Title
Application and evaluation of a pattern-based building energy model calibration method using public building datasets

Journal
Building Simulation, 15(8)

ISSN
1996-3599

Authors
Sun, Kaiyu
Hong, Tianzhen
Kim, Janghyun
et al.

Publication Date
2022-08-01

DOI
10.1007/s12273-022-0891-2

Peer reviewed

Building Technology & Urban Systems Division
Energy Technologies Area
Lawrence Berkeley National Laboratory

Application and Evaluation of a Pattern-based Building Energy Model Calibration Method using Public Building Datasets

Kaiyu Sun[1], Tianzhen Hong[1], Janghyun Kim[2] & Barry Hooper[3]

[1]Lawrence Berkeley National Laboratory, Berkeley, CA
[2]National Renewable Energy Laboratory, Golden, CO
[3]San Francisco Department of the Environment, San Francisco, CA

Energy Technologies Area
March 2022

https://doi.org/10.1007/s12273-022-0891-2

This work was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy,
Building Technologies Office, of the U.S. Department of Energy
under Contract No. DE-AC02-05CH11231.

Application and Evaluation of a Pattern-based Building Energy Model Calibration Method using Public Building Datasets

Kaiyu Sun[1], Tianzhen Hong[1*], Janghyun Kim[2], Barry Hooper[3]

[1]Lawrence Berkeley National Laboratory, Berkeley, California, USA

[2]National Renewable Energy Laboratory, Golden, Colorado, USA

[3]San Francisco Department of the Environment, San Francisco, California, USA

*Corresponding author: thong@lbl.gov

Abstract

Building performance simulation has been adopted to support decision making in the building life cycle. An essential issue is to ensure a building energy simulation model can capture the reality and complexity of buildings and their systems in both the static characteristics and dynamic operations. Building energy model calibration is a technique that takes various types of measured performance data (e.g., energy use) and tunes key model parameters to match the simulated results with the actual measurements. This study performed an application and evaluation of an automated pattern-based calibration method on commercial building models that were generated based on characteristics of real buildings. A public building dataset that includes high-level building attributes (e.g., building type, vintage, total floor area, number of stories, zip code) of 111 buildings in San Francisco, California, was used to generate building models in EnergyPlus. Monthly level energy use calibrations were then conducted by comparing building model results against the actual buildings' monthly electricity and natural gas consumption. The results showed 57 out of 111 buildings were successfully calibrated against actual buildings, while the remaining buildings showed opportunities for future calibration improvements. Enhancements to the pattern-based model calibration method are identified to expand its use for: (1) central heating, ventilation and air conditioning (HVAC) systems with chillers, (2) space heating and hot water heating with electricity sources, (3) mixed-use building types, and (4) partially occupied buildings.

Keywords: model calibration; building energy modeling; EnergyPlus; building performance simulation; monthly energy use

1. Introduction

Existing studies have indicated significant discrepancies between simulated energy use from building energy models and actual measured data. Building performance gap between the predicted and actual energy use is identified as one of the top ten challenges in the building simulation field (Tianzhen Hong, Langevin, and Sun 2018). A few related studies on the performance gaps and exploration on potential reasons leading to the gaps were reviewed. Balaras compared the calculated heating energy consumption from energy performance certificates with the actual energy use for over 8500 dwellings, and found the actual energy use is 44% lower than calculated (Balaras et al. 2016). Yoshino et al. emphasized the importance of human-influenced factors on building energy performance and concluded that the calculated energy performance is not strictly related to the actual energy use due to standardized input data (Yoshino, Hong, and Nord 2017). Yin et al. identified significant discrepancy between initial simulated data and measured energy data on a developed campus building model before calibration (Yin, Kiliccote, and Piette 2016). La Fleur et al. found that the preliminary simulations indicated a gap between the measured and predicted energy use in the case study building (La Fleur, Moshfegh, and Rohdin 2017). Cuerda et al. summarized significant differences between the simulated and actual energy consumption in dwellings with similar characteristics and explored methods for reducing the gap (Cuerda et al. 2020). Wang et al. analyzed the dynamic performance gap between expected and actual energy use of five green buildings in China and got a range of 3.0-53.5% (Wang et al. 2020). This undermines the confidence in adopting building energy simulation in the industry widely, such as during the design (Samuelson, Ghorayshi, and Reinhart 2016), retrofit_(Johnson 2017; Heo, Choudhary, and Augenbroe 2012), commissioning, and operation phases (Coakley, Raftery, and Keane 2014).

To address this issue, building energy model accuracy must be improved to reflect actual building performance, which can be achieved via model calibration. Building energy model (BEM) calibration is a process to tune model input so the simulated results (e.g., annual or monthly energy use) match the measured results based on certain criteria (Reddy 2006). BEM calibration starts with an initial building energy model, weather file, and measured building performance, e.g., annual or monthly electricity and natural gas usage. Calibration can significantly improve the accuracy and reliability of energy models that are used in processes like retrofit design optimization, commissioning (Zibin, Zmeureanu, and Love 2016) and fault detection analysis (Reddy 2006).

Calibration techniques can be grouped into either manual or automated methods. Manual calibration methods include graphical analysis and sensitivity analysis. Examples of methods used for automated calibration include Bayesian analysis, pattern matching, and multi-objective optimization (ASHRAE 2017). Manual calibration methods are labor intensive and require expertise, which is difficult to scale up in a consistent way. Automated or semi-automated calibration methods usually rely on mathematical and statistical techniques, which utilize certain optimization function(s) to reduce the difference between measured and simulated data. An objective function may be used to set a target of minimization; for example, the mean square error between the measured and simulated data. Optimization algorithms like genetic algorithm (GA) and Particle Swarm Optimization (PSO) have been used in several studies for auto-calibration.

Hong et al. used GA to perform the optimization for calibration (Taehoon Hong et al. 2017). Ramos Ruiz et al. adopted a multi-objective genetic algorithm, NSGA-II, to calibrate building envelope performance (Ramos Ruiz et al. 2016). Andrade-Cabrera et al. made augmentations to the PSO-based ensemble calibration method to improve the accuracy of non-opaque energy conservation measures (Andrade-Cabrera, Turner, and Finn 2019). The Autotune project developed by Oak Ridge National Laboratory leverages supercomputing, large simulation ensembles, and big data mining with multiple machine learning algorithms to allow auto-calibration of energy simulations (Garrett and New 2015). Automated calibration is becoming more popular due to more affordable computing resources and tools to support the model calibration.

Depending on the granularity of the available measured building performance data, different calibration approaches can be applied in a more efficient way. If only the annual electricity and gas usage data are available for a building—which is the case for buildings subject to some cities', counties', or states' public energy disclosure or benchmarking rules—calibrating such individual building energy models can easily lead to overfitting. Instead of calibrating each building individually, Chen et al. (2020) developed an automatic method to calibrate buildings at the portfolio level by learning the correlations between crucial model input parameters and the building energy use from the reference building models. They tested the method on a dataset of 112 large office buildings in San Francisco (Chen, Deng, and Hong 2020). Such portfolio-level calibration, or urban building energy model calibration, usually aims to achieve a relatively acceptable accuracy for a group of buildings from statistical perspective, rather than pursuing accuracy for each individual building. If monthly electricity and gas usage data (in utility bills) are available for a building, Bayesian analysis (Heo, Choudhary, and Augenbroe 2012), pattern matching (Sun et al. 2016), and multi-objective optimization methods can be used to calibrate the building energy model at the individual building level. Mathematical techniques are sometimes used to infer hourly energy profile from monthly utility bills (Lamagna et al. 2020). If finer granular data are available (usually limited to a few buildings), e.g., smart meter data with 15-minute electricity usage and daily natural gas usage data, calibration methods including Bayesian analysis and multi-objective optimization can be applied. The building models calibrated with the annual or monthly building energy use data are majorly used as the basis of energy retrofit analysis, while the building models calibrated with finer granular data can be used in applications that require higher accuracy, such as fault detection diagnosis and model predictive control.

The pattern matching approach developed by Sun et al. automates a process to calibrate individual energy models by identifying bias patterns of monthly simulated and measured energy use (Sun et al. 2016). It encompasses more engineering insights and experience than purely mathematical optimization-based methods for auto-calibration via pattern recognition as building energy use shows a seasonal pattern (e.g., summer cooling and winter heating) and the fundamental building physics, e.g., internal heat gains lead to more cooling in summer but less heating in winter. It addresses the issue that pure optimization based on mathematical methods lacks critical inputs from physics and engineering perspectives, which sometimes leads to unreasonable calibrated results. The method's calibration process contains four key steps: (1) running the original pre-calibrated energy model to obtain monthly simulated electricity and gas use; (2) establishing a pattern bias, either universal or seasonal, by comparing load shape patterns of simulated and actual monthly energy use; (3) using pre-programmed logic to select which parameter to tune first based on bias pattern, weather, and input parameter interactions; and (4) automatically tuning the calibration parameters and checking the progress using pattern-fit criteria. The automated calibration algorithm was implemented in the Commercial Building Energy Saver (CBES) (Tianzhen Hong et al. 2015), a web-based building energy retrofit analysis toolkit. The novelty of the developed calibration methodology lies in linking parameter tuning with the underlying logic associated with bias pattern identification. Although there are some limitations (e.g., coverage of building and system types) to the current implementation, the pattern-based automated calibration methodology can be adopted universally as an alternative to manual or hierarchical calibration approaches.

It is critical to understand how effective a method is for calibrating real buildings before deciding to adopt the method in actual projects. However, when a calibration method was developed, it was often tested and demonstrated on a single example building only; either a real building or a prototype building. For example, Raillon and Ghiaus developed an improved Bayesian calibration procedure and tested the whole procedure on a real house model (Raillon and Ghiaus 2018), and Martínez et al. compared the accuracy and robustness of 60 optimization-based calibration approaches on the BESTEST600, a white-box BEM case study defined by ASHRAE 140-2001 (Martínez et al. 2020). Chong and Chao developed a continuous Bayesian calibration method and tested the method on an actual 10-story office building in Pennsylvania (Chong and Chao 2020). Asadi et al. developed an optimization-based framework to calibrate the whole building energy model, and tested the method with a real office building located in Doha, Qatar (Asadi et al. 2019). Yin, Kiliccote, and Piette proposed an automated model calibration procedure that links the model to a real-time data monitoring system that allows the model to be updated any time, and applied the approach to a real campus building for testing (Yin, Kiliccote, and Piette 2016). Lam et al. proposed a calibration method that includes an occupant behavior data mining procedure, and tested the method on an office building in Pittsburgh (Lam et al. 2014). In summary, few studies have applied the developed calibration method to a broader set of real buildings and tested its performance in real practice. It helps to identify the limitations and potential problems by using the method to calibrate a variety of real buildings, which generally cannot be detected from a single building case study. To fill the gap, this study applied the pattern-based calibration method to a set of real buildings to evaluate the method's performance.

