we conduct case studies to compare insights from the MEMPI framework to the perceptions of facility managers overseeing 8 schools through a mixed methods qualitative and quantitative post-interview survey. Results from the case study show that facility managers' perceptions of the performance of their schools were largely accurate, yet the poor energy performance from certain pieces of building equipment and operating schedules was overlooked by building managers. Overall, the MEMPI framework aims to bridge the gap between data-driven energy management models and qualitative domain knowledge held by facility managers to provide more comprehensive insights into the energy performance of buildings.

* Corresponding author.
E-mail addresses: jmroth@stanford.edu (J. Roth), habrown@stanford.edu (H.A. Brown IV), rishee.jain@stanford.edu (R.K. Jain).

https://doi.org/10.1016/j.apenergy.2020.115435
Received 10 March 2020; Received in revised form 3 June 2020; Accepted 24 June 2020
Available online 29 July 2020

1. Introduction

Growing concerns over energy use—due to rising environmental, economic, and social pressures—have pushed countries and businesses alike to re-examine how they consume energy. Residential and commercial buildings alone account for roughly 40% of U.S. energy demand and GHG emissions, therefore improving efficiency in these two sectors can have far-reaching effects [1]. In commercial buildings, where electricity costs alone are roughly $190 billion USD a year, about 30% of this energy is estimated to be wasted, according to a recent report from the U.S. Environmental Protection Agency [2]. The potential to reduce energy consumption in new and existing building is enormous, with many opportunities for low-cost solutions, but much of these savings have yet to be realized [3].

Significant efforts across the world are underway to enhance the energy efficiency of the building sector. Specifically, in the United States, utility companies are investing $7.5 billion USD annually in energy efficiency programs, policymakers are improving building codes, and engineers and architects are creating new building designs to reduce energy demand [4]. But such efforts have limited reach because of the relatively few number of buildings that are constructed each year; most buildings in existence were erected decades ago and thus face substantial energy lock-in effects [5,6]. Fortunately, over the past decade, advancements in sensor technology have reduced costs and have enabled the collection of high-fidelity energy usage data via smart meters, which are estimated to reach a total of 1.2 billion installed worldwide by 2024 [7]. Despite the rapid growth in installations—up from 25 million installed in 2010—smart meter data has mainly been used for utility purposes. Meanwhile opportunities for monitoring and identifying sources of energy waste—through the development of a smart meter-based *Energy Management and Information Systems (EMIS)*—have lagged behind.

In general, EMIS comprise a broad range of tools and services to manage commercial building energy use and include technologies such as *Energy Information Systems (EIS), Fault Detection & Diagnosis (FDD), Energy Benchmarking, and Utility Tracking Tools*. The use of EMIS has been shown to be highly effective with energy savings as high as 20% [8]. Moreover, a study that examined over two-dozen organizations found that participants achieved year-over-year median site and portfolio savings of 17% and 8%, respectively; importantly, the participants in these studies reported that these savings would not have been possible without the EMIS [9]. However, the benefits of EMIS often come with costs associated with additional installations of proprietary equipment necessary to gather the data needed to provide actionable information that leads to these reported savings.

Numerous types of energy management information systems exist and range significantly in their temporal level of analysis as well as their spatial level of insight. On one end, *Fault Detection & Diagnosis (FDD)* analyzes sub-minute energy data to identify when specific building equipment is malfunctioning and is designed for use by facility managers (FMs). On the other end, *Energy Benchmarking* analyzes annual energy data to rank the performance of buildings, across a neighborhood or city, and is designed for use by portfolio building owners or policymakers. While such systems are valuable, there currently exist a major gap in the EMIS domain for a system that provides multi-level insights that empower facility managers (FM), the primary on-the-ground decision-makers in the energy management process. This notion is corroborated by previous work that calls for more suitable building assessment tools that focus on the operational performance of buildings and strengthens the ability of facility managers to save energy [10,11].

In this paper, we employ methods from social, building, and data sciences to understand limitations of current EMIS systems and inform the design of a new Multitiered Energy Management Performance Indicators (MEMPI) framework for characterizing the energy performance of buildings. The MEMPI framework aims to: a) approximate energy performance (i.e., benchmark) across a portfolio of buildings (policymaker and building owner focused), b) provide insights on intra-building dynamics (i.e., occupants, systems, operating states), and c) distill building and energy data into concise key performance indicators (KPIs) (facility manager focused). The MEMPI framework is designed to be flexible to any type of building so long as smart meter data and several building characteristic features can be collected, though including more variables ensures better performance approximations. Using an exploratory sequential mixed methods approach, we first collect qualitative information through semi-structured interviews of 10 facility managers and energy consultants actively working in the field [12]. Based on the information gathered from these interviews, we then design the MEMPI framework and apply it to assess energy efficiency opportunities in 569 school buildings in California, USA. We validate the MEMPI framework—which is not typically done for benchmarking models given the lack of ground truth data—by employing an explanatory sequential mixed methods approach through the use of a post-interview survey. This methodology allows us to compare MEMPI's provided insights with on-the-ground performance observations and perceptions of three facility managers operating 8 school buildings.

2. Literature review

Energy management is critical in high energy intensity industries as it can reduce costs and has substantial business implications [13]. Despite industry-wide efforts to meet high building energy efficiency standards during the design and construction phases, such as LEED and BREEAM, this emphasis is often put aside during the operational phase of a building. The practice of energy monitoring alone is seen as instrumental for raising energy efficiency awareness among tenants, and can translate to behavioral changes, energy saving competitions, and increased accountability [13,14]. Energy management has been shown to be an effective practice in multiple sectors and can be used to continuously monitor performance of buildings, improve efficiency, and reduce operating costs [15]. Despite its demonstrated benefits, the practice remains largely under-utilized given that most research on *Energy Management and Information Systems (EMIS)* focus on either the industrial or residential sector [16-18]. Fig. 1 summarizes the distinct areas of EMIS and the required data frequency needed to achieve various levels of insights—each of the systems shown requires different input data and has different goals. As such, each system has its own benefits and limitations. However, previous work, through interviews grounded in social science methods, has identified the need for a simple, unbiased, and understandable energy monitoring tool aimed at facility managers that can provide comparisons to similar buildings [19,20].

Typically only used by policymakers and building portfolio managers, energy *benchmarking* is the practice of comparing and ranking the energy use of similar buildings with the purpose of identifying inefficient buildings and top performers. Energy benchmarking is growing in popularity as cities are mandating that their largest buildings be benchmarked with the goal of transforming the real estate market to encourage energy savings. Though benchmarking has been shown to achieve savings of about 7% over four years [21], simply measuring the relative energy performance of buildings provides no insights into the drivers of energy use or sources of energy waste. In Australia, one study observed that facility managers of office buildings found the local energy benchmarking system NABERS (National Australian Built Environment Rating System) to be a useful tool to drive behavioral change and help property owners, managers, and tenants improve their sustainability performance and associated financial and reputational benefits [20]. However, interviewees also complained about the limitations of such systems to only include buildings over a certain size or buildings of a certain type; poor user experience for energy management systems is commonly found through interview-

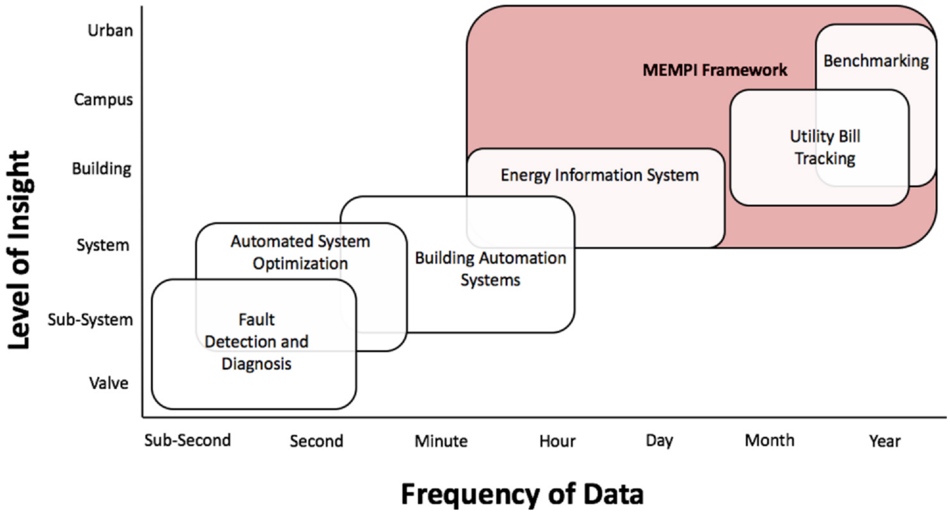


Fig. 1. Visualization of the various Energy Management and Information Systems (EMIS) employed in practice by frequency of data required and the level of insight produced. Our proposed MEMPI framework combines benefits of several current systems by using existing smart meter infrastructure, aligning incentives of the various stakeholders that use these systems, and filling in gaps that exist between Building Automation Systems and Benchmarking.

based social science research [22,23]. *Utility bill tracking* uses more granular energy data than benchmarking (monthly instead of annual energy usage) but provides no baseline or comparative insights to inform facility managers into potential energy issues—it simply provides a method to check current energy usage to historical usage of the same building [24]. Of all the types of EMIS tools, utility bill tracking is the simplest, most widely used, and cheapest, but also the least informative. Annual and monthly feedback—associated with benchmarking and utility bill tracking—can help track long-term sustainability goals, but it fails to inform FMs about daily operations, which can help them identify actionable savings opportunities.

The principle design intent of *energy information systems (EIS)* is whole-building or portfolio-level energy tracking, and automated interval data analysis to identify efficiency opportunities [25]. EIS provides higher levels of insight than benchmarking and utility bill tracking but typically requires high volumes of highly granular data from buildings across a given portfolio. An important facet of an EIS is the application of appropriate KPIs for energy efficiency [26]. A multitude of KPIs have been proposed in the literature but many of these focus solely on manufacturing or industrial applications [23]. Other studies have examined the use of KPIs for buildings at the neighborhood-scale, but these KPIs do not address daily operational energy issues happening at the building-scale [27]. Some are built based on design and construction specifications and neglect using information-rich smart meter data [28]. Studies that have focused on interviewing users have found that poorly constructed indicators inhibits effective performance evaluation and decision-support, thereby casting such systems as inadequate for effective energy management [13,26]. Further, several studies have also reported feedback issues associated with EISs [14], citing that either little feedback is being reported or that delayed feedback results in missed opportunities [13]. Widespread lack of monitoring and feedback mechanisms result in many energy issues going unidentified, wasteful user behavior being unrecognized or changed, and missed educational opportunities for facility managers, engineers, and designers [3,29,30]. Given the complexity of buildings and the multiple sources of energy demand, more immediate feedback is critical for facility manager learning. These studies that have interviewed EIS users have highlighted limitations of current systems but have not proposed alternative potential solutions.

Providing even more insight and control than EIS, *building automation systems (BAS)* are typically used by new and large buildings that have high energy costs. These systems require new equipment to be installed, such as submetering, sensors, controls, and other subsystem components. With more equipment, sensors, and data, BAS help control indoor temperature, light, and humidity through manual setpoints based on building schedules, alarming facility managers when operations are out of a predefined range [25]. Facility managers must set ranges of operation and setpoints for these systems to function, and they must manually change these parameters when schedules or weather changes. In contrast, *automated system optimization* requires a more hands-off approach from facility managers since it automatically modifies control parameters to optimize efficiency, energy use, and costs. Both solutions are expensive and require substantial equipment installation to operate properly [9]. Further, the automated system optimization system requires a deeper level of data analytics and control systems in place to optimize the operations of the systems present in the building.

The most data-intensive building energy management tool is Fault Detection & Diagnosis (FDD), which analyzes high frequency data—in order to parse out subtle operational irregularities—to identify when specific building equipment is malfunctioning. By collecting data at the second or sub-second level, FDD focuses on subsystem equipment—like valves, dampers, and motors of HVAC systems or air-handing units—due to their unit-specific and highly granular data streams; sensors for building automations systems typically collect data such as electric power usage, humidity, temperature, flow rate, pressure, and CO_2 concentration levels [31]. Machine learning algorithms are then used to tease out potential faults in equipment from variations present in normal operating conditions [32]. FDD is an area of active research where new algorithms are being examined, from support vector machines to neural networks, to identify which algorithms operate best based upon the type of data they are receiving, the equipment being monitored, and other confounding effects [33,34]. However, there has been a lack of research focusing on FDD for whole-building energy consumption using smart meter infrastructure due to the difficulty of identifying faults with lower granularity data and a lack of ground truth information at the whole building level [35].

The precipitous growth of smart meter infrastructure across the globe—with 70 million installed in the U.S. and close to 96 million in China—has spurred researchers to determine ways to create new and scalable analytical tools that help answer questions posed by utilities and policymakers. Specifically, studies have primarily focused on using smart meter data to improve forecasting [36] and customer segmentation through clustering [37,38]. Although helpful for utilities and policymakers, these tools largely overlook use-cases that may be helpful for facility managers to efficiently operate their building. Research on fault detection and diagnosis typically rely on meters with much more granular data streams than can be offered by smart meters; these studies also often look at meters at the sub-system level, requiring extra equipment [33,35]. One useful area of research using smart meter data directed at facility managers is measurement and verification (M&V), which focuses on verifying the amount of energy savings after a retrofit

[39,40]. M&V using smart meter infrastructure is particularly useful because it eliminates the need for energy consultants to visit the site multiple times to install sensors, collect data, and remove the sensors both before and after a retrofit project.

Despite the body of work that focuses on smart meter analytics, there has been a dearth of studies that have focused on using this smart meter infrastructure in commercial settings to monitor energy use in real-time and provide targeted feedback useful for facility managers. Several studies have used smart meters for energy benchmarking, but they have focused solely on measuring total building energy performance and have not looked at benchmarking other key metrics that can be derived from smart meter data streams [18,41,42]. As a result, there exists a significant opportunity to leverage high fidelity data from smart meter infrastructure to bridge the spatial and temporal insights gap between benchmarking systems and Building Automation Systems (BAS). This need is further underscored by interest from industry in which companies like Lucid, Gridium, and Aquicore are tapping into smart meter data to provide dashboards and actionable insights to transform building operations into performance improvements that save money, enhance sustainability, and streamline maintenance.

