Applied Energy 286 (2021) 116492

Contents lists available at ScienceDirect

Applied Energy

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

Occupancy data at different spatial resolutions: Building energy performance and model calibration

Adrian Chong [a,*], Godfried Augenbroe [b], Da Yan [c]

[a] Department of Building, School of Design and Environment, National University of Singapore, 4 Architecture Drive, Singapore 117566, Singapore
[b] School of Architecture, Georgia Institute of Technology, Atlanta, GA 30332-0155, USA
[c] Department of Building Science, School of Architecture, Tsinghua University, Beijing 100084, China

ARTICLE INFO

Keywords:
Occupant modeling
Building performance simulation
Building simulation
Bayesian calibration
Uncertainty analysis

ABSTRACT

Occupancy is a significant area of interest within the field of building performance simulation. Through Bayesian calibration, the present study investigates the impact of the availability of different spatial resolution of occupancy data on the gap between predicted and measured energy use in buildings. The study also examines the effect of occupancy data on the quality of the constructed prediction intervals (PIs) using the Coverage Width-based Criterion (CWC) metric. CWC evaluates the PIs based on both their coverage (correctness) and width (informativeness). This investigation takes the form of an actual building case study, with nine months of hourly measured building electricity use, WiFi connection counts as a proxy for occupancy, and actual weather data. In general, the building energy model's accuracy improves with the occupancy and plug-loads schedule derived from WiFi data. Specifically, the Coefficient of Variation Root Mean Square Error (CV[RMSE]) reduced from 37% to 24% with an exponential improvement in the PIs quality compared to the results obtained with ASHRAE 90.1 reference schedules. However, the increase in prediction accuracy shrank to 5% CV(RMSE) and a comparable CWC upon calibrating the base loads of the reference schedules. Increasing the spatial resolution from building aggregated to floor aggregated occupancy data worsened the CV(RMSE) and CWC, suggesting trade-offs between parameter uncertainty and model bias/inadequacy. These results contribute to our understanding of the interactions between model complexity, simulation objectives, and data informativeness, facilitating future discussions on the right level of abstraction when modeling occupancy.

1. Introduction

Occupants have been identified as one of the six driving factors of energy use in buildings [1] and as a source of uncertainty with a significant impact on building performance simulation (BPS) [2]. Building occupants not only act as a source of internal heat gains, but their interaction with different building systems also exerts influence on a building's heating and cooling energy (HVAC system operations and efficiency), as well as its lighting and equipment loads [3,4]. In an evaluation of 121 LEED certified buildings, Turner and Frankel [5] showed that significant variability exists between the design-phase model predictions and measured energy performance. Occupancy was hypothesized as a possible reason behind the discrepancies, citing differences in operational practices and schedules that were not anticipated during the energy modeling process. The perception that occupants causes a significant source of discrepancy between building energy simulation predictions and actual observations is further corroborated in an international survey of 274 BPS users across 37 countries [6].

1.1. Occupants and building performance simulation (BPS)

The terms occupant behavior (OB) and occupancy are used conversely in the existing literature [7]. Although there are overlaps in the methods used in their predictions, modeling behavior is a lot more complicated than occupancy (status of occupant presence or absence). In BPS literature, OB is often regarded as the quantitative description of the direct and indirect influences that building occupants exert on building energy performance. Examples include the impact that occupants have on energy use resulting from lighting and appliance usage, window opening, HVAC (Heating, Ventilation, and Air-Conditioning) usage and controls, etc [8,9]. Behavior referring to observable stimulus–response or how occupants make decisions and interact with several factors that can be separated into various biological, psychological, and social contexts [10], although important, is less common in BPS and not addressed within the current paper.

BPS is often used to quantify the impact that occupant presence and behavior might have on simulation predictions, energy-saving

* Corresponding author.
E-mail address: adrian.chong@nus.edu.sg (A. Chong).

https://doi.org/10.1016/j.apenergy.2021.116492
Received 27 July 2020; Received in revised form 18 December 2020; Accepted 11 January 2021
Available online 21 January 2021

potentials of different energy conservation measures (ECMs), and performance of various building systems. For instance, Clevenger and Haymaker [11] quantified the uncertainties in BPS due to OB and found that the predicted energy consumption could differ by more than 150% when occupant-related inputs were all maximized or minimized. Sun and Hong [12] defined three different occupant behavior styles (austerity, normal, and wasteful) and showed that OB has a significant impact on ECMs such as zonal HVAC control and mix-mode ventilation because of their strong occupant interactions. Azar and Menassa [13] performed a comprehensive sensitivity analysis on OB related parameters. Their simulation results indicate that on average, heating setpoint temperature was the most sensitive parameter in small-size buildings located in climate zone 2 and that larger buildings were more sensitive to variations in lighting and equipment usage. Using a residential prototype building in EnergyPlus, O'Neil and Niu [14] showed that a 3.76% average coefficient of variation in OB related input parameters led to about 4% impact on HVAC annual and peak energy consumption.