In this paper, we describe the pattern-based calibration methodology and workflow. Then we applied the calibration method to 111 small- and medium-sized real office and retail buildings in San Francisco individually to demonstrate the validity and process of generating and calibrating the building energy models. The workflow can be used to produce synthetic smart meter data. Findings, challenges, and future work are also discussed.

2. Methodology

This section presents the methodology of the entire model generation and calibration process, starting from understanding the real building dataset, modeling building stock from actual buildings' metadata, and performing automated pattern-based calibrations on building stock models. Figure 1 illustrates the entire workflow of the calibration process. First, the building dataset from the BayREN Integrated Commercial Retrofits (BRICR) project (Department of Energy 2019) supported by San Francisco's Existing Buildings Ordinance (San Francisco Department of the Environment 2020) was filtered based on applicable building types and data availability, as described in Section 2.1. Second, the pattern-based model calibration method implemented in the CBES tool (Hong et al. 2015) was adopted. A pipeline was set up on top of CBES to automate the calibration process of multiple buildings, where baseline models were developed in CBES using metadata inputs from the San Francisco dataset, and automated calibration was performed on filtered buildings. Lastly, calibration results were analyzed. The automated calibration pipeline can be applied to future data collected through San Francisco's Existing Buildings Ordinance. That ordinance requires annual energy benchmarking using ENERGY STAR Portfolio Manager and energy audits based on procedures defined by ASHRAE energy audit Level 1 (small-sized) or Level 2 (large-sized) for existing nonresidential buildings with 10,000 ft^2 (929 m^2) or more of space that is heated or cooled and existing multifamily residential buildings with 50,000 ft^2 (4,645 m^2) or more of space that is heated or cooled.

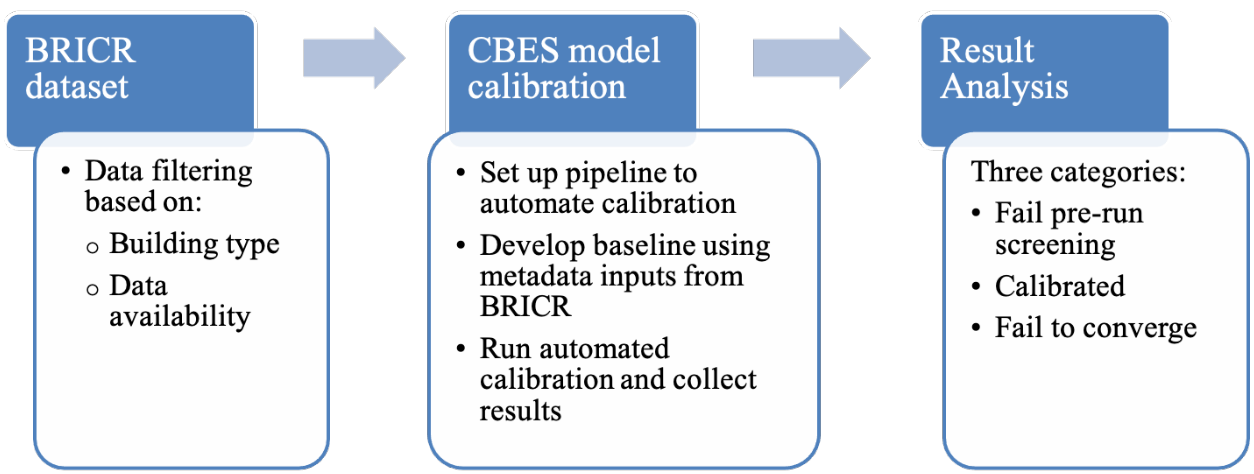


Figure 1. Overall workflow of the study.

2.1. Building dataset used for the case study

San Francisco's Existing Buildings Ordinance (San Francisco Department of the Environment 2020) applies to existing commercial buildings with 10,000 ft^2 (929 m^2) or more of spaces that are conditioned. It has two separate requirements: (1) energy benchmarking for building owners to submit monthly electricity and natural gas bills, and (2) energy audits to ensure buildings receive an energy audit by a qualified energy professional every five years. The billing data for buildings complying with the ordinance are currently being collected, and the dataset used in this study includes bills collected between January to December 2018 with high-level building attributes (building type, gross floor area, number of stories, location, vintage). Actual weather data collected at the San Francisco Airport in 2018 was used in modeling and calibration for better accuracy.

The original building dataset was filtered based on applicable building types and data availability. In terms of applicable building types, the pattern-based model calibration currently only supports small- and medium-sized office buildings, small- and medium-sized retail buildings and their mix-use types that are served by direct expansion (DX) cooling, natural gas heating systems, and natural gas powered service hot water system (SHW). The data needed for the calibration process are a complete year of monthly utility bill and a number of building attributes, including building type, year of construction, number of stories, total floor area, building height. The building attributes are used to generate the baseline models with the CBES toolkit. For the data availability of the utility bill records, there are a significant number of missing data points; some are fixable, while others are not. For example, for Building A (Figure 2), only July's natural gas consumption is missing, but it can be naturally inferred as zero from its adjacent summer months. Buildings with major data losses were not included in the model calibration. After filtering, the original 250 buildings were shortened to 111 buildings for the model calibration experiment.

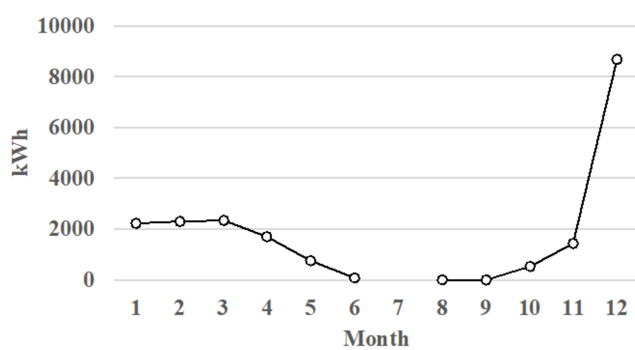
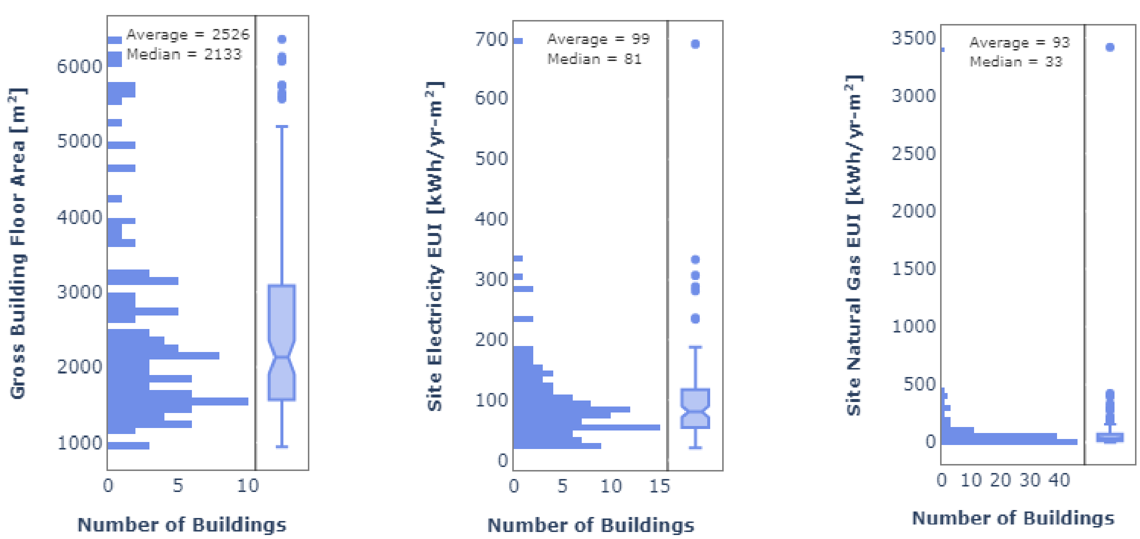


Figure 2. Natural gas bill of Building A with a missing data point.

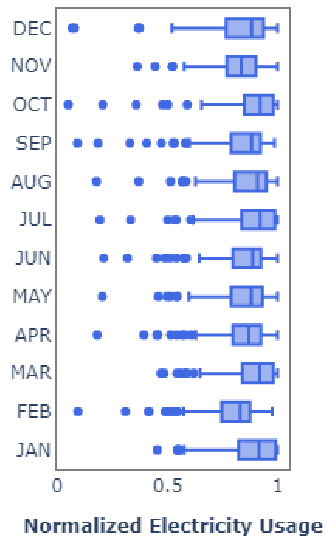
Figure 3 presents the characteristics (gross floor area and normalized annual/monthly energy consumption) of the 111 filtered buildings. Based on the primary usage of each building, 97 buildings are classified as offices, 12 buildings as retail stores, and 2 buildings as strip malls. The gross floor area of these buildings varies from 943 to 6,367 square meters (m^2) with an average floor area of 2,526 m^2.

As shown in the site electricity and natural gas use intensity (EUI) box plots (Figure 3 (b)(c)), two outliers can be visibly identified: the electricity EUI of 691 kilowatt-hours per year per square meter (kWh/yr-m^2) and the natural gas EUI of 3,421 kWh/yr-m^2. These outliers are from two separate buildings: the high electricity usage is from a strip mall, while the high gas usage is from an office building. Assumptions were made for generating building models depending on the primary usage of each building; however, the specific usages were not verified. For example, one multi-story building was classified as an office since most of it is occupied by offices. However, there is a restaurant on the first floor, and its kitchen uses gas equipment intensively. We kept the outliers so that the dataset used for this calibration case study includes the variability of actual buildings that are not captured in the model generation. This may lead to significant discrepancy between prediction and measurement, which is discussed in Sections 3.2 and 3.3.