3. Interviews

Given the complexity of building systems and the dynamic nature of energy consumption, we adopted an exploratory sequential mixed methods approach [43] to understand the needs and limitations of current energy management systems. The user insights gained from these qualitative semi-structured interviews of facility managers and energy consultants was used to inform the design of our MEMPI (Multitiered Energy Management Performance Indicators) framework. This section explains the first qualitative portion of the mixed methods approach before the user informed MEMPI framework is introduced in the following section. The semi-structured interviews aimed to better understand the roles of facility managers, the hurdles they face, and how they perceive the energy performance of their buildings [12]. We conducted a total of ten semi-structured interviews with FMs for a subset of buildings in which we have access to high-fidelity smart meter data. Nine of the interviewees were facility managers at public schools—who manage over a combined 150 schools—and one was an energy consultant who, over the course of his career, has worked with over 40 school districts throughout the state. The interviews were either conducted in person or over the phone and lasted between 45 and 75 min. The ten participants were recruited through a snowball sampling process where each consented to being recorded for the interview. The interview protocol document is attached as Appendix C and includes the list of questions we used to guide every interview. We coded the transcripts for all interviews using NVivo [44].

3.1. Insights from interviews

In each interview, the participants were asked several questions about their responsibilities to contextualize efforts spent on energy related issues. Consistent with previous literature, we found FMs must balance a variety of responsibilities such as directing personnel, managing funds, responding to work-orders, maintaining facility conditions, among others [20]. Although most FMs expressed interest in improving the energy performance of their facilities, they also admitted that these efforts are often deprioritized since other unexpected issues often arise. Furthermore, several interviewees stated they lack an incentive to prioritize energy efficiency as it does not factor into their job performance assessments. Therefore, while administrative staff might express interest in energy efficiency, they often fail to provide the means or incentives to assist FMs achieve this goal. School districts that are very serious about energy performance sometimes opt to hire a third-party energy consultant instead of handling these issues in-house. Interviewees found these consultants to be helpful, though they did express difficulty in finding a consultant that they could trust—many FMs experienced receiving phone calls from energy consultants claiming opportunities to save energy and money, but the interviewees thought that many of these callers were scammers. Beyond the barrier to initially hiring a consultant, every FM that worked with an energy consultant ended up having a positive experience.

A common barrier to increasing energy efficiency of buildings is the difficulty FMs have in detecting faults in building equipment, which was often attributed to the lack of a preventative maintenance crew, largely due to their often-prohibitive cost. These staff members help identify and fix faulty equipment in buildings, which often go unnoticed, by periodically checking equipment for issues. Without this staff, small problems can grow into costly projects, leading FMs to work in a reactionary manner rather than in the preferred preventative manner. In addition, a flurry of service requests from teachers, administrators, and other staff to resolve minor, surface-level issues instead of major, root cause issues—and a lack of budgetary appetite for capital improvements above most other priorities—results in large lists of deferred maintenance issues. Tight budgets limit many FMs from having a preventative maintenance staff, which forces them to address their backlog by going after age-driven issues (i.e., older equipment is prioritized for replacement). Although this system is the most practical given the financial and manpower constraints, it also fails to prioritize issues based on severity, which remains unknown without a preventative maintenance crew. Further, many FMs stated that they do not look at the utility bills and only hear about cost spikes if they receive a phone call from the accounting office that handles these bills. Many FMs admitted that many energy inefficiencies are likely occurring in their buildings but are not being addressed due to unobservable effects. When looking to improve energy performance of buildings, several FMs stated that they look at the utility bill price per square foot for each building in their portfolio to assess which buildings are high energy consumers as it was the easiest and most practical given resource and information constraints.

Energy waste in buildings can be typically attributed to a few sources. The major source, though, for energy waste expressed by every interviewee is the behavior of occupants in their buildings. Mitigating the effects of this source proves to be difficult due to two main factors. First, detecting the wasteful behavior is non-trivial given the complexity of buildings—from heating ventilation and cooling (HVAC) equipment to separate lighting control systems—and the dynamic nature of occupant flow and magnitude. Second, finding proper solutions to discourage the identified behavior is also challenging due to the differing requests of occupants—from temperature set-points to ventilation rates—and their often-incognizant decisions to alter control settings that can have large energy implications. Discovering that occupants are leaving lights on, opening windows and doors when the HVAC is running, or adjusting the thermostat is the first, and larger issue of the two. Many of the FMs interviewed have managed to come up with a number of solutions, both technical and non-technical, in response to these types of behaviors. For example, several FMs are working to install motion detecting lighting systems, sensors on doors that automatically shut off HVAC units, and thermostats that do not link to the control systems (i.e., they provide occupants with a false sense of control over temperature as the dial does not actually control anything). However, understanding occupant behavior and how it is affecting energy demand is the most pressing issue that FMs feel is often being overlooked.

For those that have energy management systems in place, facility managers feel that current systems are very expensive and can be cumbersome to use—several even questioned if their value outweighed the cost of the system. These insights are corroborated in a study from Curtis et al. [20], who found that FMs from mid-tier commercial office buildings desired cheap, timely, and regular feedback regarding the energy efficiency performance of their buildings. One FM noted trouble detecting whether newly commissioned HVAC units were properly installed until the EMS system was also properly commissioned. Additionally, these FMs expressed a desire for a system that uses comparisons, or some type of ranking system, to compare performance against similar buildings. Such a system could also be used to provide incentives and recognition to FMs who achieved improved energy efficiency. Further, they expressed the desire for a platform with quick, easy-to-understand, feedback and visualizations. One FM received Energy Star scores from an energy consultant and decided to focus energy efficiency retrofits on the building with the worst score. Despite the benefit provided by measuring building performance with Energy Star, this FM also indicated that initially the staff did not know where to focus their efforts and needed assistance in finding sources of inefficiency within the identified building. Typically, energy management systems are the technology of choice to satisfy this need. But those FMs interviewed highlighted that many systems are difficult to use, expensive, and lack desired features.

Overall, most FMs stated that they have very limited access to energy management tools to help them save energy in their buildings, and the several FMs that did, complained about high costs, poor user experience, and difficulty translating information into action. These sentiments point to the need for a system that leverages new high-fidelity data streams to provide deep levels of actionable insights into building operations.

4. Framework and methodology

Beyond the insights gathered from the qualitative portion of the mixed methods approach, this paper proposes the development of new key performance indicators (KPIs) for holistic, multi-level energy management that can be used by various stakeholders with different priorities. The MEMPI (Multitiered Energy Management Performance Indicators) framework is broken down into two key sections. Section 4.1 provides an overview on the mechanics of quantile regression-based benchmarking and its extension to measuring daily energy performance by leveraging smart meter data. In Section 4.2, using insights gathered from the FM interview process, we extend the smart meter-based benchmarking explained in Section 4.1—and discussed in detail in our previous work [41]—by proposing new KPIs to succinctly summarize building energy performance at different time scales. The MEMPI framework leverages high-fidelity data emerging from existing smart meter infrastructure and is designed to be applicable to any type of building in any climate. Further, the framework benchmarks energy performance across a portfolio of buildings, provides insights on intra-building dynamics (i.e., occupants, systems, operating states), and clearly translates these ideas into potential actions for policymakers, building owners, and, most notably, facility managers.

4.1. Quantile regression benchmarking and the DMT model

The interviews, and their insights detailed previously, shaped the design and development of our MEMPI framework. Due to the expressed desire of FMs for an energy management system that uses comparisons and performance rankings, we chose to build our novel framework off the basis of the established concept of energy benchmarking. In this section, we describe the mechanics of the data-driven, multi-metric, and time-varying (DMT) model that uses smart meter data to benchmark buildings at a more granular timescale than yearly, as shown in Fig. 2. The DMT model extracts energy features from smart meter data—and uses them as the dependent (target) variables—utilizing quantile regression (QR) as a basis for benchmarking as it boasts a number of benefits over other models, as outlined below.

Mechanically, quantile regression is a parametric model that takes the same functional form as ordinary least squares (OLS) regression as displayed in equation (1); however, the parameters are fit using a different cost function which is seen in equation (2), where $x_i^{'}$ is the vector of covariates with size j, β_τ is the produced vector of coefficients, and y_i is the response variable for building i. By altering the hyperparameter τ (tau) between 0 and 1, quantile regression is able to model different components of the conditional distribution of the response variable in relation to the covariates; a value of $\tau = 0.50$ corresponds to modeling the median, analogous to OLS regression that models the conditional mean of the response variable.

$$\hat{y} = \beta_0 + \beta_1 x_1 + \cdots + \beta_j x_j \tag{1}$$

$$Q(\beta_\tau) = \sum_{i:y_i \geq x_i^{'}\beta}^{N} \tau\,|y_i - x_i^{'}\beta_\tau| + \sum_{i:y_i < x_i^{'}\beta}^{N} (1-\tau)|y_i - x_i^{'}\beta_\tau| \tag{2}$$

Models are then constructed at each $\tau = \{0.01, 0.02, 0.03, \cdots, 0.99\}$ to create a total of 99 models that represent the response variable values for each quantile of the conditional distribution. The final benchmarking score is then found using equation (3), where $\tau_{closest}$ represents the tau value of the model with the smallest difference between the predicted value and the observed value. The score ranges from 1 to 99, where 1 indicates the lowest performance attainable (i.e., the bottom 1st percentile) and 99 indicates he highest performance attainable (i.e., the top 99th percentile). Because we are modeling building energy performance, our covariates include available data about buildings that normalize their energy use, such as weather, floor area, building type, etc. The major benefits of using quantile regression for benchmarking energy performance are: (a) reduced sensitivity to outlier data points; (b) ability to determine the varying effects of inputs by modeling the entire conditional distribution of the dependent variable; (c) capacity to normalize for numerous explanatory variables; (d) handles heteroskedastic data—a common feature of building energy data; (e) identification of non-linear relationships between explanatory variables and energy use—this is best elucidated through quantile regression influence plots; (f) results that are highly interpretable despite assuming non-constant relationships between explanatory variables and energy consumption. For a more detailed explanation of the benchmarking model mechanics, benefits, and use cases, please refer to our previous work [46].

$$score = (1 - \tau_{closest}) * 100 \tag{3}$$

Traditional energy benchmarking focuses on measuring and ranking the performance of buildings at the annual time scale. Applied here, that would translate to y_i representing the annual energy consumption for building i. In our previous work [41], the DMT model was proposed to extend benchmarking to multiple metrics (besides annual energy use) and at the daily time-scale by leveraging advanced metering infrastructure (AMI) and other open-data sources, such as daily weather profiles and public databases [41]. This enables y_i to represent more than just annual energy consumption, like daily or weekly energy consumption. In this paper, we propose six metrics for benchmarking, as shown in Fig. 2: *total* (*daily*) *energy, operational energy, non-operational energy, peak (max) energy demand, max ramping,* and *energy volatility*. By measuring the operational and non-operational states of a building, the model can compare the energy use performance of the whole building during times of low and high occupancy levels. Producing daily scores also allows for examination into seasonal and daily trends, allowing users to visually identify abnormal energy consumption patterns. Using the DMT framework has several key benefits over other benchmarking models: (a) measures performance at smaller time scales enabling faster delivery of feedback which can lead to increased energy savings [47]; (b) ranks energy performance of different building operating states, thereby providing insights into whole-building operations at varying scales; (c) leverages the growing network of smart meter infrastructure, eliminating cumbersome energy data gathering and providing high quality, granular, and timely energy data streams for deeper analytics; (d) utilizes other new data sources—such as granular weather data streams and open-data portals—that allow for improved weather adjustments and reduced data collection efforts; (e)

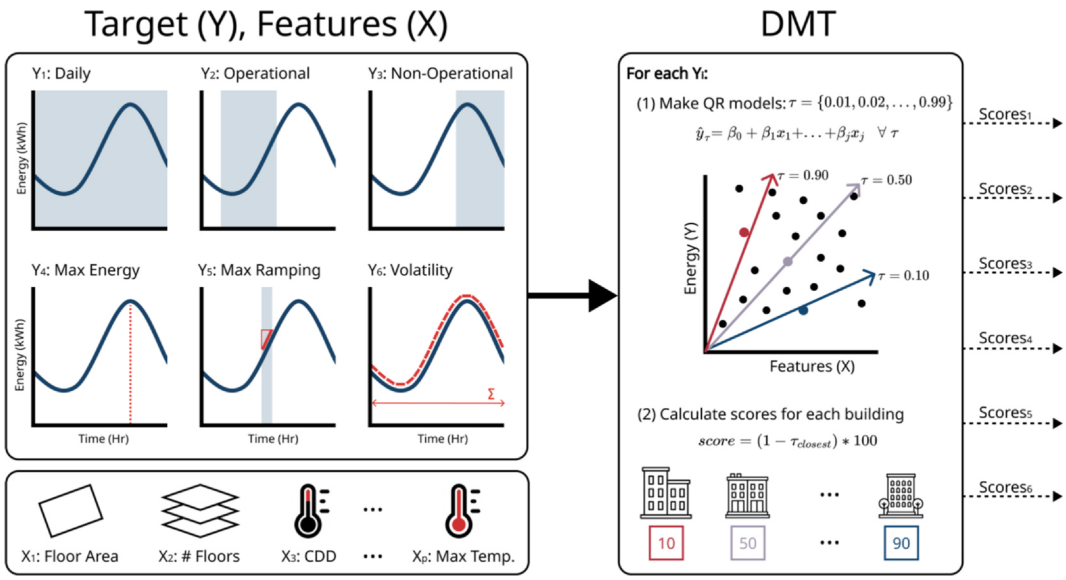


Fig. 2. Process diagram for the DMT model, which uses quantile regression benchmarking on smart meter data. Building and thermometer icons from [45].

maintains interpretability thereby enabling stakeholders with various backgrounds to use the system and gather insights on daily, weekly, and seasonal performance. For a more detailed explanation of the DMT framework and its merits, please refer to [41].

4.2. Construction of key performance Indicators (KPIs)

In this section, we propose KPIs for energy management based on the insights derived from the facility manager interviews as discussed in Section 3. The proposed KPIs are constructed from, and using, the scores produced from the DMT model, as shown in Fig. 3. By coalescing domain knowledge and insights gathered from the interviews, each KPI was built to identify building inefficiencies and/or measure levels of energy performance. Specifically, we constructed each KPI to help answer a precise question that can help point to a potential inefficiency or source of waste in the building. Our proposed KPIs are broken down to address four different scales of energy performance—daily, weekly, monthly, and yearly. Each timescale provides various levels of information that can be used by different stakeholders. Daily metrics can help FMs identify issues quickly while weekly and monthly metrics may highlight longer term or systematic issues that would otherwise be disregarded as noise when exclusively examining daily KPIs. Yearly KPIs provide a broader overview of building energy performance and can be used by building owners and policymakers to reward top performers, target inefficient buildings, and better allocate resources.