To bridge these performance gaps, there has been increasing research concerning the modeling and simulation of occupant behavior in buildings [2,15]. Occupant presence and behavior in BPS are commonly represented with diversity factors in the form of standardized hourly schedules or profiles. Default diversity factors are typically obtained from codes and standards such as the ASHRAE 90.1[1] [16] when actual schedules are unknown. A limitation of these standardized profiles is that they are not specific to the buildings being modeled and have been shown to differ substantially from actual diversity profiles [17–19]. To overcome the limitations of its generality, multiple efforts have been undertaken to derive more reliable profiles. For instance, Davis and Nutter [20] derived occupancy profiles for eight university buildings using data from various data sources (security cameras, doorway counting sensors, classroom scheduling data, and manual observations). D'Oca and Hong [21] applied a three-step data mining framework to occupancy data for 16 offices and identified four archetypal occupant patterns that can be transformed into typical working profiles of occupancy for use in BPS.

1.2. Modeling of occupants

In general, the modeling of occupancy and OB can be categorized as deterministic or stochastic. Deterministic models include schedules and deterministic rules. Schedules or diversity factors are often represented as fractions in the range $[0,1]$ and relates hourly variations in occupancy, lighting, and equipment loads to their respective maximum occupancy or peak power density. The rule-based approach prescribes a set of deterministic rules that can be used to predict OB. For example, Yu [22] applied genetic programming to learn rules to predict occupancy in a single person office based on motion sensor data. The advantage of deterministic models lies in its simplicity and that model outputs are fully determined by the parameter values and simulation conditions.

Stochastic models of occupancy and behavior are often proposed as an alternative to deterministic models to more realistically account for the stochastic nature of OB. Notably, stochastic models have been proposed to represent arrival and departure times [23], absence and presence duration [24,25], and overtime [26] more accurately. Also, considerable efforts have been directed towards developing more detailed occupant models to capture the complexity and wide variability (both spatially and temporally) in occupant behavior. Methods that have been proposed typically involves sampling from probability distributions with parameters estimated from measured data [24,26]; stochastic Markov processes that generate time-series of states based on the state attained at the previous time-step [25,27,28]; and agent based models that involve defining and modeling autonomous individuals (agents), their mutual relationship, as well as their interaction with respective building spaces and systems [29].

However, the replacement of deterministic models with more complex stochastic processes is not synonymous with better predictive accuracy. Using an out-of-sample test dataset Mahdavi and Tahmasebi [30] evaluated two existing probabilistic occupancy models and a non-probabilistic occupancy model. They found that the predictive accuracy of all three models was low in general and that the simpler non-probabilistic model performed better in view of short-term occupancy predictions. Using simulation, Gilani, O'Brien, and Gunay [31] concluded that deterministic models could reasonably represent the occupants' impact on the building's average annual energy use. However, the effect of individual occupants diminishes with building size. Additionally, increasing complexity is often associated with other issues that include onerous data requirements, an increasing likelihood of model overfitting [15,32], as well as the introduction of uncertainties that may lead to identifiability issues during the parameter estimation process [33,34].

1.3. Occupant sensing for BPS

With advancements in occupant sensing and data acquisition technology, occupant information is becoming increasingly available and more easily accessible. Sensing and data acquisition form an important aspect when modeling OB in many studies. Examples include using an extensive network of environmental sensors to model occupancy [28], power consumption data of household appliances to characterize OB [35], and WiFi technologies to detect occupancy patterns [36]. Not surprisingly, models that are based on measured data also performs the best [15].

According to Melfi et al. [37], occupancy sensing can be categorized as either explicit or implicit. Implicit occupancy sensing involves the use of existing building infrastructure to measure occupancy because the interactions between occupants and the building systems can be used to determine occupancy information. On the contrary, explicit occupancy sensing involves the addition of specific occupancy sensing and detection capabilities and is often synonymous with high installation and maintenance costs [38]. The different occupancy monitoring and detection technologies include surveys, motion sensors (e.g., passive infrared (PIR) and ultrasonic detectors), vision-based sensors (e.g., video cameras), and radio frequency (RF) based sensors (e.g., ultra-wideband (UWB), radio-frequency identification (RFID), Wireless Local Area Network (WLAN) or WiFi) [39,40].

The availability of occupant information provides an opportunity to quantify its contribution towards bridging the performance gap more robustly. Specifically, the measurements provide a means to empirically test the impact that various spatial resolutions of occupancy information have in reducing the discrepancies between simulation predictions and real measurements. For instance, using post-occupancy surveys to quantify OB, Yu, Du, and Pan [41] showed that the prediction accuracy of the energy model could be improved by integrating the survey results into the simulation model. Sangogboye [42] showed that using higher resolution occupancy data does not improve the accuracy of the energy model that was previously calibrated on lower resolution data. Using data from 16 single occupied private offices, O'Brien et al. [43] concluded that modeling occupancy from aggregated data tends to suppress inter-occupant diversities. The authors argue that although the high-level metrics showed that there is no strong evidence suggesting the suppression of inter-occupant diversities, occupancy profiles generated from aggregated data do not contain the same level of diversities observed in individual occupant profile.