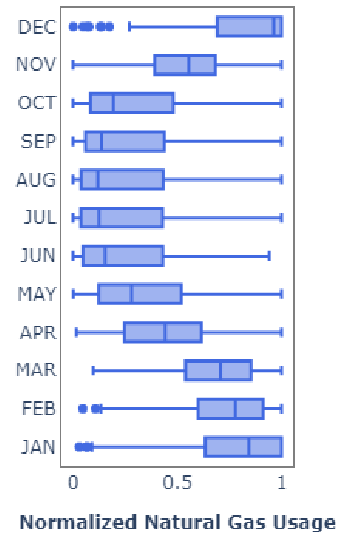
Figure 3 (d)(e) illustrate normalized monthly consumption (electricity and natural gas) for all 111 buildings, represented with box plots. Monthly consumption was normalized based on the maximum monthly consumption of the year for each building. As shown from the outliers (dots in the box plot), the monthly consumption profiles for some buildings are significantly different from the interquartile range. These differences show real variability in building operations. For example, buildings that use electricity instead of natural gas for heating will show high electrical demand in the winter and will show less seasonal dependency on natural gas. While the scope of this study does not include modeling variability in buildings, this is also an important aspect for modeling the building stock.



(a) (b) (c)



(d)



(e)

Figure 3. Characteristics of the 111 filtered buildings: (a) gross floor area, (b) annual site electricity EUIs, (c) annual natural gas EUIs, (d) monthly normalized electricity usage, and (e) monthly normalized natural gas usage.

2.2. Building stock modeling

A detailed building model can be created if sufficient information of the building is available for defining numerous parameters in the building model: location of the building, properties of the building envelope (e.g., roofs, walls, windows), the orientation of the building, type and specification of the HVAC system, building operation schedule, occupant behavior, and other factors. These parameters can be extracted from the building through expensive effort such as performing a site audit, communication with the building owner/operator, and understanding construction/mechanical drawings of the building. While the effort enables the creation of accurate building models, this is not an efficient or cost-effective approach for modeling a building stock.

In order to generate models for a building stock efficiently with reasonable inputs, the CBES Toolkit (Tianzhen Hong et al. 2015) was used to develop baseline models using metadata inputs from the San Francisco dataset. The metadata include mainly high-level attributes of the buildings, i.e., building type, year of construction, number of stories, total floor area, and building height. The CBES Toolkit references different sources for generating representative modeling assumptions in the California region: California Title 24 building energy standards of different vintages (e.g., California Energy Commission 2013), the U.S. Department of Energy reference building models (Deru et al. 2011), and the Database for Energy Efficiency Resources (DEER) (California Public Utilities Commission 2020). These modeling assumptions include envelope performance, internal load settings (occupant, lighting, plug load, hot water), ventilation requirements, HVAC efficiencies, water heating efficiencies, and various operational schedules. They serve as default inputs that correspond to various building attributes above.

The prototype building models represent small- and medium-sized office and retail buildings in all 16 climate zones in California at six vintages that determine the building standard of Title 24.

For the model calibration study, the baseline models were generated based on prototypes, and further customized with other available information such as number of floors, building height, and total floor area.

2.3. Building model calibration

This study adopted the pattern-based automated model calibration method, developed by Sun et al. (Sun et al. 2016) and implemented in CBES (Tianzhen Hong et al. 2015), to demonstrate the validity and the process of the model calibration. The approach compares the monthly electricity and gas consumption profiles generated from building simulations versus actual utility data. The novelty of this methodology relies on the auto-identification of patterns and the logic in selecting to-be-tuned parameters according to the specific patterns. This approach was developed to generate a calibrated model using minimal inputs, which aligns with the limited available data from real buildings that can be used in calibration. The only inputs required for the building model calibration process were the monthly electricity bill, monthly natural gas bill, weather data, and the original baseline model. The process begins by using the input information and relies on parameter auto-tuning that iteratively alleviates discrepancies between the simulated and actual monthly energy use profiles. The selection of the parameter is dependent upon the specific characteristics of the pattern mismatch. This iterative process of shape identification, parameter selection, and parameter tuning occurs until pattern convergence within a specified tolerance is found. Figure 4 shows an overview of the general logic of the automated calibration process.

Generic pattern biases are identified and used at the start of each calibration step. A *pattern bias* refers to the difference between the monthly simulated results and measured data within a whole year. The two distinctly different generic biases are the universal bias and the seasonal bias. A *universal bias* theoretically occurs when the monthly electricity or natural gas bill is consistently higher or lower than the simulated results (Figure 5). A *seasonal bias* has the common characteristics of the monthly electricity or natural gas bill being partially higher or lower than the simulated results depending on the season (Figure 6). The selection of to-be-tuned parameters depends on the pattern biases of both electricity and natural gas, climate type, and baseline characteristics. Five climate types were defined based on heating degree-day and cooling degree-day: (1) hot all year round, (2) cold all year round, (3) hot summer & cool winter, (4) warm summer & cold winter, and (5) mild (Sun et al. 2016). The tuning ranges of the parameters are usually within ±50% of the baseline values. Meanwhile, each parameter has its own physical thresholds predefined in the algorithm. For example, the adjusted lighting power density of office buildings should be between 27.8 W/m^2 representing inefficient incandescent lighting and 6.5 W/m^2 representing energy efficient lighting such as LED. Table 1 shows the general list of 16 changeable parameters. For a certain climate type paired with a certain HVAC system type, a change in a single parameter leads to a specific combination of pattern changes in the monthly electricity and gas consumption. Based on engineering experience, the patterns would first inform higher-level tuning directions, i.e., identify the most relevant end-use categories and the associated subset of parameters. For example, in climates of hot summer & cool winter or warm summer & cold winter, a universal bias in electricity consumption is most likely caused by inappropriate assumption of internal loads such as lighting, plug load, and occupants, rather than that of HVAC systems; while in mild climate, the same bias could also be caused by inappropriate HVAC settings. Then, for different bias patterns coupled with different climate types and HVAC system types, sensitivity analysis was performed on the above identified subset of parameters to help determine which parameter could lead to the largest bias with the same tuning percentage. Sensitivity analysis was performed by tuning different input parameters using EnergyPlus, using a medium-size, three story, rectangular office building (4982 m^2 conditioned floor area), with 5-zones per floor (one central zone and four perimeter zones) as the baseline model. Accordingly, the priority lists under different conditions were established and consolidated in a database, based on which the parameters to tune are selected. Therefore, under each pattern combination, a limited number of possible parameters are identified, and then sorted by the above priority list to determine the next to-be-tuned parameter (Sun et al. 2016).

Table 1. List of parameters that potentially can be adjusted during the calibration process (Sun et al. 2016).

Category	Parameters
Internal loads	Occupant density
	Lighting power density
	Electric equipment power density
	Outdoor air flow rate
	Infiltration rate
HVAC system	Cooling equipment efficiency
	Heating equipment efficiency
	Fan efficiency
	Cooling set point (schedule)
	Heating set point (schedule)
	Economizer status
Construction	Window U-value
	Window SHGC
Schedules	HVAC operation schedule
	Lighting schedule
	Electric equipment schedule

In the CBES calibration module, a prescreening function was implemented to check if the utility data were within the calibration capability. The function calculates the upper and lower boundaries of energy efficiency of each building, which the calibration algorithm is able to reach with the current tuning ranges of all parameters. If the building's utility data are within the boundaries, the program will enter into the automated calibration process and end with two possible results: either the calibration succeeds or fails to converge. Otherwise, the utility bill is beyond the calibration capability, so the program will stop before starting the actual calibration process and continue to the next building.

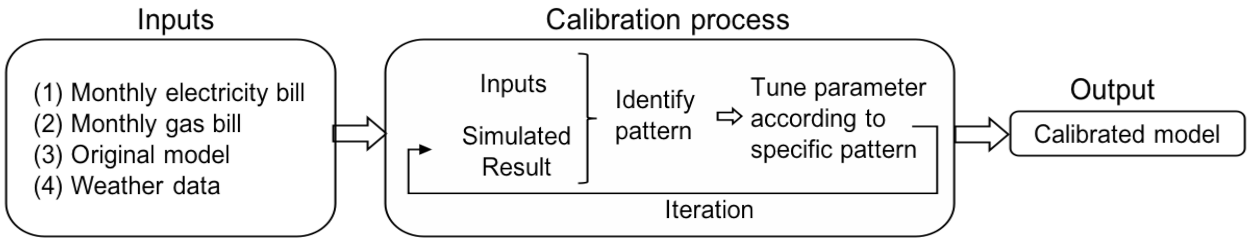


Figure 4. Inputs, calibration process, and outputs for the automatic calibration.

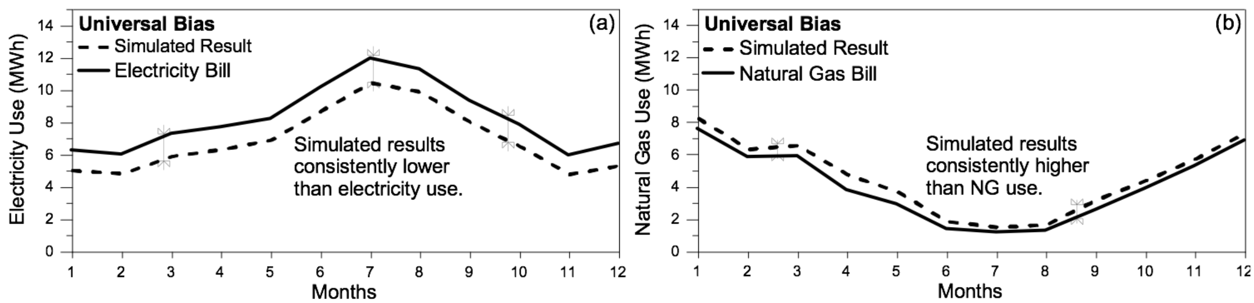


Figure 5. Examples of a Universal Bias existing between simulated vs. actual profiles for (a) electricity use and (b) natural gas use.