Each KPI is constructed using three different methodologies: (a) directly benchmarked energy data using quantile regression-based benchmarking and the DMT framework as outlined in Section 4.1 which normalize building energy performance based on available data such as weather (temperature), floor area, building type, etc. This will be referred to as DMT-KPIs; (b) combinations (i.e., summations) of DMT-KPIs, typically over time, to show more persistent and systematic issues of energy waste. Because these KPIs are built using DMT-KPIs, they also normalize performance by the same data. These combinations of benchmarks will be referred to as COB-KPIs; (c) models constructed to highlight relationships between DMT- or COB-KPIs with other raw data. These benchmarking relationships will be referred to as BR-KPIs. Table 1 summarizes all of the proposed KPIs, how they are constructed, and what questions they aim to help answer.

The proposed KPIs were selected to provide desired feedback—or address specific concerns—for FMs, building owners, and policymakers. Fig. 4 is an example output of the energy performance labels produced by the MEMPI framework using the proposed KPIs. The motivation for the construction of each KPI is as follows (see Table 1 for more details):

1. ***Trends in energy performance:*** *Most generally, building energy performance can be examined at different time-scales, from daily to yearly (DMT^{Daily}_{Energy}, COB^{YMW}_{Energy}), and between weekdays and weekends ($DMT^{Daily}_{Weekend}$) to identify changes in performance over time. Facility managers are often also interested in examining the difference in operational performance to disentangle energy performance at different times of day (DMT^{MWD}_{Op}, DMT^{MWD}_{NonOp}, $COB^{Daily}_{Op\_NonOp}$, $COB^{YMW}_{Op\_NonOp}$).*
2. ***Comparative feedback:*** *The DMT-KPIs (DMT^{Daily}_{Energy}, DMT^{MWD}_{Op}, DMT^{MWD}_{NonOp}, DMT^{MWD}_{Peak}, $DMT^{Daily}_{Ramping}$, $DMT^{Daily}_{Volatility}$) provide normative feedback that directly ranks the energy performance of the specified dependent variable among multiple buildings. Normative KPIs provide FMs with reference points that helps them understand the potential energy savings achievable since it is comparing their building to peer buildings. Normative feedback was identified as a desire of FMs through the interviews and has been shown to be an effective form of feedback to encourage energy savings by contextualizing current and historical consumption in relation to a user's peer* [48].
3. ***Occupant effects:*** *Several metrics were designed to help facility managers pinpoint potential occupancy behavior issues, like overriding thermostat settings, using high energy consuming equipment, and abnormal use of equipment ($DMT^{Daily}_{Volatility}$, DMT^{Daily}_{Peak}, $COB^{YMW}_{Op\_vol}$). These behaviors can result in volatile energy use at the building-level as equipment is ramped up or down without notice. $DMT^{Daily}_{Ramping}$ measures the increase in energy consumption between two sequential periods of time, indicating if there is any large jump in energy use. Further, this aids policymakers that must grapple with ramping constraints of power generation that has recently been further exacerbated with growing concerns around intermittent renewable energy sources* [49-51].
4. ***Control systems:*** *Apart from operational* (i.e., *daytime) use of energy, non-operational energy consumption can help FMs identify control issues, like when systems are unintentionally left running at night ($COB^{YMW}_{Op\_vol}$, $COB^{YMW}_{NonOp\_vol}$, $DMT^{Daily}_{Weekend}$). By measuring the energy performance of both operating and non-operating states of a building, FMs can watch for spikes in energy use during times when they know the building is unoccupied and systems should be off.*
5. ***Demand response (DR) potential:*** *Volatility, often also referred to as variability, in energy use ($DMT^{Daily}_{Volatility}$, $COB^{YMW}_{E\_vol}$, $COB^{YMW}_{Op\_vol}$, $COB^{YMW}_{NonOp\_vol}$) is a key index in evaluating the potential of demand response, where consumers with lower entropy have relatively similar*

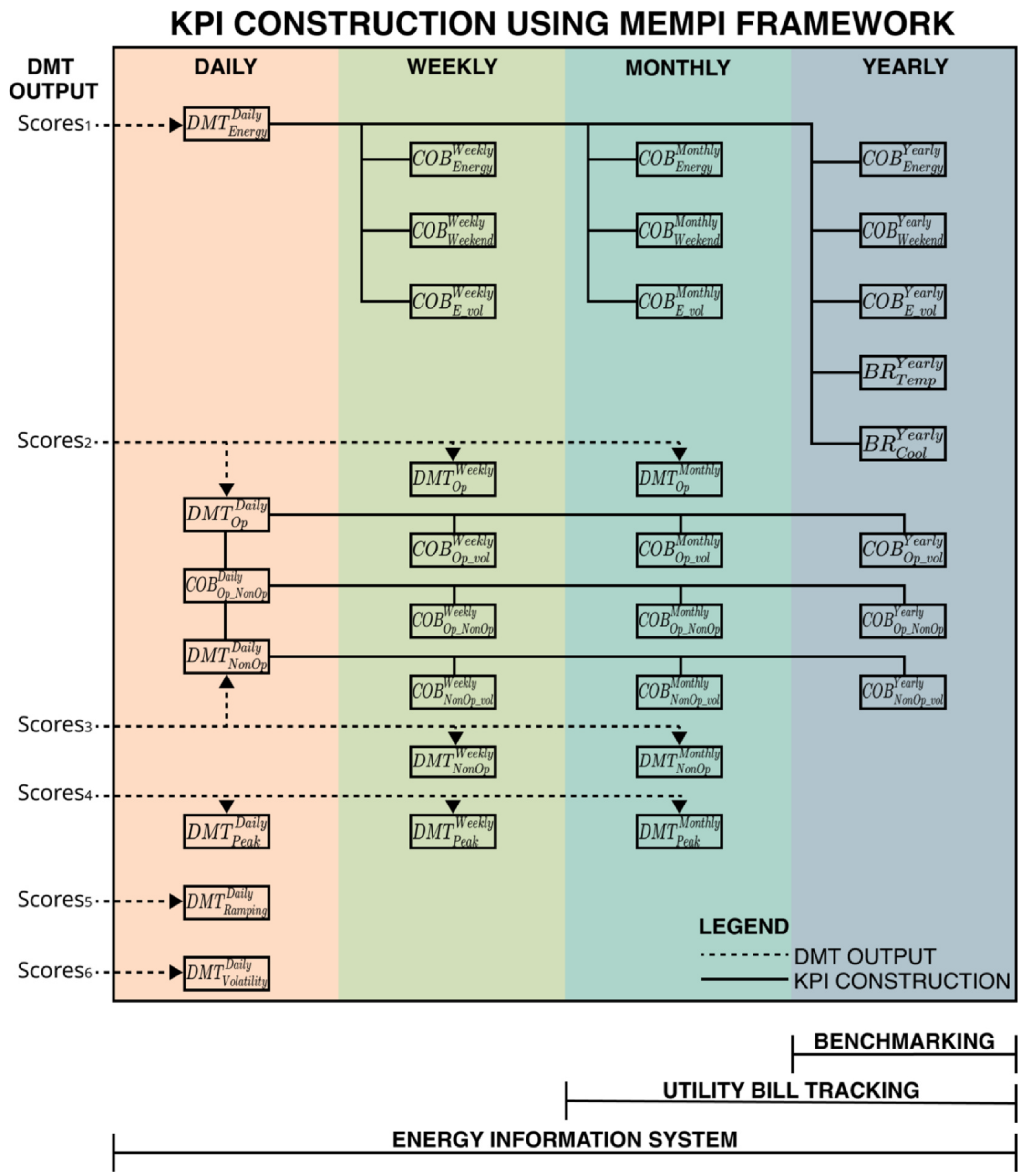


Fig. 3. Breakdown of the KPIs—and their dependencies—based on their timescale and which energy management practices they replace. Scores from the DMT model are directly used as KPIs and indirectly used to construct other KPIs. See Table 1 for details, information, and KPI equations.

patterns of energy consumption—and viewed as suitable for incentive-based demand response due to their predictability to follow control commands—while consumers with high entropy are suitable for price-based demand response due to their flexibility to adjust their load based on change in price [38]. Peak load monitoring (DMT^{MWD}_{Peak}) is also important for FMs that are looking to reduce demand charges of their building; the associated KPI can help them pinpoint which days are problematic and isolate scheduling issues that can be causing the spike in energy use.

6. ***HVAC / Building envelope:*** *Measuring the relationship between daily energy performance scores and temperature (BR^{Yearly}_{Temp}, BR^{Yearly}_{Cool}) can help FMs further isolate issues associated with HVAC units or building envelope since hot days correspond with increased use of air conditioning which is known to be a high energy consumer."*

The proposed MEMPI framework fills a gap in the scattered, and often ill-defined, space that comprises building management systems (BMS), energy management systems (EMS), and energy benchmarking, as discussed in Section 2 and shown in Fig. 1. Measuring how these metrics change over time, or combinations of them, can provide additional insights that may be overlooked if only looking at a single KPI independently. The main benefit of the proposed framework is that it encompasses three distinct and central types of energy management practices—energy information systems, utility bill tracking, and energy benchmarking. Each of these practices is targeted at different stakeholders. Energy information systems are primarily used by facility managers to track building performance and identify potential issues; utility bill tracking is primarily used by building owners who want to manage energy costs and understand energy use patterns between buildings in their portfolio over extended periods of time; benchmarking is primarily used by policymakers who want to track the energy performance of city, or state, building stocks and encourage competition between buildings to drive energy savings. As shown in Fig. 4, the KPIs from the MEMPI framework provide insights at multiple timescales and for multiple stakeholders. By satisfying the main needs of each stakeholder, and aligning their incentives under one system, the

Table 1
Summary of the proposed KPIs, the questions they aim to answer, and how they are constructed.

KPI Name	Equation	Frequency	Inputs	Output Type	Summary	Questions Answered by KPI
Full name DMT_{Energy}^{Daily}	$y_{Energy}^{Daily} = \beta_0 + \beta_1 x_1 + \cdots + \beta_i x_i$	Daily	Building characteristics, weather, demographics [x_i]	DMT Benchmark	Daily energy performance—This metric provides quicker feedback than utility bills to help FMs identify issues occurring in their buildings.	Was the building utilized more than usual today?
Full name DMT_{Op}^{Daily}	$y_{Op\_Energy}^{Daily} = \beta_0 + \beta_1 x_1 + \cdots + \beta_i x_i$	Daily	Building characteristics, weather, demographics[x_i]	DMT Benchmark	Daily operational performance—this metric helps FMs identify performance of their building when occupancy is high and systems are running.	Were there more occupants in the building than usual? If not, was occupant behavior an issue today?
Full name DMT_{NonOp}^{Daily}	$y_{NonOp\_Energy}^{Daily} = \beta_0 + \beta_1 x_1 + \cdots + \beta_i x_i$	Daily	Building characteristics, weather, demographics [x_i]	DMT Benchmark	Daily non-operational performance—this metric helps FMs identify issues during periods of low occupancy when building equipment should not be running.	Were any systems left on during the night? Was there an event that ran late into the night?
Full name DMT_{Peak}^{Daily}	$y_{Max\_Energy}^{Daily} = \beta_0 + \beta_1 x_1 + \cdots + \beta_i x_i$	Daily	Building characteristics, weather, demographics [x_i]	DMT benchmark	Daily peak performance—this metric helps FMs identify days that contribute to their peak demand charges, enabling them to identify scheduling or occupant issues.	Did today affect my peak power usage for my utility bill? Are all my systems turning on at the same time?
Full name $DMT_{Ramping}^{Daily}$	$y_{Max\_Ramp}^{Daily} = \beta_0 + \beta_1 x_1 + \cdots + \beta_i x_i$	Daily	Building characteristics, weather, demographics [x_i]	DMT benchmark	Daily ramping performance (for time period with max ramping)—this metric helps FMs identify building equipment with known schedules that are high energy consumers.	If I know the schedule of certain building systems, do they seem to be large users of energy?
Full name $DMT_{Volatility}^{Daily}$	$\sum_{t=2}^{D} \lvert y_t - y_{t-1} \rvert = \beta_0 + \beta_1 x_1 + \cdots + \beta_i x_i$	Daily	Building characteristics, weather, demographics [x_i]	DMT benchmark	Daily volatility performance—this metric helps FMs understand if demand response programs may help them save money, and which type of demand response is more suitable.	Does occupant behavior seem to be driving energy consumption in this building?
Full name $COB_{Op\_NonOp}^{Daily}$	$DMT_{Op}^{Daily} - DMT_{NonOp}^{Daily}$	Daily	DMT benchmarks	Combination of Benchmarks	Daily difference in operational vs. non-operational performance—this metric reveals occupant behavior (OB) or faulty equipment (FE) issues.	High Value—Were any systems left on at night? Low Value (very negative)—Were building systems being utilized heavily during the day? Were there higher occupancy-driven effects than normal?
Full name BR_{Temp}^{Yearly}	$DMT_{Energy}^{Daily} = \beta_0 + \beta_1 T_1$	Yearly	Temperature [T_1], DMT benchmarks	Benchmarking Relationship	This metric model the relationship between temperature and DMT scores and helps reveal HVAC or building envelope issues.	Low Value—Is the AC unit inefficient or the building envelope leaky?
Full name BR_{Cool}^{Yearly}	$DMT_{Energy}^{Daily} = \beta_0 + \beta_1 C_1$	Yearly	Cooling Degree Day (CDD) [C_1], DMT benchmarks	Benchmarking Relationship	This metric model the relationship between temperature and DMT scores and helps reveal HVAC or building envelope issues.	Low Value—Is the AC unit inefficient or the building envelope leaky?
Full name $COB_{Op\_NonOp}^{YMW}$	$\frac{1}{N}\sum_{t=1}^{N} COB_{Op\_NonOp}^{Daily}$	Yearly, monthly, or weekly. Indicated by YMW in KPI superscript	*COB*-KPIs	Combination of Benchmarks	This metric measures the long-term trend between operational and non-operational performance and reveals systemic control systems issues related to nighttime scheduling (if high value) or inefficient building systems (if low value).	High Value—Are systems being left on at night consistently (control issues)? Low Value (very negative)—Are the systems that are known to run during the day very inefficient?
Full name COB_{Energy}^{YMW}	$\frac{1}{N}\sum_{t=1}^{N} DMT_{Energy}^{Daily}$	Yearly, monthly, or weekly. Indicated by YMW in KPI superscript	DMT-KPIs	Combination of benchmarks	This metric measures yearly energy performance. I.e., traditional energy benchmarking	How (in)efficient is this building?
Full name $COB_{Weekend}^{YMW}$	$\frac{1}{N}\sum_{t=1}^{N} DMT_{Energy}^{Daily} - \frac{1}{M}\sum_{t=1}^{M} DMT_{Energy}^{Daily}$	Yearly, monthly, or weekly. Indicated by YMW in KPI superscript	DMT benchmarks. N = weekdays M = weekend days	Combination of benchmarks	This metric measures the weekly difference between weekday and weekend performance and reveals control issues related to weekend scheduling.	Are systems being left on during the weekend? High Value—Weekends worse Low Value—Weekends better
Full name $COB_{E\_vol}^{YMW}$	$-\sum_{t=1}^{N} (DMT_{Energy}^{Daily} * \log(DMT_{Energy}^{Daily}))$	Yearly, monthly, or weekly. Indicated by YMW in KPI superscript	DMT-KPIs	Combination of benchmarks	Entropy of daily energy performance for one year.	High Entropy—Customer is suitable for price-based demand response? Low Entropy—Customer suitable for incentive-based demand response?
Full name $COB_{Op\_vol}^{YMW}$	$-\sum_{t=1}^{N} (DMT_{Op}^{Daily} * \log(DMT_{Op}^{Daily}))$	Yearly, monthly, or weekly. Indicated by YMW in KPI superscript	DMT-KPIs	Combination of benchmarks	Entropy of daily operational performance for one year.	High Entropy—How much do occupants drive energy use in this building compared to others?

(*continued on next page*)

Table 1 (*continued*)

KPI Name	Equation	Frequency	Inputs	Output Type	Summary	Questions Answered by KPI
Full name $COB^{YMW}_{NonOp\_vol}$	$-\sum_{t=1}^{N}(DMT^{Daily}_{NonOp} * \log(DMT^{Daily}_{NonOp}))$	Yearly, monthly, or weekly. Indicated by YMW in KPI superscript	DMT-KPIs	Combination of benchmarks	Entropy of daily non-operational performance for one year.	High Entropy—Indicates that control issues are not systematic but that building systems are frequently being used during non-operational hours

proposed MEMPI framework can provide a low-cost solution to energy management that has the potential to result in large energy savings.

Using KPIs can streamline decision-making, enable better preventative maintenance practices, and provide timely and targeted feedback, all of which are currently significant issues highlighted by the FMs we interviewed. Rather than focus on alternative display layouts and new charts to show FMs, which has been the focus of several studies at the residential level [52], we aim to explore the benefits of KPIs derived from smart meter data streams from commercial and SMB (small and medium business) buildings. Though KPI construction is a high area of interest in the academic community for energy management in the industrial sector, there have been few studies that have extended it to the commercial building sector [15]. KPI development for energy management has gained popularity because it synthesizes copious data into one number that is quickly absorbed, easily understood, and highlights a critical area of performance that is indicative of potential issues. These KPIs provide more insightful information than utility bills and utilize in-place and low-cost smart meter infrastructure.

5. Case studies

Validating building energy performance (i.e., benchmarking) is a critical challenge in the field and a major reason so few studies have attempted it rigorously [53]. Unlike vehicles—which have a straightforward definition for efficiency as defined by the distance traveled per unit of energy—buildings vary dramatically in size, type, climate conditions, and offer a wide array of services. Buildings rarely exist to serve one singular purpose, thereby making the exercise in defining efficiency subjective. There is no objective definition when people value natural light and indoor air quality—among many other services—quite differently. Therefore, no "ground truth" data exists, but there are several reasonable alternatives that can be explored. Conducting energy audits for all buildings in a dataset would provide rich information to compare to benchmarking results; however, collecting this data for a large dataset would be cost- and time-prohibitive. As a result, we adopted a case study approach to study the efficacy of the MEMPI framework and contextualize the data-driven results among perception and insights of real facility managers.

Although not a statistically driven method, case studies have long been an effective method to study, assess, and validate energy efficiency technologies and analytical methods for buildings [16,54-57]. We applied the MEMPI framework to real smart meter data from 569 school buildings in California, USA and then employed an explanatory sequential mixed methods approach [43]— through the use of a post-interview survey—to compare the insights produced from the MEMPI framework to those seen by the facility managers who operate the buildings. This mixed qualitative and quantitative approach allows us to compare MEMPI's provided insights with on-the-ground performance observations and perceptions of three facility managers operating eight case-study buildings. The goal of the case studies is to compare and contrast the insights gathered from the proposed MEMPI framework, and associated KPIs, to those obtained from the FMs through the post-interview survey.

5.1. Data and modeling

We applied our MEMPI framework to 15-min interval electricity usage data from 569 public schools across the state of California, USA (see Fig. 5 for a map of the geographic distribution). Electricity consumption data for these schools was collected using data released from Proposition 39 program administered by the California Energy Commission (CEC) for the 2014–2015 academic year (July 2014 to July 2015). Data on building characteristics was collected from the California Department of Education (CDE) and Federal Census Bureau, all publicly available online. Daily weather data was gathered from the National Oceanic and Atmospheric Administration (NOAA) and merged

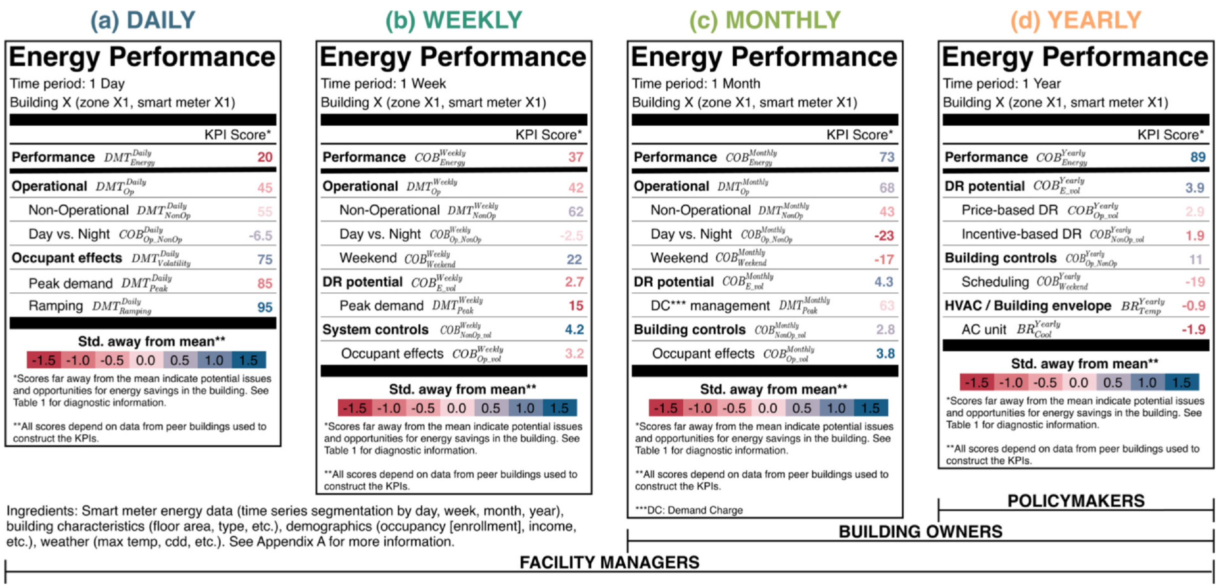


Fig. 4. Example output of the performance labels created by the KPIs from the MEMPI framework. Note, the scores in the above labels are merely exemplary to showcase a realistic output label.

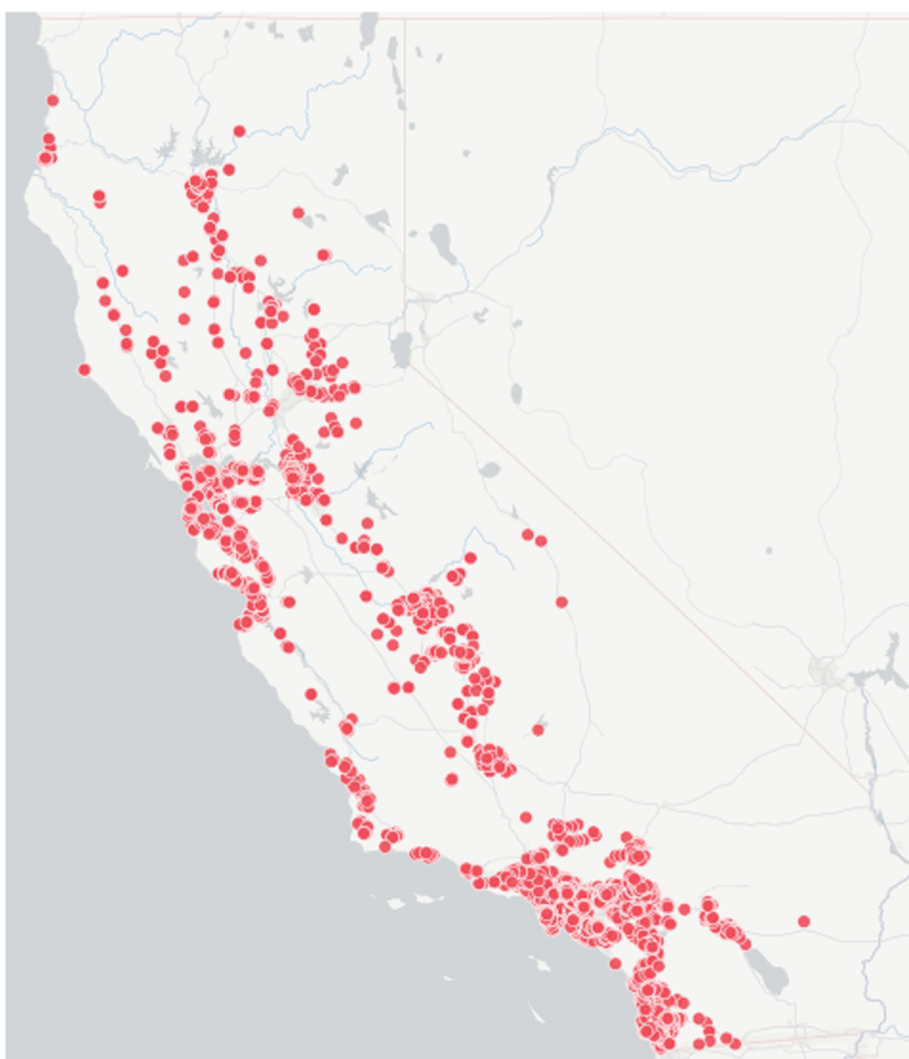


Fig. 5. Map showing the location of every school included in the study.

at the zip code level. We focused our validation and testing of MEMPI on California public schools due to the fact that all schools are engaged in relatively similar activities, have similar operating schedules, and are subject to similar funding trends throughout the state. The MEMPI framework can, in fact, be used for buildings of any type or location, but using one building type (e.g., schools) in one US state allows for a more rigorous validation process. Information and descriptions about each of the 34 independent variables used in the study can be found in Appendix A, and include data on building characteristics, occupant demographics, and daily weather patterns. Predictive mean matching was used to impute any missing values [58]. Due to the high number of variables in the dataset, a forward stepwise variable selection process was used to select the top 10 most influential variables for each month. More details about the DMT modeling approach can be found in our previous work [41].

Daily energy performance (i.e., results depicted in Fig. 4a) was then found by modeling each month of the year independently to account for seasonal changes in energy drivers. Energy data from each of the 569 schools for every day of the month was used as the dependent variables (as depicted in Fig. 2 and further outlined in Table 1) while building characteristics, school demographics, and weather data were used as the independent variables. Quantile regression models were built for each of the 6 DMT-KPIs for each month of the 2014–2015 academic year, producing 6 DMT-KPI's for each day of the year in each school. For these case studies, we defined the operational state to be between 7am–3pm and the non-operational state to be between 10pm and 4am, given that school schedules are consistent and widely known; other definitions for operational and non-operational can be used in the proposed MEMPI framework depending on if the schedule is known or can be extracted from the smart meter data. Using the 6 DMT-KPIs, the other KPIs were then constructed using the equations outlined in Table 1. In total, each school has 6 DMT-KPIs for each day of the year plus an additional 8 COB- and BR-KPIs that summarize yearly performance.

Although performance labels are produced weekly and monthly—and host their own set of benefits measuring performance at this timescale—we decided to focus on yearly performance labels. The historical nature of our data risks facility managers improperly recalling the energy use dynamics of their buildings in the past at shorter timescales, like weeks or months. Fig. 6 shows the distributions for each of the eight yearly KPIs for every school in the dataset. The $COB_{Op\_vol}^{Yearly}$ has a median value of 3.84 while $COB_{NonOp\_vol}^{YMW}$ has a median of 3.28, indicating that the operational state has higher variability than the non-operational, and validating the FM perception that occupants have a large effect on building energy consumption. Fig. 6 also contextualizes the results later discussed in Section 5.3.

5.2. Post-Interview survey

In order to provide a basis for comparing the results of the MEMPI framework to perceived on-the-ground conditions at the schools, an extensive follow-up post-interview survey was administered to the same ten FMs that we interviewed and who are responsible for eight schools in our dataset. In this post-interview survey, we collected information regarding their opinions about the energy performance of the schools that they managed. Details of the post-interview survey questionnaire can be found in Appendix B. Our aim was to collect data on the FMs' judgements about the conditions and (in)efficiencies of systems within particular schools. For example, beyond their views on overall energy performance of specific schools, we sought information on HVAC equipment, lighting systems, control settings, sensitivity to hot days, age of equipment, building envelope, and more. With this information provided by the FMs, we aim to compare their perceived issues to those identified by MEMPI framework and associated KPIs. Given that our data was for the 2014–2015 academic year, we framed our post-interview survey to capture school conditions during the same time period, ensuring that we are comparing information from the same period. Of the ten FMs that received the extensive post-interview survey, three completed all of the questions, providing us detailed information on eight schools, cumulatively, that they managed. Please refer to Fig. 6 to observe the scores for all 569 school buildings to further contextualize these results.