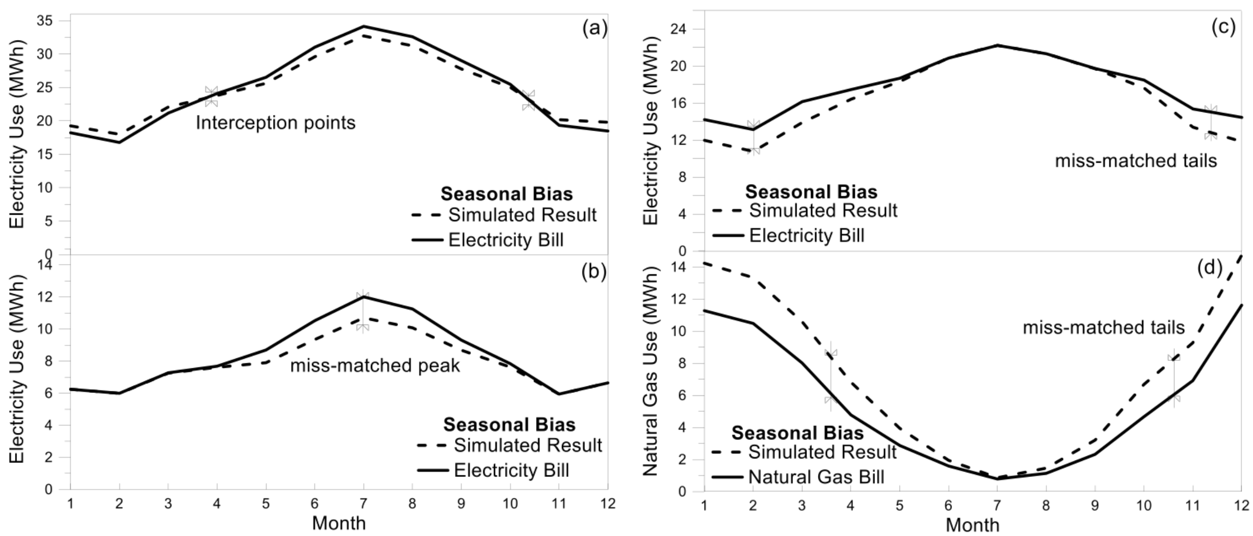


Figure 6. Examples of three most common Seasonal Bias shapes (a) being partially higher and lower than simulated results with interception points, (b) mis-matched peak (c) mis-matched tails for an electricity profile and (d) the miss-matched tails for a natural gas profile.

The normalized mean bias error (NMBE) and coefficient of variation of the root mean square error (CVRMSE) were adopted from ASHRAE Guideline 14 (ANSI/ASHRAE 2014) as the metrics to determine the calibration status. NMBE and CVRMSE are calculated using equations (1) and (2), respectively. NMBE at 5% and CVRMSE at 15% were used as the calibration criteria for monthly data. Considering that the buildings from the dataset are all located in San Francisco, their electricity consumption is usually much higher than their consumption of natural gas, thus natural gas is much more sensitive to discrepancy. In this case, it would be overkill to require that NMBE and CVRMSE of both electricity and natural gas reach the calibration criteria. Therefore, NMBE and CVRMSE of monthly source energy were used as the calibration criteria instead. As source energy is the weighted sum of electricity and natural gas consumption, using source energy as the only metric may lead to discrepancies on both ends. Therefore, an extra constraint is added: that NMBE of electricity consumption, which is the major energy source, should be no larger than 10%. This constraint ensures that both the total source energy and the major energy source fit the measured data well. The site to source energy conversion factors for electricity and natural gas are 3.167 and 1.084, respectively (U.S. Department of Energy 2018).

$$NMBE = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)}{n \times \bar{y}} \times 100 \tag{1}$$

$$CVRMSE = 100 \times [\sum_{i=1}^{n}(y_i - \hat{y}_i)^2/n]^{1/2}/\bar{y} \tag{2}$$

where $\hat{y}_i$ is the predicted results at month i, y_i is the measured data at month i that is used for calibration, n represents the number of months, and $\bar{y}$ is the average of the measured data.

In this study, a pipeline was set up on top of CBES to automate the calibration process of multiple buildings. The high-level building attributes from the San Francisco dataset, including building type, vintage, total floor area, number of stories, and zip code, were used to generate the baseline models as well as the corresponding tuning ranges for calibration parameters in CBES. Automated model calibration was then performed on the 111 selected buildings, and the calibration results were analyzed (see Section 3).

3. Results

First, an initial round of automated calibration was performed on the 111 selected buildings to demonstrate the application and evaluate the effectiveness of the pattern-based calibration approach. According to Section 2.3, there are three possible calibration outcomes: (1) calibration succeeds (the algorithm was able to find a solution that matches simulation results with the utility data based on the defined criteria), (2) calibration fails to converge, this happens when the building's utility bill is within the calibration capability and normal calibration process was performed, however, the algorithm tried all the applicable calibration logic but still did not find a viable solution that meets the calibration criteria, and (3) prescreening fails, i.e., the utility bill is beyond (too high or too low) the calibration capability, the program skips the normal calibration process for this building and continues to the next building. Overall, 47 buildings were successfully calibrated, 16 buildings failed to converge, and 48 buildings' utility bills were beyond calibration capability.

3.1. Successfully calibrated examples

Table 2 and Figure 7 illustrate the detailed calibration process of a successfully calibrated example, Building B. It is a two-story office building with a total floor area of 2,832 m^2 and a total building height of 15.1 meters. The building was built in 1923. After the seven tuning steps listed in Table 2, the building model was successfully calibrated, with the source energy NMBE at -0.8%, CVRMSE at 6.1%, and electricity NMBE at 0.9%.

First, the baseline results were identified as having a universal bias in electricity use and a seasonal bias in gas use. In a mild climate, this bias combination is most likely caused by inaccurate assumptions of the cooling system due to two reasons: (1) the cooling system consumes electricity and operates almost all year round, and its wrong estimation can lead to universal bias in electricity use. Meanwhile, it has minimal impact on heating energy use, which consumes natural gas, or (2) inaccurate assumptions of internal loads, such as lighting, plug load, and occupants can also result in universal bias in electricity use; the gas usage throughout the year will be affected, so internal loads assumptions are second priority. Therefore, the parameters of the cooling system are the top priorities for addressing this bias combination, including cooling COP, economizer, fan efficiency, and cooling setpoint. The sequence of the selected four parameters was predetermined by sensitivity analysis performed on a medium-sized office building (Sun et al. 2016). After tuning these parameters, the NMBE and CVRMSE metrics for the annual source energy shrunk significantly, from 194.4% and 196.1% to 120.8% and 127.7%, but the bias patterns remained at the same combination. Then the internal load parameters, with the second priority, were tuned until a solution was found. The tuned parameters include lighting power density, electric equipment power density, and lighting schedule. The final NMBE and CVRMSE were 4.7% and 13.2%.

Table 2. Tuning steps and the calibration index after each tuning step for Building B.

Steps	Calibration actions	Parameter value before tuning	Parameter value after tuning	NMBE_source (%)	CVRMSE_source (%)	NMBE_electricity (%)
0	Baseline			194.4	196.1	256.2
1	Increase cooling COP (unit: 1)	3.07	4.605	193	203.3	254.5
2	Add economizer	1	0	150.3	159.2	200.7
3	Increase fan efficiency (unit: 1)	0.536	0.804	138.9	147.1	185.1
4	Increase cooling setpoint (unit: °C)	24.48	28.05	120.8	127.7	166.2
5	Decrease lighting power density (unit: W/m^2)	21.39	5	57.8	62.2	82
6	Decrease electric equipment power density (unit: W/m^2)	14.68	5	11	16.9	17.7
7	Decrease lighting schedule	0.265	0.169	4.7	13.2	8.8

	(average running percentage, unit: 1)					

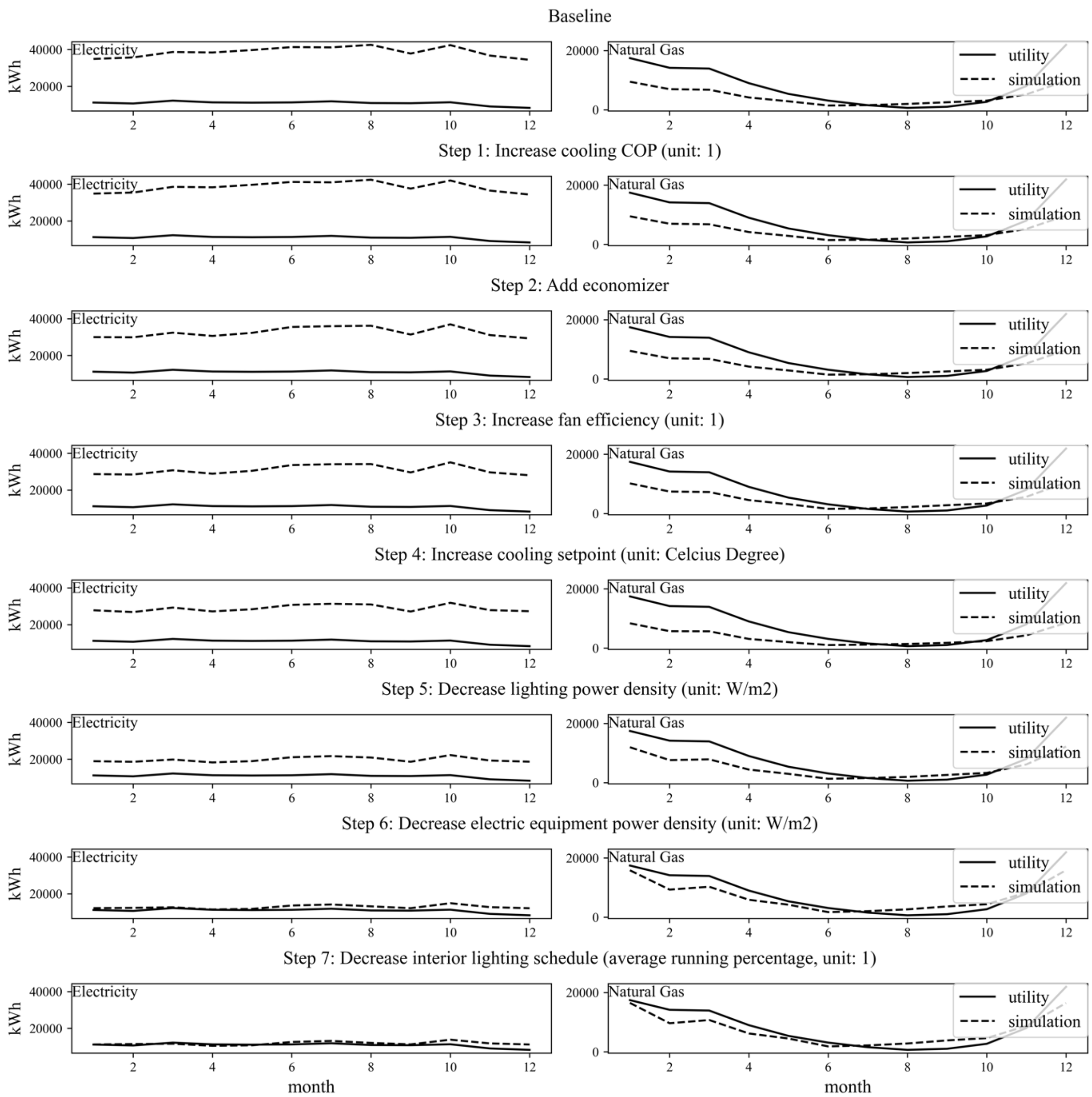


Figure 7. Detailed calibration process and results from each step for Building B.