5.3. Results & discussion

In this section, we look primarily at the responses from 10 post-interview survey questions, which are highlighted in their abbreviated format in Table 2 with their relevant KPIs. We asked the FMs about school-specific information that we could then compare with our KPIs. The following list first highlights the post-interview survey question

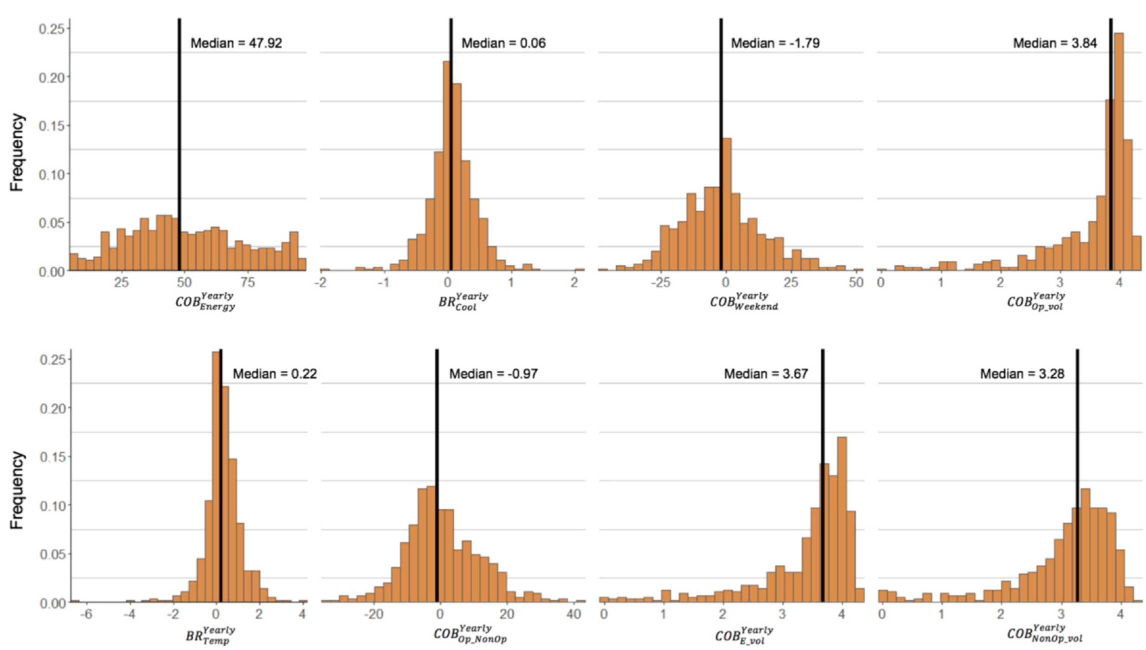


Fig. 6. Histograms for each yearly KPI showing their distributions and median values for all 569 schools.

Table 2
Post-interview survey questions (shortened) and their relevant KPIs. These questions are used to compare to the performance labels outputted from the MEMPI framework.

Question Number	Question (abbreviated)	Relevant Yearly KPIs	Relevant Non-Yearly KPIs	Range
Q1	Overall energy efficiency?	COB^{Yearly}_{Energy}		1–5
Q2	Building envelope efficiency?	COB^{Yearly}_{Energy}, $COB^{Yearly}_{E\_vol}$, BR^{Yearly}_{Temp}, BR^{Yearly}_{Cool}		1–5
Q3	Lighting controls capability?	COB^{Yearly}_{Energy}, $COB^{Yearly}_{Op\_NonOp}$, $COB^{Yearly}_{Weekend}$	COB^{Daily}_{Energy}, $COB^{Weekly}_{Weekend}$, $COB^{Daily}_{Op\_NonOp}$	1–5
Q4	Lighting system efficiency?	COB^{Yearly}_{Energy}, $COB^{Yearly}_{E\_vol}$, $COB^{Yearly}_{Op\_vol}$, $COB^{Yearly}_{NonOp\_vol}$		1–5
Q5	HVAC controls capability?	COB^{Yearly}_{Energy}, $COB^{Yearly}_{Weekend}$, BR^{Yearly}_{Temp}, BR^{Yearly}_{Cool}, $COB^{Yearly}_{Op\_NonOp}$	$COB^{Weekly}_{Weekend}$, $COB^{Daily}_{Op\_NonOp}$	1–5
Q6	HVAC system efficiency?	COB^{Yearly}_{Energy}, $COB^{Yearly}_{E\_vol}$, $COB^{Yearly}_{Op\_vol}$, BR^{Yearly}_{Temp}, BR^{Yearly}_{Cool}		1–5
Q7	AC performance during hot days?	COB^{Yearly}_{Energy}, BR^{Yearly}_{Temp}, BR^{Yearly}_{Cool}		1–4
Q8	How much do occupants affect energy performance?	COB^{Yearly}_{Energy}, BR^{Yearly}_{Temp}, BR^{Yearly}_{Cool}, $COB^{Yearly}_{E\_vol}$, $COB^{Yearly}_{Op\_vol}$, $COB^{Yearly}_{NonOp\_vol}$, $COB^{Yearly}_{Weekend}$, $COB^{Yearly}_{Op\_NonOp}$	COB^{Daily}_{Energy}, $COB^{Daily}_{Op\_NonOp}$	1–5
Q9	Building equipment automation capability?	COB^{Yearly}_{Energy}, $COB^{Yearly}_{Op\_NonOp}$, $COB^{Yearly}_{Weekend}$	$COB^{Daily}_{Op\_NonOp}$, COB^{Daily}_{Energy}, $COB^{Weekly}_{Weekend}$	1–4
Q10	Building equipment age?	COB^{Yearly}_{Energy}, $COB^{Yearly}_{E\_vol}$, $COB^{Yearly}_{Op\_NonOp}$	COB^{Daily}_{Energy}, $COB^{Weekly}_{Weekend}$, $COB^{Daily}_{Op\_NonOp}$	1–50

and then the corresponding related KPIs. All questions are rated on a 1–5 scale where 5 is very energy *efficient* and 1 is very *inefficient,* except for questions seven, nine, and ten (Q7, Q8, Q10). Q7 (i.e., AC performance during hot days?) is ranked on a 1–4 scale where 1 indicates that the AC unit always breaks down and 4 indicates that the AC unit never has problems. Q9 is also on a 1–4 scale where 1 indicates that no systems are automated while a 4 indicates that all systems are automated. Finally, Q10 asks for the average building equipment age and therefore is on a scale of 1–50. Detailed analysis of how KPIs aligned with each facility manager's perceptions is discussed in detail in the following section for each case study school.

5.3.1. Facility manager a (FM-A)

Background: FM-A has been employed by the school district for less than four years and has less than four years of experience as a facility manager. Managing over ten schools, FM-A views energy efficiency as one of his/her main job responsibilities and ranks energy efficiency as very important to the school district's leadership team. FM-A pointed to occupant behavior as a large source of energy waste, and receives feedback from faculty, staff, and students about 2–3 times a week, thus the volatility KPIs and trends of daily KPIs can help this FM better track occupancy behavior. The FM provided school specific information for two of the schools he/she (school buildings A1, A2) manages and as a result we were able to compare the results of the MEMPI framework to these two specific schools. Fig. 7 shows the yearly energy performance label produced from the MEMPI framework and compares them to the survey answers provided by facility manager A.

School Building A1: The KPIs for school building A1 indicate that this school may have an inefficient AC unit and/or building envelope due to the fact that BR^{Yearly}_{Temp} is over one standard deviation away from the mean. Additionally, the $COB^{Yearly}_{Op\_NonOp}$ KPI also indicates that the building performs worse during the operating state than the non-operating state. Post-interview survey responses from FM-A paint a somewhat similar picture. Though FM-A rated the building as having average efficiency (Q1 scored 3/5), they also indicated that the building had a poor building envelope (Q2 scored 2/5), very poor lighting controls (Q3 scored 1/5), and said that the AC unit breaks down often during hot days (Q7 scored 2/4). FM-A also indicated that the building equipment is fairly old, with an average age of equipment around 30 years. A deeper analysis of the $COB^{Daily}_{Op_NonOp}$KPI for school building A1 (Fig. 8a) reveals that the second half of the year has a worse operational performance, especially in the summer months, further pointing to an inefficient AC unit. Using the $COB^{Daily}_{Op_NonOp}$KPI, FM-A knows which days have worse operational performance and can more quickly identify potential operational issues in the building before the end-of-month utility bill.

School Building A2: For school building A2, the COB^{Yearly}_{Energy} KPI indicates that the building is very efficient (with over 1.5 standard deviations away from average) and has relatively low occupant-driven energy issues as $COB^{Yearly}_{E\_vol}$, $COB^{Yearly}_{Op\_vol}$, and $COB^{Yearly}_{NonOp\_vol}$ are all more than 1.5 standard deviations away from the average. However, FM-A post-interview survey responses indicated that the building is relatively inefficient (Q1 scored 2/5), has both poor HVAC unit and HVAC controls (Q5 and Q6 scored 2/5) and that the AC units breaks down often during hot days (Q7 scored 2/4). FM-A did also note that the equipment in this school is much newer than school building A1, with an average age of equipment around 20 years. Fig. 8b shows the COB^{Daily}_{Energy}KPI for every day in the year for school building A2, where the summer months have a relatively worse performance than the rest of the year—an important insight that was missed by the FM.

Summary: The perceptions of FM-A are mixed when compared to the insights found using the KPIs. For school building A1, FM-A felt that the AC unit was somewhat problematic, which was corroborated with the result from the BR^{Yearly}_{Temp} KPI. Although FM-A felt the building had average energy performance, they proceeded to indicate that nearly every system in the building faced significant issues, which is corroborated with the COB^{Yearly}_{Energy} KPI. However, FM-A had very different perceptions about school building A2 than what the KPIs indicated. FM-A thought the school was inefficient but the KPIs indicate differently. Nevertheless, FM-A also recognized that the equipment was much newer than school building A1 which could be a source of the overall more energy efficient building. Given the relative inexperience of FM-A (< 4 years' experience), and that about half of what FM-A perceived to be true was reflected in the KPIs, our framework could help FM-A corroborate intuitions (e.g., the poor building envelope of school building A1) or investigate where his/her perceptions of building operations are misaligned with the data (e.g., the overall efficiency of school building A2).

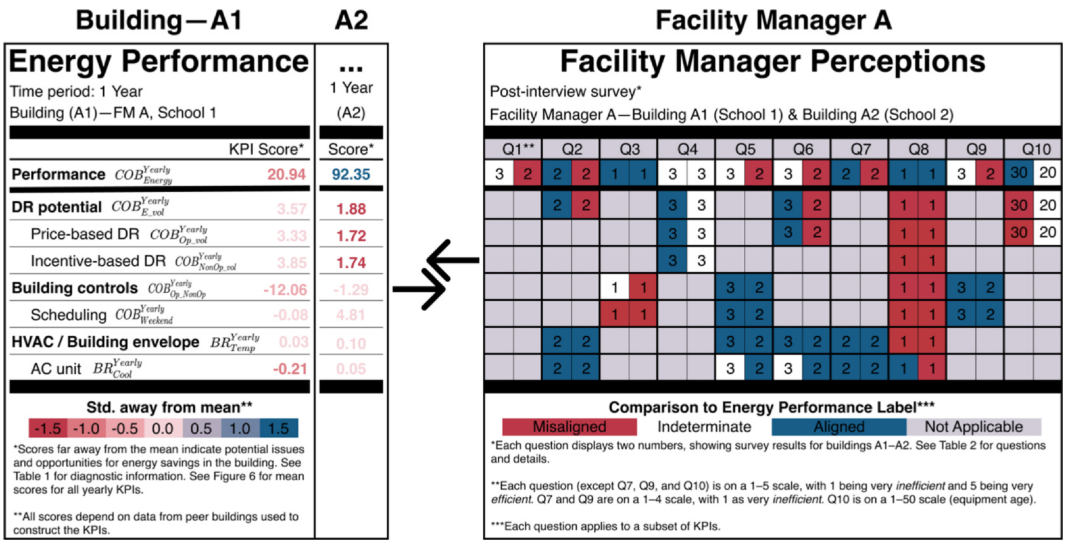


Fig. 7. Comparing the energy performance labels—produced by the MEMPI framework—against the perceptions of facility manager A for two buildings. The cells are highlighted to indicate how far away each value is from the mean value for that KPI for all 569 school buildings.

5.3.2. Facility manager B (FM-B)

Background: FM-B has much more experience than FM-A with 8–12 years in his/her current job and 20 + years as a facility manager. FM-B is responsible for the operation of over 35 buildings. FM-B stated that energy efficiency is a secondary, or tertiary, responsibility and is somewhat important to the district administration staff. However, FM-B felt that energy efficiency is very important to him/her, and received feedback from faculty, staff, and students around 2–3 times per month. We received school specific information for two of the buildings in FM-B's portfolio (B3 and B4). Fig. 9 shows the yearly energy performance labels produced from the MEMPI framework and compares them to the survey answers provided by facility manager B.

School Building B3: Inspecting the KPIs displayed in Fig. 9, school building B3 is indicated to be inefficient with a COB_{Energy}^{Yearly} KPI score of 23.84, have a well-working AC unit with a BR_{Temp}^{Yearly} KPI score of 0.66, and have poor weekday performance compared to weekends with a $COB_{Weekend}^{Weekly}$ KPI score of −15.96. Based upon the post-interview survey results, FM-B perceives this building to have average efficiency (Q1 scored 3/5), thinks the building equipment can handle most hot days (Q7 scored 3/4), but wrote that many building systems are the sources of frequent complaints. Most of the equipment in the building has automated controls (Q9 scored 3/4), as the equipment is relatively new with an average age of just 5 years (Q10). Fig. 10a shows COB_{Energy}^{Daily} for school building B3, where weekday and weekend days are differentiated by shape, the blue line is the moving average, and the orange line is the average score which is equal to COB_{Energy}^{Yearly}. School building B3 and B4 show the same pattern in their moving averages, where they both experience spikes in scores in February, April, October, and December. These times align with the school vacations for this district and highlights the large effect that occupants have on school energy consumption.