Figure 8 illustrates the detailed calibration process of another successfully calibrated model, Building F. It is a two-story retail building with a total floor area of 1,263 m^2 and a total building height of 9.6 meters. The building was built in 1920. This building model was successfully calibrated after seven tuning steps. It started with a bias pattern combination different from Building B: universal bias in both electricity and gas use. Specifically, the simulated electricity and gas are both higher than the utility bill data throughout the year. The most common reason for this symptom is that the assumed HVAC operating hours are longer than reality. So the algorithm first tried decreasing the HVAC running time. The new bias pattern after the first tuning step turns out to be the same type as the starting status of Building B, i.e., universal bias in electricity use and seasonal bias in gas use. So the following calibration logic is similar. The final NMBE and CVRMSE of the source energy are 4.5% and 6.4%, and electricity NMBE is 4.1%.

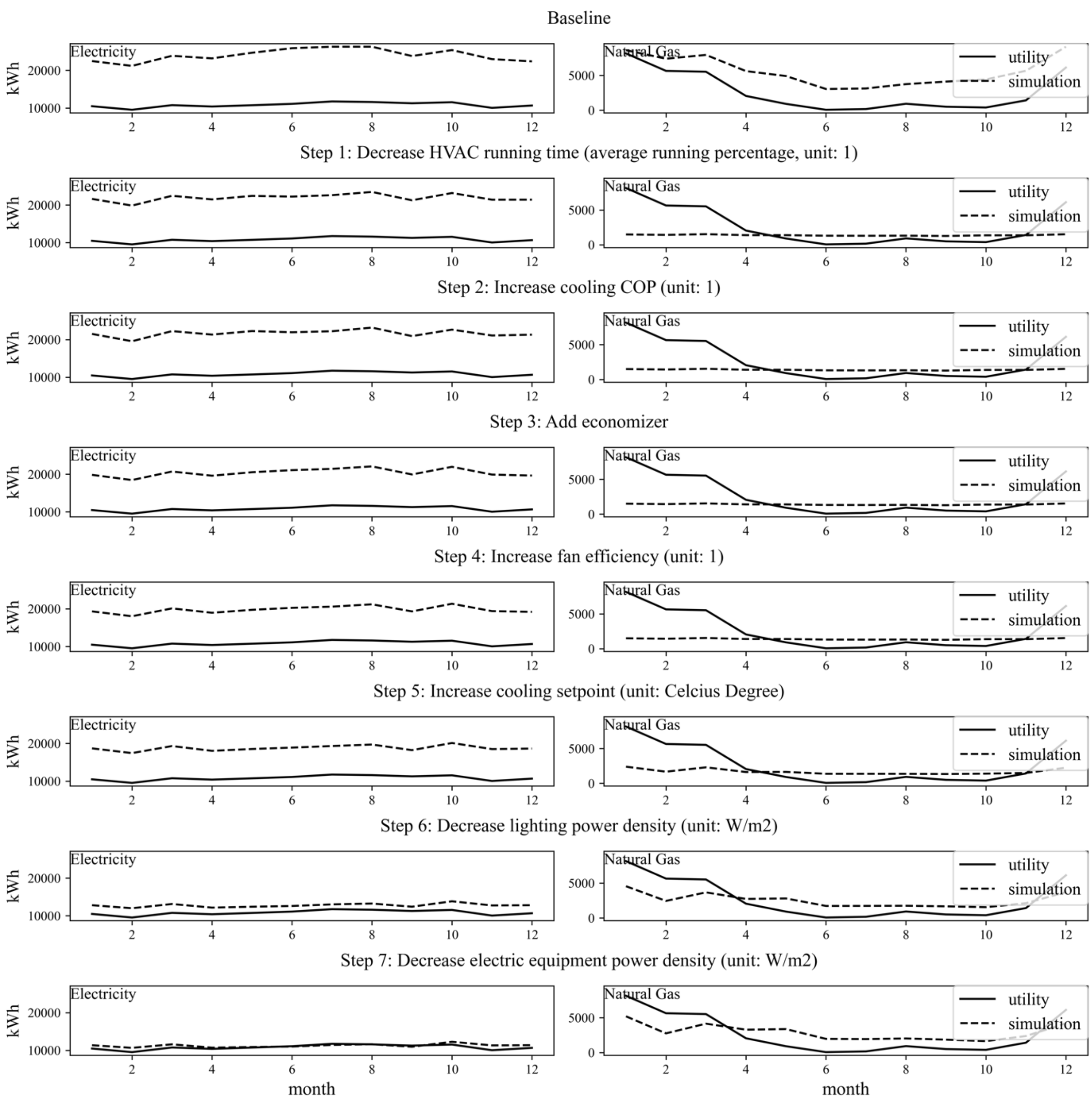


Figure 8. Detailed calibration process and results from each step for Building F.

3.2. Analysis on cases failed to converge

There are potentially three major reasons leading to a failed calibration: (1) an abnormal utility bill, (2) natural gas is not the only source for the space heating or SHW system, and (3) insufficient/mismatched building information. These are discussed below.

(1) Abnormal utility bill

Among the buildings that failed to be calibrated, an abnormal utility bill was common. A utility bill is considered to be abnormal when the energy profile shows a significantly different pattern than consistent normal operation throughout the year, which is the default assumption in the simulation. Figure 9 and Figure 10 illustrate examples of buildings with abnormal utility bills. The electricity consumption of Building C, a medium-sized office building, dropped drastically from above 20,000 kWh in October to less than 2,000 kWh in December, and its natural gas consumption in February was almost triple that of other winter months, i.e., January, November, and December. Similarly, the electricity consumption of Building D, a medium-sized retail building, suddenly dropped to half in September and increased back to normal the following month.

These abnormal utility data could be caused by many reasons, such as incorrect data logging, short-term building shutdown, new tenant moving in, and other factors. It is challenging to practically match simulation results with measured data without knowing the specifics of the building's operation. To solve this problem in the future, we may try to investigate the underlying reason of the abnormal data points, or we may use the utility data from another year that possibly represent more regular energy consumption behavior, if available, for calibration. Nonetheless, it should be noted that we might not be able to model 100% of these buildings because of various unexpected interventions (e.g., temporary vacancy, major maintenance) may occur in real buildings. However, the majority of use cases (e.g., calculating energy savings from potential HVAC retrofit) resulting from building energy modeling will be beneficial even if the baseline model only captures the normal behavior of buildings. Therefore, it is suggested to filter out buildings with abnormal utility data via outlier detection prior to calibration, to improve the efficiency and reliability of this automated calibration method. It also helps to provide building energy models that represent typical and normal behavior of the building stock. On the other hand, the concept of the variability (e.g., daily operating hour differences throughout the year) in buildings should not be disregarded along with the abnormal behavior because these are important aspects of buildings as granular models are being preferred.

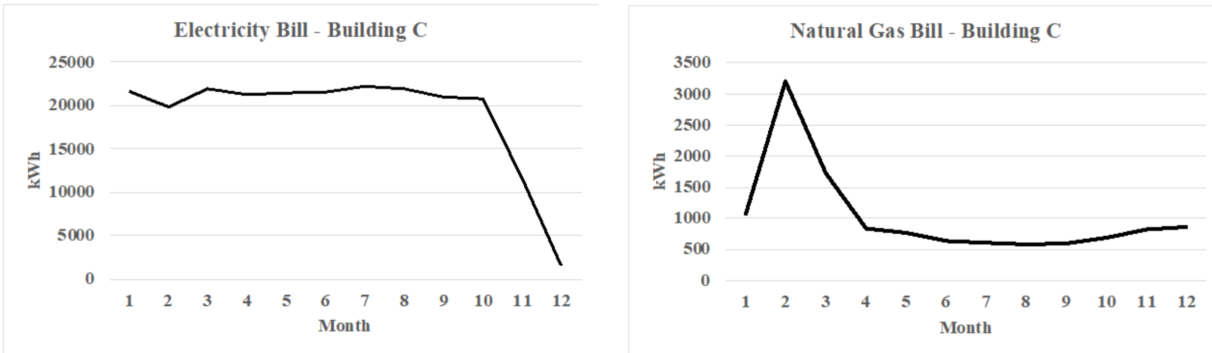


Figure 9. Monthly utility bill of Building C with abnormal data points.

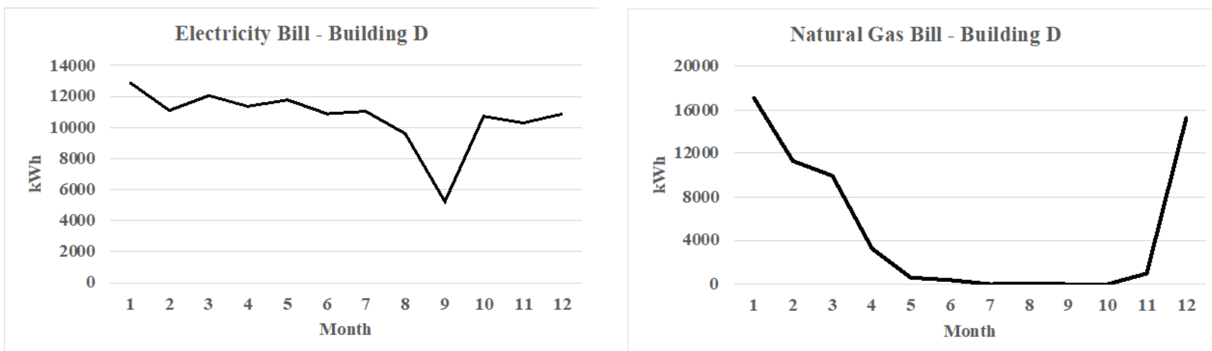


Figure 10. Monthly utility bill of Building D with abnormal data points.

(2) Natural gas is not the only source for the heating or SHW system

The current pattern-based calibration approach is only applicable to buildings that are served by DX cooling and natural gas heating systems, as well as natural gas powered SHW systems; the algorithm is not able to calibrate buildings that use other heating sources. For mild climates like San Francisco (and as shown in Figure 3 (d)), the electricity consumption profile should be either consistent throughout the year or peak in summer, depending on the cooling system operation, and the natural gas should peak in winter and minimize in summer. If the utility data are showing significantly different patterns, it may indicate that natural gas is not the only heating source. For example, as shown in Figure 11, Building E is consuming much more electricity in winter than summer, meanwhile natural gas also peaks in winter. There is a high chance that this building uses both natural gas and electricity for heating/SHW, in which case the current algorithm is not able to find a viable solution.

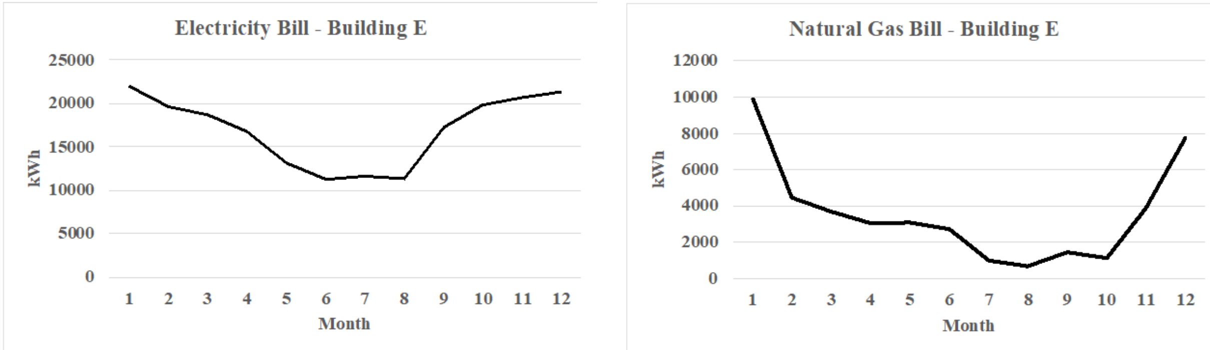


Figure 11. Monthly utility bill of Building E.

(3) Insufficient/mismatched building information

As described in Section 2.2, the baseline model in CBES was generated based on Title 24 standards and customized with available building information, such as total floor area and number of floors in this case. In fact, a number of characteristics that significantly affect buildings' actual energy consumption (such as window-to-wall ratio, building orientation, aspect ratio, exterior shading, shading from adjacent buildings, whether the building shares walls with adjacent buildings, and whether the building has been retrofitted) are not available in the building dataset. For example, if a building shares walls with adjacent buildings (Figure 12), its overall envelope performance is very different from that of stand-alone buildings, which is the default assumption of the CBES baseline. If an old building has been retrofitted, its performance should be much better than the default assumption of an older vintage. The insufficient information makes it more difficult for the simulation to capture the energy pattern of actual buildings. In future work to improve the success rate and accuracy of model calibration, it will be critical to use various techniques to obtain more information to improve baseline model inputs.

It is also possible that the building information was not collected in a way that can be used directly in energy models. For example, in an office and retail mixed-use building, the reported meter represents the retail section, which includes multiple retail stores, but the floor area collected covers the entire building. In this case, the model created based on the reported floor area will overpredict energy consumption compared with the actual meter data. This could be another reason for the problems described in Section 3.3.

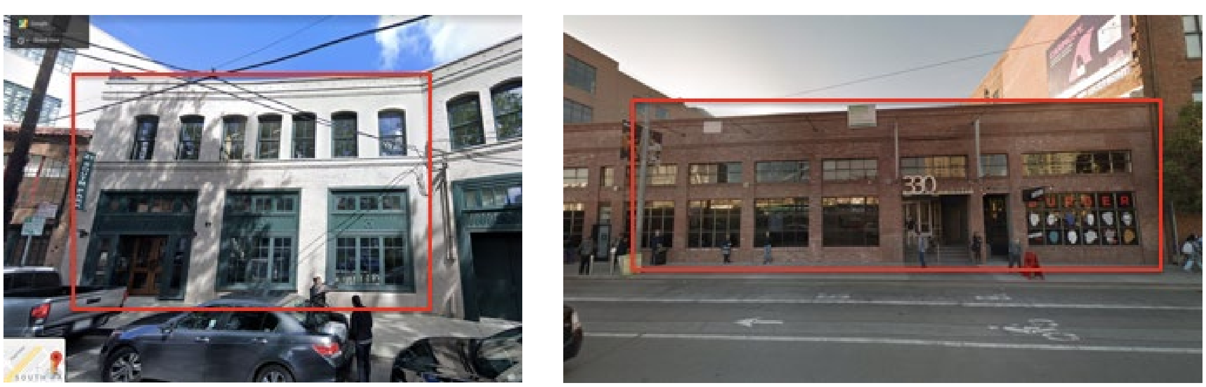


Figure 12. Examples of buildings with adjacent walls (Buildings F and G).

3.3. Analysis on cases beyond calibration capability

When the utility bill is beyond the calibration capability, the prescreening process fails, then the program will skip the normal calibration process for this building and continue to the next building. In this situation, generally two scenarios come into play:

(1) Scenario 1: Simulated baseline energy use is much less than the utility bill

This scenario mostly happens for natural gas – there are 23 occurrences on natural gas and only 1 on electricity in the dataset. In general office and retail buildings, space heating and SHW systems are the only natural gas consumers. Since the prescreening function has already considered the extreme scenarios where all parameters are set to their range limits to reach the largest heating energy and SHW consumption, a reasonable explanation for this excessive gas consumption is that the buildings may have extra gas equipment (e.g., commercial kitchen) and use them very often.

Take Buildings H and I, which did not pass prescreening, as examples. They are both identified as office buildings in the dataset. However, their measured natural gas consumptions are as high as 507 and 22 times, respectively, of the simulation results. These buildings were further investigated through an energy audit. It was found that restaurants actually occupy a large portion of the building areas, both over 30%, as shown in Table 3, although the office is the primary property use. Typical restaurants use a very large amount of natural gas for cooking, hot water, and heating. The restaurant in Building I even uses outdoor gas heaters to heat their patio, which further increases their gas usage. Another type of example is Building J, which is an office building. Its major tenant is a software company, but it has in-house cafeteria as an employee perk, which is not uncommon in the San Francisco Bay Area.

Table 3. Property function area allocation of Buildings H and I.

	Office area (m^2)	Restaurant area (m^2)	Restaurant area percentage (%)
Building H	849	451	34.7
Building I	1,836	887	32.6

(2) Scenario 2: Simulated baseline energy is much larger than the utility bill

This scenario happens for both electricity and natural gas – there are 20 occurrences on electricity and 10 occurrences on natural gas in the dataset. For some buildings, the actual building function is different from what is labeled in the dataset. For example, Building K was identified as a retail building in the dataset; however, an energy audit revealed that the building is in fact half retail and half warehouse. In this case, the building would consume much less in reality than it would be expected to consume as a 100% retail building.

Many other buildings are found to have multiple tenants through Google search of the associated addresses. Since CBES assumes the buildings are fully occupied by default, there is a high chance that these buildings are not fully occupied throughout the year, which leads to overestimation. An experiment was conducted (Section 3.4) to prove this assumption. Another possibility is that the utility bill in the dataset may not be the sum of all the tenants in the buildings. This needs to be verified with further investigation.

3.4. Improved results from considering partial occupancy

We performed an experiment on the assumption that the multi-tenant buildings' low energy use might result from being partially occupied. Take Building L as an example, it is a medium-sized office building and failed prescreening in the first round of calibration runs. Its electricity and natural gas bills are only half and one-eighth of the baseline simulation results, respectively. With the original tuning ranges of the parameters, even the lower boundary of the calibration capability is higher than the utility data. To simulate the partially occupied condition, the lower limits of occupant-related parameters, including occupant density, HVAC operation percentage, and lighting and plug load operation percentage, are further reduced, as listed in Table 4. With the modified tuning ranges, automated calibration was rerun on Building L, and it was successfully calibrated, as illustrated in Figure 13.

Table 4. Modification of occupancy-related parameters.