School Building B4: For school building B4, the KPIs suggest that the building is inefficient with a COB_{Energy}^{Yearly} KPI score of 27.63 and has a non-operating state that is worse than the operating state with a $COB_{Op\_NonOp}^{Yearly}$ KPI score of 11.46 (more than one standard deviation below average). The post-interview survey results share a similar story where FM-B perceives this building to have average efficiency (Q1 scored 3/5), believes building equipment can handle most days (Q7 scored 3/4) but, again, states that most systems are the sources of frequent complaints. Similar to school building B3, school building B4 has mostly automated controls in-place where building equipment is about 5 years old on average. Fig. 10b shows $COB_{Op\_NonOp}^{Daily}$ for school building B4, where the blue line is the moving average and the orange line is the average score which is equal to $COB_{Op\_NonOp}^{Yearly}$. This plot shows that the operating performance is superior to the non-operating performance, which aligns with the perception that the HVAC equipment keeps up on most days but is in contrast to the perception that the building equipment is often faulty.

Summary: The perceptions of FM-B, based on the completed post-interview survey, of the schools that they manage are somewhat inconsistent. For both buildings, FM-B ranked the buildings as being average in terms of energy efficiency, but then proceeded to state that nearly all the building systems are the source of frequent complaints. The KPIs suggest that both buildings are fairly energy inefficient, with

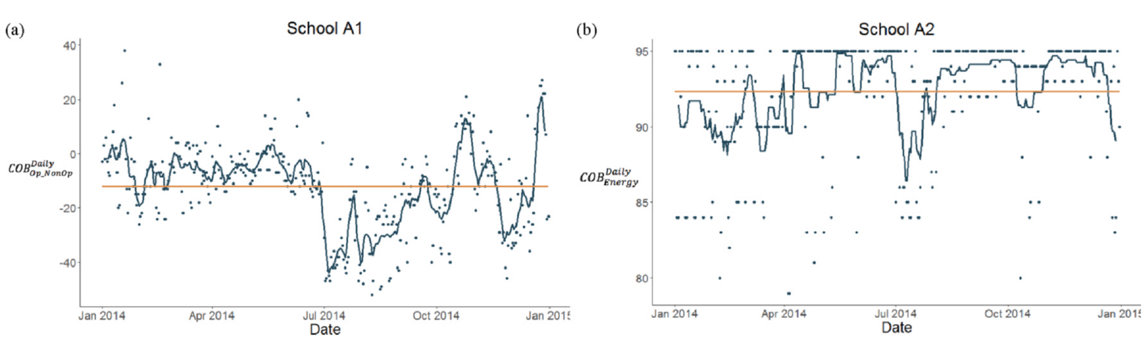


Fig. 8. Scatterplots of School A1 and School A2 with (a) highlighting trends related to operational performance ($COB_{Op\_NonOp}^{Daily}$) and (b) highlighting trends related to seasonality of overall efficiency (COB_{Energy}^{Daily}). The blue lines are the 7-day moving average while the orange lines are the average scores, which correspond to the yearly KPI equivalent.

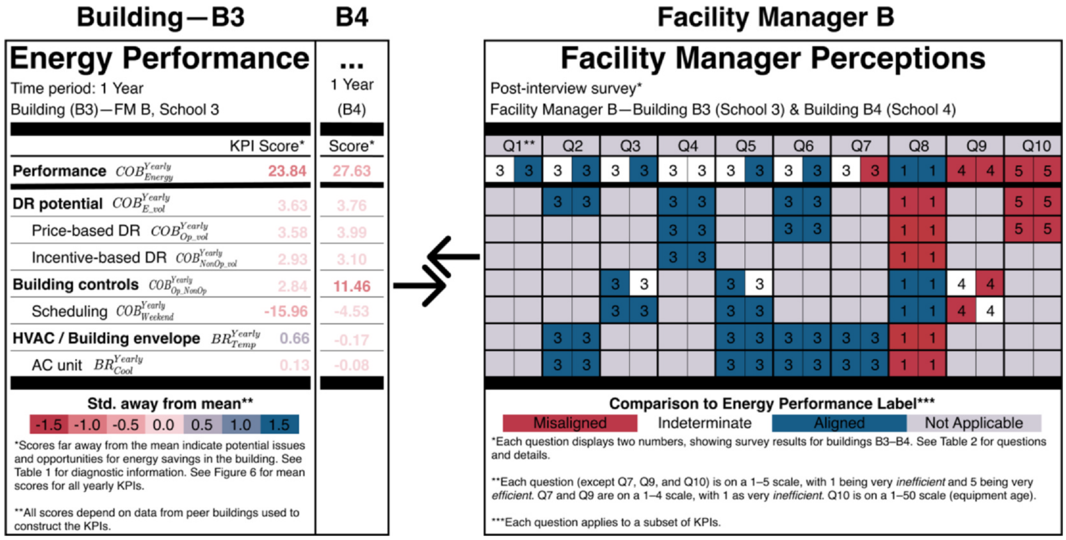


Fig. 9. Comparing the energy performance labels—produced by the MEMPI framework—against the perceptions of facility manager B for two buildings. The cells are highlighted to indicate how far away each value is from the mean value for that KPI for all 569 schools.

school building B3 having more issues that arise during weekdays while school building B4 instead appears to have issues during non-operational hours. Neither of these issues were identified by the facility manager in the post-interview survey. Although FM-B has significantly more experience than FM-A, they are also responsible for managing nearly four times as many schools so may not be as aware of energy issues occurring in each of their buildings. During the interviews, facility managers expressed their desire for quick, easily understood, and contextualized feedback about the energy performance of their buildings, which can be achieved by displaying these KPIs, and is further supported by previous work [3,48]. While post-interview survey results from FM-B were somewhat inconsistent, it highlights the fact that FMs are often not aware of issues in their buildings and need a tool like the MEMPI framework.

5.3.3. Facility manager C

Background: FM-C has less than four years of experience as an FM at the current school district but has between 4 and 8 years of experience as a FM in general. FM-C is responsible for managing just over 20 schools in the district. Building energy efficiency is sometimes considered part of their responsibilities and is considered somewhat important by the district administration. FM-C says that maximizing energy performance of their buildings is important and says that they receive energy feedback daily from faculty, staff, and students. We received school-specific information for four building in FM-C's portfolio. Fig. 11 shows the yearly energy performance labels produced from the MEMPI framework and compares them to the survey answers provided by facility manager C.

School Building C5: For school building C5, the KPIs suggest that the building is relatively inefficient with a COB_{Energy}^{Yearly} KPI score of 31.62, has a worse non-operating state than operating state with a $COB_{Op\_NonOp}^{Yearly}$ KPI score of 6.29, and has poor weekday performance compared to weekends with a $COB_{Weekend}^{Weekly}$ KPI score of -18.18, which can also be seen in Fig. 12a. In the post-interview survey, FM-C perceived the building to be inefficient (Q1 scored 2/5) and have poor HVAC, HVAC controls, building envelope, and lighting (Q2–Q6 all scored 2/5). The FM stated that the equipment is old, with an average age of about 30 years (Q10), but that the equipment could handle most hot days (Q7 scored 3/4). The KPIs corroborate the FM's perception about this school being inefficient and having poor control systems, leading to the poor non-operating state performance and negative $COB_{Weekend}^{Weekly}$ KPI scores.

School Building C6: For school building C6, the KPIs suggest that the building is inefficient with a COB_{Energy}^{Yearly} KPI score of 36.40, is very temperature sensitive with a BR_{Temp}^{Yearly} KPI score of -0.92, and has poor weekday performance compared to weekend performance with a $COB_{Weekend}^{Weekly}$ KPI score of -17.10. In the post-interview survey, FM-C perceives this building to be inefficient (Q1 scored 2/5) and have very poor HVAC, HVAC controls, and a building envelope (Q2, Q5, and Q6 scored 1/5). The temperature sensitivity, due to the poor HVAC system can be seen in Fig. 12b where the COB_{Energy}^{Daily} is plotted against the mean daily temperature and the slope represents BR_{Temp}^{Yearly} indicating that the school is sensitive to hot days. The FM stated that the building equipment is about 40 years old on average (Q10) which may be contributing to the poor HVAC performance and overall building inefficiency.

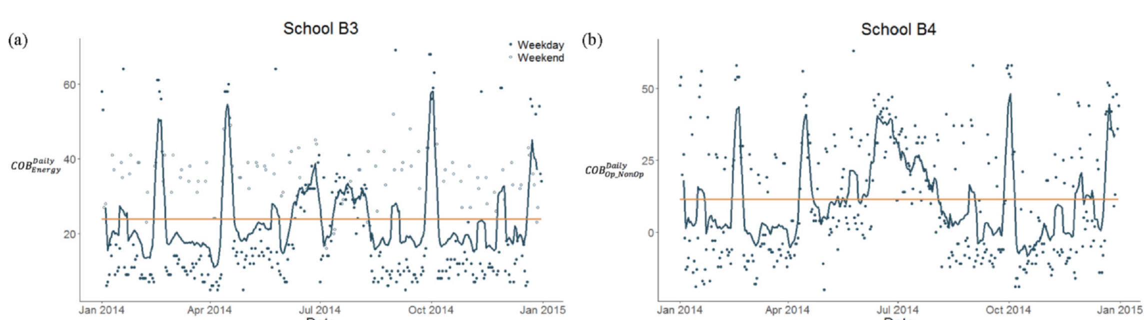


Fig. 10. Scatterplots of School B3 and School B4 with (a) indicating trends related to discrepancies between weekday/weekend overall efficiency (COB_{Energy}^{Daily}) and (b) indicating trends for $COB_{Op\_NonOp}^{Daily}$ scores. The blue lines are the 7-day moving average while the orange lines are the average scores, which correspond to the yearly KPI equivalent.

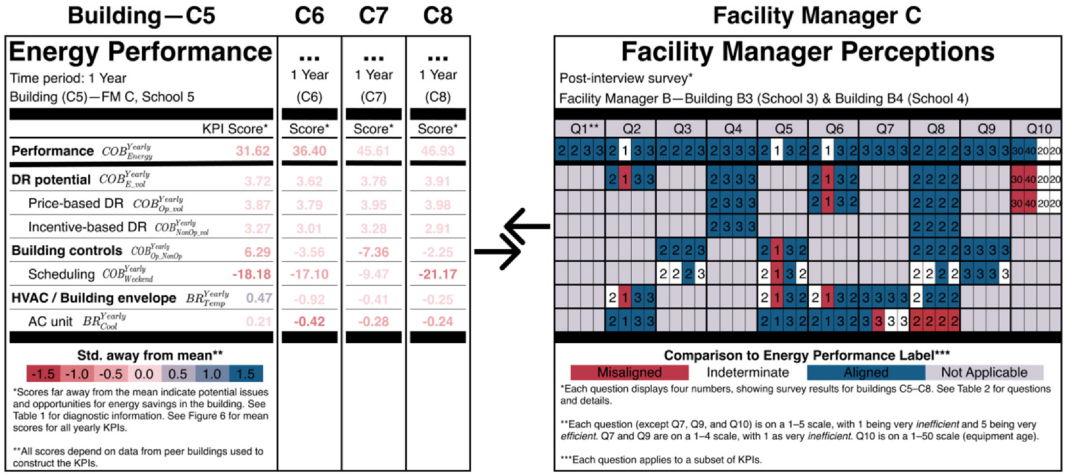


Fig. 11. Comparing the energy performance labels—produced by the MEMPI framework—against the perceptions of facility manager C for four buildings. The cells are highlighted to indicate how far away each value is from the mean value for that KPI for all 569 school.

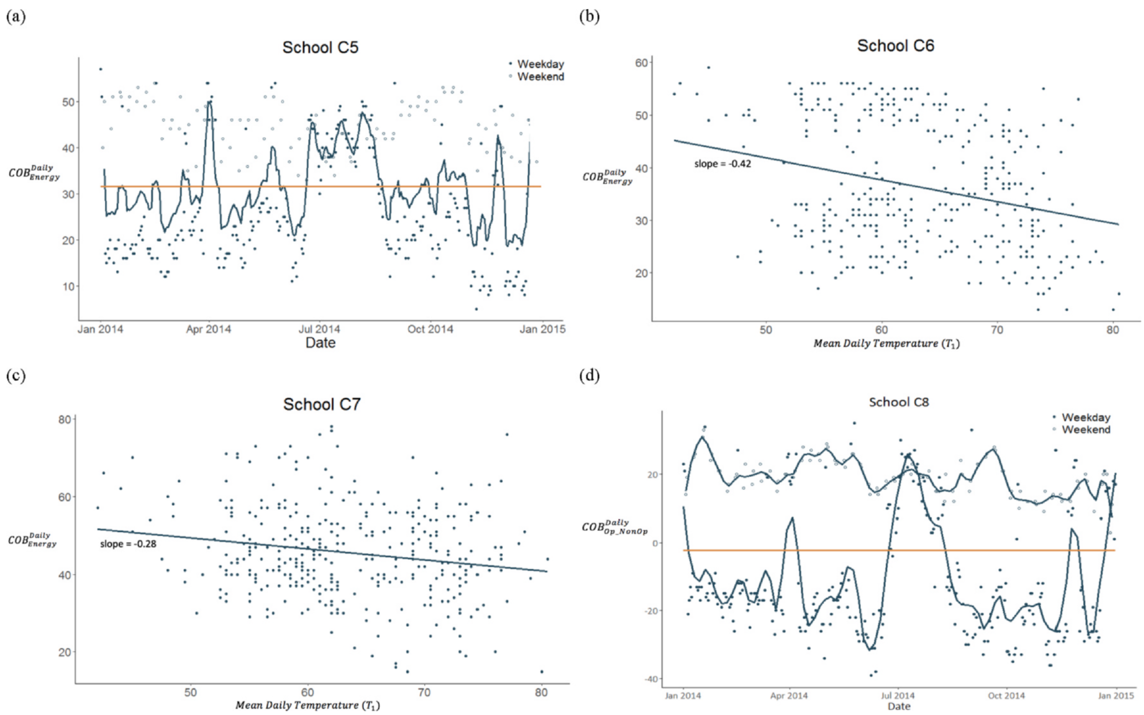


Fig. 12. Scatterplots of Schools C5, C6, C7 and C8. (a) indicates that C5 has poor weekday performance compared to weekends (COB^{Daily}_{Energy}). (b) indicates that C6 is sensitive to hot days. (c) indicates that C7 has a high sensitivity to temperature. (d) indicates that C8 weekdays performs worse during operating hours on weekdays but better on weekends. In (a) and (d), the blue lines are the 7-day moving average while the orange lines are the average scores.

School Building C7: For school building C7, the KPIs suggest the building to be somewhat temperature sensitive and to have an average level of overall energy efficiency. Post-interview survey results show a similar outcome with an average overall efficiency (Q1 scored 3/5), poor lighting controls (Q3 scored 2/5), and ability to handle most hot days (Q7 scored 3/4). Equipment is about 20 years old on average (Q10). Although the FM's perception of the school's overall efficiency aligns with the KPIs, the BR^{Yearly}_{Temp} KPI, as shown in Fig. 12c, indicates a high sensitivity to temperature which seems to have gone unnoticed, as the FM indicated the HVAC equipment to be average (Q6 scored 3/5). This implies that the MEMPI framework may be able to add extra insight on hot days to the FM.

School Building C8: Finally, school building C8 KPIs suggest a building that is slightly temperature sensitive and has large weekday issues despite having an average overall average energy efficiency level. FM-C also feels this building to be average in energy efficiency (Q1 scored 3/5) but have poor HVAC and lighting controls (Q3 and Q5 scored 2/5). Average equipment age for school building C8 is about 20 years (Q10) and FM-C feels like this building can also handle most hot days (Q7 scored 3/4). In Fig. 12d, the $COB^{Daily}_{Op\_NonOp}$ is plotted, where there are two blue-lines which represent the moving average for the weekdays and weekends, respectively. Even though the $COB^{Yearly}_{Op\_NonOp}$ KPI—as represented by the orange line—provides little insight, the $COB^{Daily}_{Op\_NonOp}$ KPI shows a large difference between the weekdays and weekends. Specifically, weekdays have a worse operational state while the weekends have a worse non-operational state, corroborating the FM views that the building has poor lighting controls. Furthermore, these large daily and weekday/weekend differences in energy performance, if promptly given to the FM, can help him/her identify operational inefficiencies which can lead to greater savings.

Summary: With post-interview survey information for four schools, this FM seemed to show high levels of awareness of their schools' performance, as nearly all their post-interview survey results were corroborated with what the KPIs suggested. KPIs for school buildings C5 and C6 indicate that they have below average energy efficiency while indicating school buildings C7 and C8 were average, exactly what was said by FM-C in the post-interview survey. Further, FM-C ranked school C6 as having the worst HVAC and HVAC controls of all four schools, which was reflected in the poor BR_{Temp}^{Yearly} KPI score. Each of the schools had KPIs that suggested worse weekday performance, however, FM-C provided mixed results for inefficient building systems (controls vs. the units themselves) for each of the schools. Inefficient systems could be the primary drivers of inefficiency during the operating or daily hours, or they could be driven by fluctuations in occupant behavior as it has been shown to impact energy usage significantly [52]. Regardless, the KPI framework presented here showed to largely align with the perceptions of FM-C.

5.3.4. Discussion

Overall, our results indicate that the yearly KPIs generally align with the perceptions of the three facility managers. However, several conditions of performance were overlooked by the FMs. Generally, the KPIs highlighted inefficiencies in areas of building equipment automation, lighting controls, and HVAC controls that were not indicated as issues by the three facility managers. The results of our post-interview survey highlight the FMs desire and ability to understand their efficiency performance through quantitative metrics, such as the KPIs from the MEMPI framework, rather than solely on intuition or personal perception. In addition, these KPIs can inform FMs about occupant behavior, which has been shown in numerous studies—and was also corroborated in our interviews—to have a large effect on building energy use. An increased understanding of how energy is being wasted can therefore help drive behavior change [59-62]. Although the proposed framework offers a multitude of benefits, we doubt it would provide more useful insights than an expensive, propriety BMS system that requires installation. Instead, our framework offers a low-cost alternative that aligns incentives from multiple stakeholders by leveraging existing smart meter infrastructure. But without any proper feedback, energy issues go unidentified and wasteful user behavior is unrecognized and unchanged [47].

To track long-term sustainability goals, which is an important practice for achieving energy savings [63], FMs can use the annual and monthly KPIs. To gain insight during times of different occupancy levels and identify issues with control settings, schedules, and different pieces of building equipment, FMs can use the KPIs that measure different operational states. To modify their daily operations, which would ultimately lead to energy savings [47,64,65], FMs can use the timely feedback provided by daily KPIs. As shown by the daily KPIs in Figs. 8, 10, 12 high variability exists in energy performance day-to-day. By measuring performance every day, the MEMPI framework provides deeper insight into building dynamics that are difficult to capture when examining a building for only one day, which is the case for traditional energy audits—they lack insight into building operations over any period of time and only provide a snapshot, though detailed, of energy performance [66]. Because facility managers have many responsibilities—with energy efficiency being only one of them—the MEMPI framework can help them validate the energy performance of their buildings and be alerted to issues that were previously overlooked.

6. Limitations and future work

While our proposed MEMPI framework addresses many of the concerns highlighted by facility managers, its introduction also precipitates several limitations. First, it is important to note that proprietary building management systems—which often include new equipment, devices, sensors, and more—likely offer better opportunities to help facility managers save energy. However, as noted in the interviews, many building owners cannot afford to install such systems meaning that facility managers must do without. Furthermore, the facility manager (FM) views' are predominately associated with their time working at public schools throughout the state of California. FMs that work in the private sector, or in other states or countries, may have a different experience than expressed by our interviewees. The ten participants were also recruited through a snowball sampling process where there is an added risk of self-selection bias in the results, as those who agreed to be interviewed nominated themselves willingly. Finally, given the data that we use for our quantitative analysis and KPI construction is historical, from the 2014–2015 academic year, the post-interview survey results obtained from the FMs rely on their recall of school conditions from two years prior. This prohibited us from validating the insights derived from the daily KPIs. Further user testing can be applied to see if FMs notice more issues that inform their actions and lead to increased energy savings. More up-to-date data would allow us to examine the efficacy of the proposed framework to display some of its other mentioned benefits.

Like other smart meter research, this system could be extended to be more useful to utilities and energy service providers (ESCOs) by finding patterns in the KPIs for large portfolios of buildings. For example, different clustering techniques could be examined to determine if they can help identify customers for various energy efficiency programs. By providing an alternative way of quickly distilling profile load types for customer segregation, utilities may find overlooked patterns in the data that better highlights energy performance that can be used for more targeted programs. Further, this framework could be used with forecasting algorithms to better identify buildings that are likely to experience a decrease in energy performance. For example, daily scores from multiple buildings could be compiled—potentially achieved through clustering algorithms if the dataset is large—thereby creating a panel dataset, which could be used as inputs into a machine learning algorithm that would produce future predicted scores. If the model predicts a decrease in performance, the facility manager could be alerted ahead of time. Such forecasting algorithms could help facility managers with preventative maintenance if they know their buildings are likely to experience breakdowns in equipment in the near future.

7. Conclusions & implications

In this paper, we integrate methods from social, building and data sciences to contribute:

1) Insights into the limitations of current EMIS systems based on in-depth qualitative interviews of 10 facility managers and energy consultants.
2) A novel data-driven Multitiered Energy Management Performance Indicator (MEMPI) framework that takes a facility manager centric viewpoint, embeds lessons learned from the industrial energy management research field to create Key Performance Indicators (KPIs) and leverages high-fidelity data emerging from existing smart meter infrastructure.
3) Comparative analysis of the MEMPI framework using real smart meter data from 569 buildings in California, USA and case studies of 8 case study buildings. Results indicated that the yearly KPIs generally align with the perceptions of the three facility managers. However, several conditions of performance were overlooked by the FMs. Generally, the KPIs highlighted inefficiencies in areas of building equipment automation, lighting controls, and HVAC controls that were not indicated as issues by the three facility managers.

Overall, the proposed MEMPI framework aims to bridge the gap between data-driven energy management models and qualitative domain knowledge held by facility managers to provide more comprehensive insights into the energy performance of buildings. Additionally, the MEMPI framework combines benefits of benchmarking, utility bill tracking, and energy information systems into one low-cost system that aligns incentives of multiple stakeholders including policymakers, building owners, and facility managers. Given that extensive smart meter infrastructure is already in place, we propose using new high-fidelity consumption data to provide a comprehensive view of energy use that can be tailored to inform a range of stakeholders—policymakers, facility managers, and building owners. Policymakers are pushing to have more buildings undergo energy benchmarking given its potential to help enhance building energy efficiency. Facility managers require a low-cost energy management system that provides comparisons to other similar buildings (i.e., benchmarking), quickly provides relevant information about energy performance, and provides actionable insights that can be immediately implemented. Building owners require a system that can provide high-level temporal insights into their building portfolio that can help them prioritize resource allocation and investments related to energy efficiency. In the end, this work aims to provide a framework for better understanding the energy performance of existing building stock and provide a pathway for more informed energy efficiency decision-making across policymakers, facility managers and building owners.

CRediT authorship contribution statement

Jonathan Roth: Conceptualization, Methodology, Writing - original draft. **Howard Alexander Brown IV:** Data Collection, Writing - review & editing. **Rishee K. Jain:** Conceptualization, Project administration, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would also like to thank Eric Hagerman for his support in the execution of this study. This material is based upon the work supported in part by the Stanford School of Engineering under a Terman Faculty Fellowship, the Precourt Institute for Energy and the United States National Science Foundation under Grant Nos. 1642315, 1941695. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Appendix A

Variable Name	Characteristic and Explanation
District Type	District Ownership Type Description
Educational Type	Educational Option Type Description
Charter School	A “Y” or “N” value indicating whether a school is a charter school in the current academic year.
High Grade	Highest grade offered
Enrollment	A total count of K-12 students enrolled (primary or short-term) on Census Day (the first Wednesday in October). These data were submitted to CALPADS as part of the annual Fall 1 submission.
Total free meal count	Of the *Enrollment (K-12)*, a total unduplicated count of students who meet household income or categorical eligibility criteria for free meals based on one or more of the following reasons: (1)applying for the National School Lunch Program (NSLP); (2) submitting alternative household income forms; (3) student homeless or migrant statuses in CALPADS; (4) being “directly certified” through CALPADS as participating in California's food stamp or CalWORKs programs during July - November; or (5) being identified through the weekly CALPADS Foster Matching process or matched by the LEA through the CALPADS online match process as being in Foster Placement or Foster Family Maintenance on Census day. The *Free Meal Count (K-12)* is not displayed on any CALPADS report; however, this count represents the official *Free Meal Count (K-12)* for the academic year.
Percent eligible free	The percent of students eligible for free meals. [*Free Meal Count (K-12)* divided by *Enrollment (K-12)*].
FRPM count	Of the *Enrollment (K-12)*, a total unduplicated count of students who meet household income or categorical eligibility criteria for free or reduced meals (FRPM) based on one or more of the following reasons: (1) applying for the National School Lunch Program (NSLP); (2) submitting alternative household income forms; (1) student homeless or migrant statuses in CALPADS; (4) being “directly certified” through CALPADS as participating in California's food stamp or CalWORKs programs during July - November; or (5) being identified through the weekly CALPADS Foster Matching process or matched by the LEA through the CALPADS online match process as being in Foster Placement or Foster Family Maintenance on Census day.
Percent eligible FRPM	The percent of students eligible for free or reduced price meals (FRPM). [*FRPM Count (K-12)* divided by *Enrollment (K-12)*].
EDP 365	The total cost for the current expense of education.
Current expense ADA	Total ADA (average daily attendance) is defined as the total days of student attendance divided by the total days of instruction. This is the total cost of the ADA.
Current expense per ADA	The total cost per ADA or the EDP_365 divided by the Current Expense ADA.
School Type	The type of school as either “High School”, “Unified”, or “Elementary”
Area (sqft)	Total area of the school building(s) in square feet
Median Household Income	The median household income for the zip code taken from the US Census Bureau
Temperature max	The maximum temperature recorded during the day in Fahrenheit
Temperature min	The minimum temperature recorded during the day in Fahrenheit
Temperature mean	The average daily temperature for the day in Fahrenheit
Dewpoint	The average daily dewpoint temperature for the day in Fahrenheit
Temperature wetbulb	The average daily wetbulb temperature for the day
Heating degree day (HDD)	Number of degrees that the day's average temperature was below 65 degrees Fahrenheit
Cooling degree day (CDD)	The number of degrees that the day's average temperature was above 65 degrees Fahrenheit
Total precipitation	The total amount of precipitation (water equivalent) in inches
Standard pressure	The total standard pressure for the day in Hg
Result speed	The resulting wind speed for the day
Average wind speed	The daily average wind speed in miles per hour
Max5speed	The max speed of wind with a duration of 5 min
Max2speed	The max speed of wind with a duration of 2 min
Temperature mean squared	The average daily temperature squared
Heating degree day squared (-HDD_2)	The heating degree day (HDD) squared

Cooling degree day squared (CDD_2)	The cooling degree day (CDD) squared
Temperature mean natural log	The natural log transformation of the average daily temperature
Weekend	Dummy variable where "1" is a weekend and "0" is a weekday
Enrollment per area	The total enrollment per unit area (Students per square foot)

Appendix B

Table of the survey questions—see attached document.

Appendix C

Interview protocol—see attached document.