	Original lower limit of tuning range	Reduced lower limit of tuning range
Occupant density (person/m^2)	0.05	0.01
HVAC operation percentage (%)	23	0
Lighting and plug load operation percentage (%)	10	0

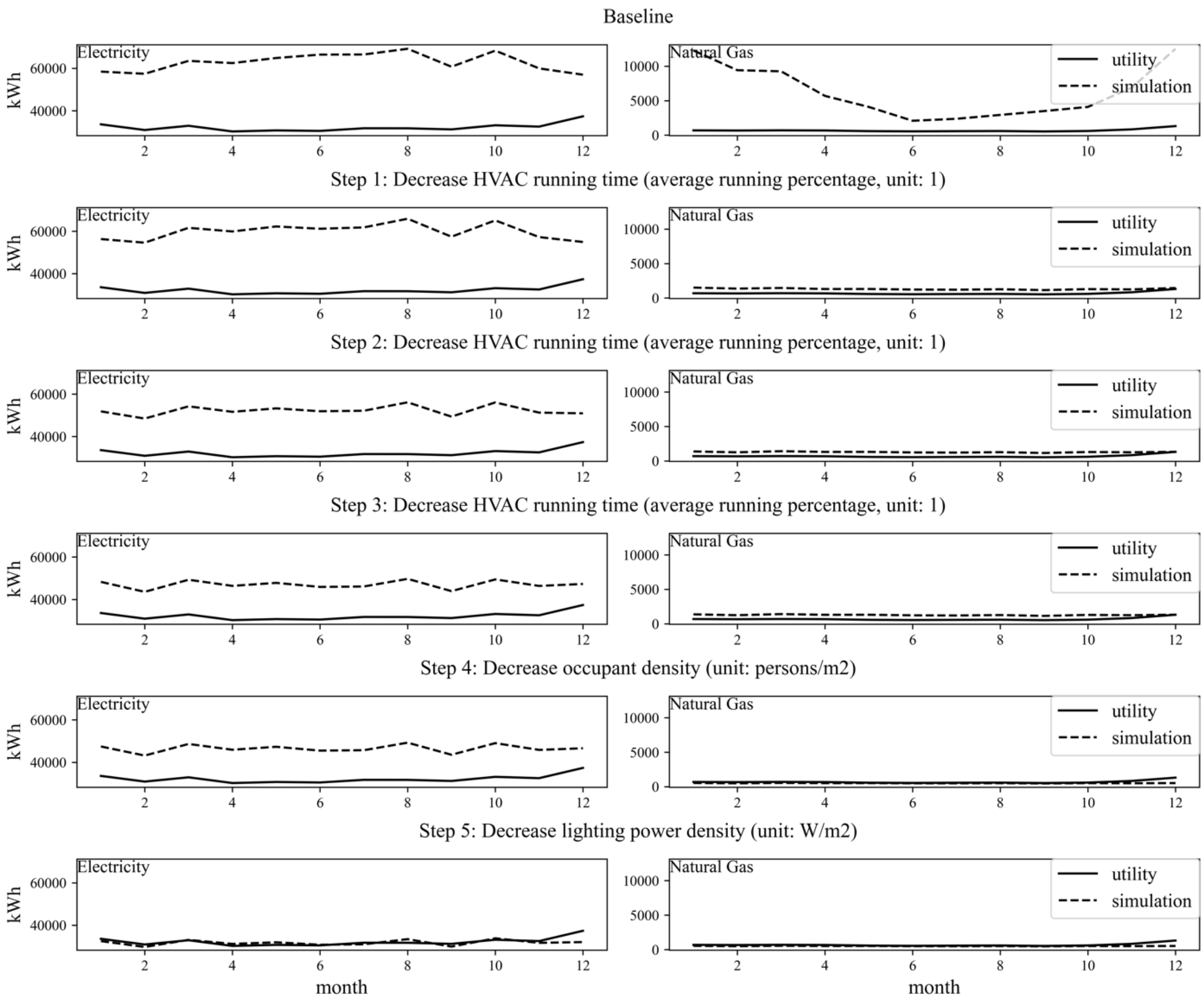


Figure 13. Detailed calibration process and results at each step for Building L. There are three separate steps of "decrease HVAC running time" due to limited marching size.

Taking this a step further, automated calibration was rerun on all 111 buildings with the modified tuning ranges. Overall, 57 buildings were successfully calibrated, 35 buildings failed to converge, and 19 buildings' utility bills were beyond calibration capability. The calibration success rate was improved from 42.3% to 51.4%. Therefore, for multi-tenant buildings, it is critical to take into consideration the possibility of partial occupancy rate in the model calibration.

Figure 14 illustrates the calibration metrics results, i.e., NMBE and CVRMSE, of the total source energy use, of the successfully calibrated 57 buildings and 35 buildings that failed to converge. The results of the last tried step are shown for the unsuccessful buildings. The 19 buildings that didn't pass prescreening were not included in the figure because they skip the calibration process. For the successful buildings, the NMBE are all within ±5% and CVRMSE are all less than 15%. As their NMBE results are scattered evenly from -5% to 5%, the method doesn't lead to systematic errors.

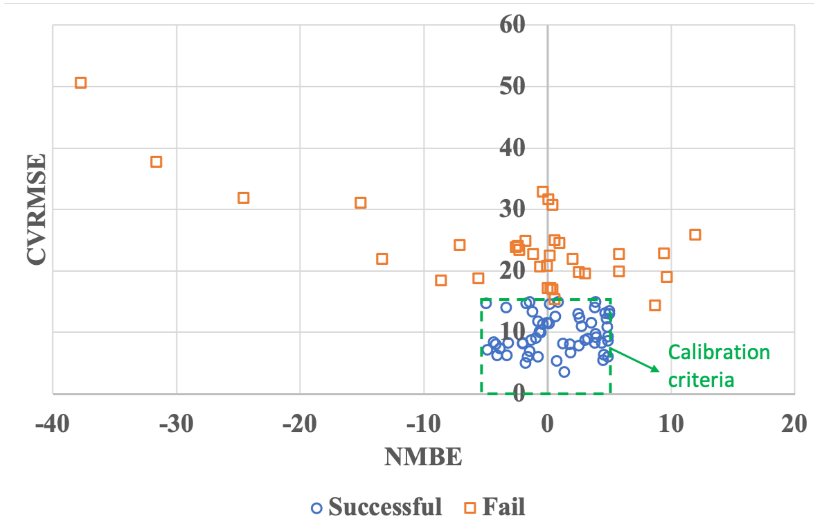


Figure 14. NMBE and CVRMSE results of the successfully calibrated 57 buildings and the failed 35 buildings.

4. Discussion

As the baseline models are generated using only four pieces of high-level building information (i.e., year built, total floor area, use type, and location), many model inputs are inferred by the corresponding minimum requirements based on applicable building energy codes and standards—in this case, California's Title 24 standards. As almost every existing building has gone through some changes (retro-commissioning, retrofit, or other operational changes), the latest information from an energy audit, building permit, or operator's logbook will help refine the baseline models before the calibration process. This not only saves calibration time but also increases the number of models that can be calibrated. It should be noted that efforts in getting the facts of buildings and their energy systems and operations are crucial in creating a reasonable building energy model. Calibration is an engineering technique that cannot replace the dedicated effort of getting factual inputs and modelers' domain expertise.

As discussed in Section 3.2, the more information we have, the more likely we are able to calibrate the model with a higher accuracy. It is simple and straightforward to search information for a couple of buildings manually, but it becomes labor intensive for a larger group of buildings. A systematic method has been developed to integrate a city's public records into data models that can be adopted in urban scale modeling (Chen et al. 2017; Chen et al. 2018). It becomes much easier to obtain geographical information (such as adjacent walls, building orientation, aspect ratio, and shading between adjacent buildings) on the urban scale than from a single building perspective. Such information can then be used as simulation inputs to improve the simulation accuracy, as well as the calibration success rate.

The 111 buildings are a subset of a larger dataset. They were selected for this study partly due to the limitations of the current pattern-based model calibration, which only applies to buildings served by DX cooling and natural gas heating systems, as well as natural gas powered SHW systems. Central chiller systems or electric (resistance type or heat pump) heating systems are not supported yet. More HVAC system types will be included in the algorithm to expand its applicability in future. We found it not unusual that some buildings are only partially occupied, which is consistent with previous research (Ding et al. 2022; Lu et al. 2021) and also in line with the emerging teleworking mode due to pandemic (Kharvari, Azimi, and O'Brien 2022). This is another important factor to be considered in model calibration. Besides, occupant behaviors impact the energy efficiencies in buildings significantly (Tang, Wang, and Sun 2021; Zhou et al. 2022) and should also be considered in model calibration in future research.

The pattern-based automated calibration method provides detailed step-by-step tuning of key model parameters and the progress of calibration using the ASHRAE Guideline 14 metrics, enabling modelers to double check the calibration process to ensure its validity. This calibration method requires a complete year of 12 monthly electricity and natural gas consumption data. As utility bill data are practically available for every building with its own meters due to the utility billing purpose (getting access to the data is a separate issue), the calibration method can be potentially adopted widely.

A future work of interest is to test the calibrated models using another year's monthly energy use data and weather data if available. The calibrated models should have reasonable accuracy in predicting the monthly energy use if the building characteristics and operations remain the same or barely changed.

Other calibration methods should be considered if the available building energy use data are different, e.g., smart meter data providing 15-minute electricity use for individual buildings, and city's public building benchmarking ordinance providing annual energy use for groups of buildings subjected to the ordinance (usually over certain threshold of building size / total conditioned floor area). Sometimes building energy system operational data, e.g., indoor air temperature and humidity, CO_2 concentration, supply air flow rate and supply air temperature, are available, which can be used in other calibration methods. High resolution building energy use data and system operation data may be needed to calibrate models that target use cases, e.g., fault detection and diagnostics, demand response, and model predictive controls.

5. Conclusion

Automated energy model calibration using widely available monthly utility bills for buildings is a crucial step for ensuring the accuracy of the simulated synthetic smart meter data for a group of buildings. This paper demonstrates the pattern-based method can be used to calibrate the majority of building models that are within the method's capabilities (i.e., certain building types and HVAC system types) and have normal utility bill data. Improvements to the pattern-based calibration method for future consideration include: (1) extension to cover large-sized buildings that use central chiller plant, (2) use of electricity (either electrical resistance or electric heat pump) for space heating and/or service water heating, and (3) a method to pre-screen validity of utility bill data to detect anomalies.

Future work will package the pattern-based calibration in an OpenStudio measure (Goldwasser et al. 2016) for public use. With more than 25 U.S. cities and states implementing public building energy benchmarking and disclosure programs, much more energy data (at the granularity of annual electricity and annual natural gas usage) for individual buildings has become available. A new model calibration method will be developed and tested for a large group of buildings using the annual energy data. Effective calibration methods based on real smart meter data is another future research topic.

Acknowledgments

This research was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Office of Building Technologies of the United States Department of Energy, under Contract No. DE-AC02-05CH11231.

References

Andrade-Cabrera, C, W J N Turner, and D P Finn. 2019. “Augmented Ensemble Calibration of Lumped-Parameter Building Models.” *Building Simulation* 12 (2): 207–230. doi:10.1007/s12273-018-0473-5.

ANSI/ASHRAE. 2014. “ASHRAE Guideline 14-2014 Measurement of Energy, Demand, and Water Savings.” *American Society of Heating, Refrigerating and Air-Conditioning Engineers*.