References

[1] U.S Energy Information Administration, Annual Energy Outlook 2016 with projections to 2040, Washington DC, 2016. doi:EIA-0383 (2016).
[2] About the Commercial Buildings Integration Program | Department of Energy, US Dep. Energy. (2019). https://www.energy.gov/eere/buildings/about-commercial-buildings-integration-program (accessed March 7, 2019).
[3] Carrie Armel K, Gupta A, Shrimali G, Albert A. Is disaggregation the holy grail of energy efficiency? The case of electricity. Energy Policy. 52 2013:213–34. https://doi.org/10.1016/j.enpol.2012.08.062.
[4] 2016 State of the Efficiency Program Industry: Budgets, Expenditures, and Impacts, Boston, MA, 2017. https://library.cee1.org/system/files/library/13159/2016_CEE_Annual_Industry_Report.pdf (accessed March 7, 2019).
[5] Balaras CA, Gaglia AG, Georgopoulou E, Mirasgedis S, Sarafidis Y, Lalas DP. European residential buildings and empirical assessment of the Hellenic building stock, energy consumption, emissions and potential energy savings. Build. Environ. 2007;42:1298–314. https://doi.org/10.1016/j.buildenv.2005.11.001.
[6] Reyna JL, Chester MV. The Growth of Urban Building Stock: Unintended Lock-in and Embedded Environmental Effects. J. Ind. Ecol. 2015;19:524–37. https://doi.org/10.1111/jiec.12211.
[7] WoodMac: Smart Meter Installations to Surge Globally Over Next 5 Years | Greentech Media, (n.d.). https://www.greentechmedia.com/articles/read/advanced-metering-infrastructure-to-double-by-2024 (accessed November 25, 2019).
[8] Harris N, Shealy T, Kramer H, Granderson J, Reichard G. A framework for monitoring-based commissioning: Identifying variables that act as barriers and enablers to the process. Energy Build. 2018;168:331–46. https://doi.org/10.1016/J.ENBUILD.2018.03.033.
[9] Granderson J, Lin G. Building energy information systems: synthesis of costs, savings, and best-practice uses. Energy Effic. 2016. https://doi.org/10.1007/s12053-016-9428-9.
[10] Shrubsole C, Hamilton IG, Zimmermann N, Papachristos G, Broyd T, Burman E, et al. Bridging the gap: The need for a systems thinking approach in understanding and addressing energy and environmental performance in buildings. Indoor Built Environ. 2018. https://doi.org/10.1177/1420326X17753513.
[11] Goulden M, Spence A. Caught in the middle: The role of the Facilities Manager in organisational energy use. Energy Policy. 2015;85:280–7. https://doi.org/10.1016/J.ENPOL.2015.06.014.
[12] Galvin R. How many interviews are enough? Do qualitative interviews in building energy consumption research produce reliable knowledge? J. Build. Eng. 2015;1:2–12. https://doi.org/10.1016/J.JOBE.2014.12.001.
[13] Sivill L, Manninen J, Hippinen I, Ahtila P. Success factors of energy management in energy-intensive industries: Development priority of energy performance measurement. Int. J. Energy Res. 2013;37:936–51. https://doi.org/10.1002/er.2898.
[14] Poulsen RT, Johnson H. The logic of business vs. the logic of energy management practice: understanding the choices and effects of energy consumption monitoring systems in shipping companies. J. Clean. Prod. 2016;112:3785–97. https://doi.org/10.1016/J.JCLEPRO.2015.08.032.
[15] Li Y, O'Donnell J, García-Castro R, Vega-Sánchez S. Identifying stakeholders and key performance indicators for district and building energy performance analysis. Energy Build. 2017;155:1–15. https://doi.org/10.1016/J.ENBUILD.2017.09.003.
[16] May G, Barletta I, Stahl B, Taisch M. Energy management in production: A novel method to develop key performance indicators for improving energy efficiency. Appl. Energy. 2015;149:46–61. https://doi.org/10.1016/J.APENERGY.2015.03.065.
[17] Boyd G, Dutrow E, Tunnessen W. The evolution of the ENERGY STAR energy performance indicator for benchmarking industrial plant manufacturing energy use. J. Clean. Prod. 2008;16:709–15. https://doi.org/10.1016/j.jclepro.2007.02.024.
[18] D'Oca S, Corgnati SP, Buso T. Smart meters and energy savings in Italy: Determining the effectiveness of persuasive communication in dwellings. Energy Res. Soc. Sci. 2014;3:131–42. https://doi.org/10.1016/J.ERSS.2014.07.015.
[19] Andersson E, Arfwidsson O, Thollander P. Benchmarking energy performance of industrial small and medium-sized enterprises using an energy efficiency index: Results based on an energy audit policy program. J. Clean. Prod. 2018;182:883–95. https://doi.org/10.1016/J.JCLEPRO.2018.02.027.
[20] Curtis J, Walton A, Dodd M. Understanding the potential of facilities managers to be advocates for energy efficiency retrofits in mid-tier commercial office buildings. Energy Policy. 2017;103:98–104. https://doi.org/10.1016/J.ENPOL.2017.01.016.
[21] U.S.E.P. Agency, ENERGY STAR score technical reference, (2014) 1–14. https://portfoliomanager.energystar.gov/pdf/reference/ENERGY STAR Score.pdf (accessed September 7, 2017).
[22] Lavy S, Garcia JA, Dixit MK. KPIs for facility's performance assessment. Part II: identification of variables and deriving expressions for core indicators 2014;32:263–2772. https://doi.org/10.1108/F-09-2012-0067.
[23] Bunse K, Vodicka M, Schönsleben P, Brülhart M, Ernst FO. Integrating energy efficiency performance in production management – gap analysis between industrial needs and scientific literature. J. Clean. Prod. 2011;19:667–79. https://doi.org/10.1016/J.JCLEPRO.2010.11.011.
[24] J. Granderson, M.A. Piette, B. Rosenblum, L. Hu, D. Harris, P. Mathew, P. Price, G. Bell, S. Katipamula, M. Brambley, Energy Information Handbook Applications for Energy-Efficient Building Operations, (2011). https://cloudfront.escholarship.org/dist/prd/content/qt03z8k1v3/qt03z8k1v3.pdf (accessed July 23, 2018).
[25] J. Granderson, Energy Management and Information Systems (EMIS), (n.d.). http://eis.lbl.gov/pubs/emis-tech-class-framework.pdf (accessed June 18, 2018).
[26] Schulze M, Nehler H, Ottosson M, Thollander P. Energy management in industry - A systematic review of previous findings and an integrative conceptual framework. J. Clean. Prod. 2016;112:3692–708. https://doi.org/10.1016/j.jclepro.2015.06.060.
[27] M. Short, M. Dawood, T. Crosbie, N. Dawood, M. Ala-Juusela, Visualization tools for energy awareness and management in energy positive neighbourhoods, 2014.
[28] Xu PP, Chan EHW, Qian QK. Key performance indicators (KPI) for the sustainability of building energy efficiency retrofit (BEER) in hotel buildings in China. Facilities. 2012;30:432–48. https://doi.org/10.1108/02632771211235242.
[29] Rohdin P, Thollander P. Barriers to and driving forces for energy efficiency in the non-energy intensive manufacturing industry in Sweden. Energy. 2006;31:1836–44. https://doi.org/10.1016/J.ENERGY.2005.10.010.
[30] Ahmad MW, Mourshed M, Mundow D, Sisinni M, Rezgui Y. Building energy metering and environmental monitoring – A state-of-the-art review and directions for future research. Energy Build. 2016;120:85–102. https://doi.org/10.1016/J.ENBUILD.2016.03.059.
[31] Li S, Wen J. A model-based fault detection and diagnostic methodology based on PCA method and wavelet transform. Energy Build. 2014;68:63–71. https://doi.org/10.1016/j.enbuild.2013.08.044.
[32] Magoulès F, Zhao H, Elizondo D. Development of an RDP neural network for building energy consumption fault detection and diagnosis. Energy Build. 2013;62:133–8. https://doi.org/10.1016/j.enbuild.2013.02.050.
[33] Yu Y, Woradechjumroen D, Yu D. A review of fault detection and diagnosis methodologies on air-handling units. Energy Build. 2014;82:550–62. https://doi.org/10.1016/j.enbuild.2014.06.042.
[34] Liang J, Du R. Model-based Fault Detection and Diagnosis of HVAC systems using Support Vector Machine method. Int. J. Refrig. 2007;30:1104–14. https://doi.org/10.1016/j.ijrefrig.2006.12.012.
[35] Capozzoli A, Lauro F, Khan I. Fault detection analysis using data mining techniques for a cluster of smart office buildings. Expert Syst. Appl. 2015;42:4324–38. https://doi.org/10.1016/j.eswa.2015.01.010.
[36] Jain RK, Smith KM, Culligan PJ, Taylor JE. Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy. Appl. Energy. 2014;123:168–78. https://doi.org/10.1016/j.apenergy.2014.02.057.
[37] Haben S, Singleton C, Grindrod P. Analysis and Clustering of Residential Customers Energy Behavioral Demand Using Smart Meter Data. IEEE Trans. Smart Grid. 2016;7:136–44. https://doi.org/10.1109/TSG.2015.2409786.
[38] Y. Wang, Q. Chen, T. Hong, C. Kang, Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges, (n.d.). doi:10.1109/TSG.2018.2805***.
[39] Granderson J, Touzani S, Fernandes S, Taylor C. Application of automated measurement and verification to utility energy efficiency program data. Energy Build. 2017;142:191–9. https://doi.org/10.1016/j.enbuild.2017.02.040.
[40] Heo Y, Zavala VM. Gaussian process modeling for measurement and verification of building energy savings. Energy Build. 2012;53:7–18. https://doi.org/10.1016/j.enbuild.2012.06.024.
[41] J. Roth, R.K. Jain, Data-driven, multi-metric, and time-varying (DMT) building energy Benchmarking using smart meter data, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 10863 LNCS (2018) 568–593. doi:10.1007/978-3-319-91635-4_30.

[42] Francisco A, Mohammadi N, Taylor JE. Evaluating Temporal Shifts in City Scale Building. Energy Benchmarks 2018 (accessed April 29, 2019).
[43] Halcomb E, Hickman L. Mixed methods research. Nurs. Stand. 2015;29:41–7. https://doi.org/10.7748/ns.29.32.41.e8858.
[44] A. Castleberry, NVivo 10 [software program]. Version 10. QSR International; 2012., Am. J. Pharm. Educ. 78 (2014) 25. doi:10.5688/ajpe78125.
[45] Free Icons for Everything - Noun Project, (n.d.). https://thenounproject.com/ (accessed June 3, 2020).
[46] Roth J, Rajagopal R. Benchmarking building energy efficiency using quantile regression. Energy. 2018;152:866–76. https://doi.org/10.1016/J.ENERGY.2018.02.108.
[47] Karlin B, Zinger JF, Ford R. The effects of feedback on energy conservation: A meta-analysis. Psychol. Bull. 2015;141:1205–27. https://doi.org/10.1037/a0039650.
[48] Jain RK, Taylor JE, Peschiera G. Assessing eco-feedback interface usage and design to drive energy efficiency in buildings. Energy Build. 2012;48:8–17. https://doi.org/10.1016/j.enbuild.2011.12.033.
[49] Burlig F, Knittel C, Rapson D, Reguant M, Wolfram C. Machine Learning from Schools about. Energ Effi 2017;30:2017) (accessed September.
[50] Zhang J, Florita A, Hodge B-M, Lu S, Hamann HF, Banunarayanan V, et al. A suite of metrics for assessing the performance of solar power forecasting. Sol. Energy. 2015;111:157–75. https://doi.org/10.1016/j.solener.2014.10.016.
[51] E. Delarue, C. De Jonghe, R. Belmans, W. D'haeseleer, Applying portfolio theory to the electricity sector: Energy versus power, Energy Econ. 33 (2011) 12–23. doi:10.1016/J.ENECO.2010.05.003.
[52] Khosrowpour A, Jain RK, Taylor JE, Peschiera G, Chen J, Gulbinas R. A review of occupant energy feedback research: Opportunities for methodological fusion at the intersection of experimentation, analytics, surveys and simulation. Appl. Energy. 2018;218:304–16. https://doi.org/10.1016/J.APENERGY.2018.02.148.
[53] Yang Z, Roth J, Jain RK. DUE-B: Data-driven urban energy benchmarking of buildings using recursive partitioning and stochastic frontier analysis. Energy Build. 2018;163:58–69. https://doi.org/10.1016/j.enbuild.2017.12.040.
[54] Rudberg M, Waldemarsson M, Lidestam H. Strategic perspectives on energy management: A case study in the process industry. Appl. Energy. 2013;104:487–96. https://doi.org/10.1016/j.apenergy.2012.11.027.
[55] Buonomano A, Calise F, Ferruzzi G, Palombo A. Dynamic energy performance analysis: Case study for energy efficiency retrofits of hospital buildings. Energy. 2014;78:555–72. https://doi.org/10.1016/j.energy.2014.10.042.
[56] Murray SN, Walsh BP, Kelliher D, O'Sullivan DTJ. Multi-variable optimization of thermal energy efficiency retrofitting of buildings using static modelling and genetic algorithms - A case study. Build. Environ. 2014;75:98–107. https://doi.org/10.1016/j.buildenv.2014.01.011.
[57] Thollander P, Backlund S, Trianni A, Cagno E. Beyond barriers - A case study on driving forces for improved energy efficiency in the foundry industries in Finland, France, Germany, Italy, Poland, Spain, and Sweden. Appl. Energy. 2013;111:636–43. https://doi.org/10.1016/j.apenergy.2013.05.036.
[58] Landerman LR, Land KC, Pieper CF. An Empirical Evaluation of the Predictive Mean Matching Method for Imputing Missing Values. Sociol. Methods Res. 1997;26:3–33. https://doi.org/10.1177/0049124197026001001.
[59] Yan D, O'Brien W, Hong T, Feng X, Burak Gunay H, Tahmasebi F, et al. Occupant behavior modeling for building performance simulation: Current state and future challenges. Energy Build. 2015;107:264–78. https://doi.org/10.1016/J.ENBUILD.2015.08.032.
[60] D'Oca S, Hong T, Langevin J. The human dimensions of energy use in buildings: A review. Renew. Sustain. Energy Rev. 2018;81:731–42. https://doi.org/10.1016/J.RSER.2017.08.019.
[61] Schultz PW, Nolan JM, Cialdini RB, Goldstein NJ, Griskevicius V. The Constructive, Destructive, and Reconstructive Power of Social Norms. Psychol. Sci. 2007;18:429. https://doi.org/10.1111/j.1467-9280.2007.01917.x.
[62] McCalley LT, Midden CJH. Energy conservation through product-integrated feedback: The roles of goal-setting and social orientation. J. Econ. Psychol. 2002;23:589–603. https://doi.org/10.1016/S0167-4870(02)00119-8.
[63] Becker LJ. Joint Effect of Feedback and Goal Setting on Performance: A Field Study of Residential Energy Conservation. J. Appl. Psychol. 1978;63:428–33 (accessed September 7, 2017). http://psycnet.apa.org/fulltext/1979-09988-001.pdf.
[64] Bull R, Janda KB. Building Research & Information Beyond feedback: introducing the "engagement gap" in organizational energy. management 2017. https://doi.org/10.1080/09613218.2017.1366748.
[65] Houde Sebastien, Todd Annika, Sudarshan Anant, Carrie Armel K. Real-time Feedback and Electricity Consumption: A Field Experiment Assessing the Potential for Savings and Persistence. EJ 2013;34(1). https://doi.org/10.5547/ISSN0195-6574-EJ10.5547/01956574.34.110.5547/01956574.34.1.4.
[66] C. Deb, S.E. Lee, Determining key variables influencing energy consumption in office buildings through cluster analysis of pre- and post-retrofit building data, Energy Build. 159 (2018) 228–245. doi:10.1016/J.ENBUILD.2017.11.007.