Asadi, Somayeh, Ehsan Mostavi, Djamel Boussaa, and Madhavi Indaganti. 2019. “Building Energy Model Calibration Using Automated Optimization-Based Algorithm.” *Energy and Buildings* 198 (September). Elsevier Ltd: 106–114. doi:10.1016/j.enbuild.2019.06.001.

ASHRAE. 2017. “Chapter 19: Energy Estimating and Modeling Methods.” In *2017 ASHRAE Handbook—Fundamentals*. ASHRAE.

Balaras, Constantinos A., Elena G. Dascalaki, Kalliopi G. Droutsa, and Simon Kontoyiannidis. 2016. “Empirical Assessment of Calculated and Actual Heating Energy Use in Hellenic Residential Buildings.” *Applied Energy* 164. Elsevier Ltd: 115–132. doi:10.1016/j.apenergy.2015.11.027.

California Energy Commission. 2013. *2013 Building Energy Efficiency Standards for Residential and Nonresidential Buildings*. https://ww2.energy.ca.gov/title24/2013standards/.

California Public Utilities Commission. 2020. “The Database for Energy Efficient Resources (DEER).” http://www.deeresources.com/.

Chen, Yixing, Zhang Deng, and Tianzhen Hong. 2020. “Automatic and Rapid Calibration of Urban Building Energy Models by Learning from Energy Performance Database.” *Applied Energy* 277 (March). Elsevier Ltd: 115584. doi:10.1016/j.apenergy.2020.115584.

Chong, Adrian, and Song Chao. 2020. “A Framework for the Continuous Calibration Of Building Energy Models With Uncertainty.” In *Proceedings of Building Simulation 2019: 16th Conference of IBPSA*, 16:4562–4569. doi:10.26868/25222708.2019.210577.

Coakley, Daniel, Paul Raftery, and Marcus Keane. 2014. “A Review of Methods to Match Building Energy Simulation Models to Measured Data.” *Renewable and Sustainable Energy Reviews* 37. Elsevier: 123–141. doi:10.1016/j.rser.2014.05.007.

Cuerda, Elena, Olivia Guerra-Santin, Juan José Sendra, and Fco. Javier Neila. 2020. “Understanding the Performance Gap in Energy Retrofitting: Measured Input Data for Adjusting Building Simulation Models.” *Energy and Buildings* 209 (February): 109688. doi:10.1016/j.enbuild.2019.109688.

Department of Energy. 2019. “BayREN Integrated Commercial Retrofits (BRICR).” https://www.energy.gov/eere/buildings/downloads/bayren-integrated-commercial-retrofits-bricr.

Deru, M., K. Field, D. Studer, K. Benne, B. Griffith, P. Torcellini, B. Liu, et al. 2011. *U.S. Department of Energy Commercial Reference Building Models of the National Building Stock*. NREL/TP-5500-46861, 1009264. doi:10.2172/1009264.

Ding, Y, S X Han, Z Tian, J Yao, W Y Chen, and Q Zhang. 2022. “Review on Occupancy Detection and Prediction in Building Simulation.” *Building Simulation* 15 (3): 333–356. doi:10.1007/s12273-021-0813-8.

Garrett, Aaron, and Joshua New. 2015. “Scalable Tuning of Building Models to Hourly Data.” *Energy* 84. Elsevier Ltd: 493–502. doi:10.1016/j.energy.2015.03.014.

Heo, Y., R. Choudhary, and G.A. Augenbroe. 2012. “Calibration of Building Energy Models for Retrofit Analysis under Uncertainty.” *Energy and Buildings* 47. Elsevier B.V.: 550–560. doi:10.1016/j.enbuild.2011.12.029.

Hong, Taehoon, Jimin Kim, Jaemin Jeong, Myeonghwi Lee, and Changyoon Ji. 2017. “Automatic Calibration Model of a Building Energy Simulation Using Optimization Algorithm.” *Energy Procedia* 105. The Author(s): 3698–3704. doi:10.1016/j.egypro.2017.03.855.

Hong, Tianzhen, Jared Langevin, and Kaiyu Sun. 2018. “Building Simulation: Ten Challenges.” *Building Simulation* 11 (5): 871–898. doi:10.1007/s12273-018-0444-x.

Hong, Tianzhen, Mary Ann Piette, Yixing Chen, Sang Hoon Lee, Sarah C. Taylor-Lange, Rongpeng Zhang, Kaiyu Sun, and Phillip Price. 2015. “Commercial Building Energy Saver: An Energy Retrofit Analysis Toolkit.” *Applied Energy* 159 (December): 298–309. doi:10.1016/j.apenergy.2015.09.002.

Johnson, Nicolas. 2017. “Building Energy Model Calibration for Retrofit Decision Making by Nicolas R . Johnson A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Mechanical Engineering Thesis Committee : Sung Yi , Chair Elliott Ga.”

Kharvari, F, S Azimi, and W O’Brien. 2022. “A Comprehensive Simulation-Based Assessment of Office Building Performance Adaptability to Teleworking Scenarios in Different Canadian Climate Zones.” *Building Simulation*. doi:10.1007/s12273-021-0864-x.

La Fleur, Lina, Bahram Moshfegh, and Patrik Rohdin. 2017. “Measured and Predicted Energy Use and Indoor Climate before and after a Major Renovation of an Apartment Building in Sweden.” *Energy and Buildings* 146. Elsevier B.V.: 98–110. doi:10.1016/j.enbuild.2017.04.042.

Lam, Khee Poh, Jie Zhao, Erik B. Ydstie, Jason Wirick, Meiwei Qi, and Jihyun Park. 2014. “An Energyplus Whole Building Energy Model Calibration Method for Office Buildings Using Occupant Behavior Data Mining and Empirical Data.” In *2014 ASHRAE/IBPSA-USA Building Simulation Conference*, 160–167.

Lamagna, M, B Nastasi, D Groppi, M M Nezhad, and D A Garcia. 2020. "Hourly Energy Profile Determination Technique from Monthly Energy Bills." *Building Simulation* 13 (6): 1235–1248. doi:10.1007/s12273-020-0698-y.

Lu, X, F Feng, Z H Pang, T Yang, and Z O'Neill. 2021. "Extracting Typical Occupancy Schedules from Social Media (TOSSM) and Its Integration with Building Energy Modeling." *Building Simulation* 14 (1): 25–41. doi:10.1007/s12273-020-0637-y.

Martínez, Sandra, Pablo Eguía, Enrique Granada, Amin Moazami, and Mohamed Hamdy. 2020. "A Performance Comparison of Multi-Objective Optimization-Based Approaches for Calibrating White-Box Building Energy Models." *Energy and Buildings* 216. doi:10.1016/j.enbuild.2020.109942.

Raillon, L., and C. Ghiaus. 2018. "An Efficient Bayesian Experimental Calibration of Dynamic Thermal Models." *Energy* 152. Elsevier Ltd: 818–833. doi:10.1016/j.energy.2018.03.168.

Ramos Ruiz, Germán, Carlos Fernández Bandera, Tomás Gómez-Acebo Temes, and Ana Sánchez-Ostiz Gutierrez. 2016. "Genetic Algorithm for Building Envelope Calibration." *Applied Energy* 168: 691–705. doi:10.1016/j.apenergy.2016.01.075.

Reddy, T Agami. 2006. "Literature Review on Calibration of Building Energy Simulation Programs." *ASHRAE Transactions* 112 (1): 226–240. doi:Article.

Samuelson, Holly W., Arash Ghorayshi, and Christoph F. Reinhart. 2016. "Analysis of a Simplified Calibration Procedure for 18 Design-Phase Building Energy Models." *Journal of Building Performance Simulation* 9 (1): 17–29. doi:10.1080/19401493.2014.988752.

San Francisco Department of the Environment. 2020. "San Francisco's Existing Buildings Ordinance." https://sfenvironment.org/article/san-franciscos-existing-commercial-buildings-ordinance.

Sun, K., T. Hong, S.C. Taylor-Lange, and M.A. Piette. 2016. "A Pattern-Based Automated Approach to Building Energy Model Calibration." *Applied Energy* 165. doi:10.1016/j.apenergy.2015.12.026.

Tang, R, S W Wang, and S B Sun. 2021. "Impacts of Technology-Guided Occupant Behavior on Air-Conditioning System Control and Building Energy Use." *Building Simulation* 14 (1): 209–217. doi:10.1007/s12273-020-0605-6.

U.S. Department of Energy. 2018. *EnergyPlus V9.2 Documentation: Output Details and Examples*. *U.S. Department of Energy*. https://bigladdersoftware.com/epx/docs/9-2/output-details-and-examples/index.html.

Wang, D, X F Pang, W Wang, Z W Qi, Y Ji, and R X Yin. 2020. "Evaluation of the Dynamic Energy Performance Gap of Green Buildings: Case Studies in China." *Building Simulation* 13 (6): 1191–1204. doi:10.1007/s12273-020-0653-y.

Yin, Rongxin, Sila Kiliccote, and Mary Ann Piette. 2016. "Linking Measurements and Models in Commercial Buildings: A Case Study for Model Calibration and Demand Response Strategy Evaluation." *Energy and Buildings* 124 (July). Elsevier Ltd: 222–235. doi:10.1016/j.enbuild.2015.10.042.

Yoshino, Hiroshi, Tianzhen Hong, and Natasa Nord. 2017. "IEA EBC Annex 53: Total Energy Use in Buildings—Analysis and Evaluation Methods." *Energy and Buildings* 152 (March 2013). Elsevier B.V.: 124–136. doi:10.1016/j.enbuild.2017.07.038.

Zhou, Xin, Shuai Tian, Jingjing An, Da Yan, Lun Zhang, and Junyan Yang. 2022. "Modeling Occupant Behavior's Influence on the Energy Efficiency of Solar Domestic Hot Water Systems." *Applied Energy* 309 (March): 118503. doi:10.1016/j.apenergy.2021.118503.

Zibin, Nicholas, Radu Zmeureanu, and James Love. 2016. “Automatic Assisted Calibration Tool for Coupling Building Automation System Trend Data with Commissioning.” *Automation in Construction* 61. Elsevier B.V.: 124–133. doi:10.1016/j.autcon.2015.10.004.