[1] http://sspc901.ashraepcs.org/documents/Addendum_an_Sched_and_Load.pdf.

1.4. Aims and objectives

The importance and originality of this study are that it explores the effect of including different resolutions of occupant information on model predictive performance. Although the characterization of OB's impact on BPS is not new, a forward or classical modeling approach is typically employed, with most studies focusing on the modeling methodology of occupants and its resulting impact on BPS. It is also not evident from existing literature if standardized profiles can be tuned to provide models that are good enough for predicting building energy usage. If standardized profiles are lacking, at what spatial resolution should occupant information be included in BPS to reduce the gap between simulated and measured energy use in buildings given an adequate calibration procedure? Monitoring occupancy at high spatial resolution (i.e., for a large number of zones) is an expensive endeavor, and there is a need to balance practicality with accuracy.

In this study, we employ an inverse empirical approach to quantify the effect of including occupant information on BPS predictive performance. Bayesian calibration is used for uncertainty quantification and to construct prediction intervals (PIs) given the stochastic nature of OB and because BPS predictions are influenced by many known and unknown factors. The purpose of modeling and quantifying uncertainties is to support any claims about the simulation predictions' accuracy robustly. Therefore, the objectives of this paper are:

- To evaluate the impact of including different spatial resolutions of occupancy information on model predictive performance.
- To elucidate the impact of base loads on predictive performance. Specifically, we would like to evaluate the improvements in predictive performance if the base loads of standardized profiles were calibrated.
- To propose a quantitative measure for the evaluation of BPS simulation PIs against measured data. Most studies in BPS literature evaluate PIs using their mean value, which diminishes the purpose of constructing PIs in the first place [44].

2. Method

A combination of data acquisition, modeling, and calibration will be used to quantify the effect of occupancy data on matching building energy simulation models to measured data. As illustrated in Fig. 1, the framework used for the evaluation of different spatial resolutions of occupancy data on the efficacy of the calibration can be summarized as:

1. Acquire occupant count information, total building energy consumption data, and the corresponding weather data.
2. Create occupant information at different spatial resolution.
3. Calibrate building energy model with different spatial resolution of occupant information.
4. Evaluate and compare "calibrated" model's performance.

2.1. BayesIan calibration and meta-modeling

The impact that the availability of different spatial resolutions of occupant presence information would have on the predictive performance of the building energy model is evaluated within a Bayesian calibration framework. The motivation behind a Bayesian approach is to construct prediction intervals and quantify the uncertainties associated with the predictions of the realizations. The pioneering work by Kennedy and O'Hagan [45] laid the foundation for Bayesian calibration that explicitly models uncertainty in model inputs, the discrepancy between the simulator and the actual physical system, and observation errors (Eq. (1)).

$$y(x) = \eta(x,t) + \delta(x) + \epsilon(x) \tag{1}$$

where, $y(x)$ is the observed field measurement, $\eta(x,t)$ is the output of the building energy simulation given observable inputs x and unknown calibration parameters t, $\delta(x)$ is the model bias or inadequacy or discrepancy, and $\epsilon(x)$ (assumed to be i.i.d and normally distributed) is used to model observation errors. The purpose of including the discrepancy term is to find the posterior distribution representative of the true but unknown values of the calibration parameters t. By incorporating model inadequacy δ, we believe that the calibration quality would improve since the model would better represent the actual physical system.

Since the iterative calibration process can be computationally intensive, a Gaussian process (GP) emulator or metamodel is used to emulate the energy model and combine measured data with simulation data [46].

We use a GP metamodel because it has been shown to provide the highest accuracy compared to other commonly used statistical models despite its longer run-time [47]. Maximin Latin hypercube sampling [48] was used to determine the different values of calibration parameters that would be used for the construction of the GP metamodel. Hamiltonian Monte Carlo (HMC), a Markov chain Monte Carlo (MCMC) method, is used to sample from the posterior probability distributions for more efficient sampling and better convergence [49,50]. Gelman Rubin statistics ($\hat{R}$ within 1 ± 0.1) and trace plots of multiple MCMC chains were used to ensure adequate convergence to a common stationary distribution. The source code, along with details on the implementation of Bayesian calibration utilizing a GP metamodel, can be found in Chong and Menberg [33].

2.2. Calibrating load profiles/schedules

Internal loads refer to occupants, equipment (plug and process), and lighting loads. In this study, we quantify the effects on the model's predictive performance when the schedule base load fraction is modeled as a calibration parameter. Calibrating the schedule base load fraction increases the degrees of freedom for matching the simulation predictions to the measured data. Base load fraction here refers to the minimum schedule fraction. Even though an hourly schedule is used, we do not calibrate every hourly schedule value because calibrating too many parameters results in an ill-posed inverse problem. This is due to overparameterization that could lead to issues of identifiability [51]. Also, simpler models generalize better. More calibration parameters mean more complex models that tend to over-fit the measured data.

Fig. 2 shows an illustration of varying the base load fraction of the ASHRAE 90.1 [16] equipment schedule while keeping the schedule's shape. We do this by taking advantage of the `Energy Management System` feature in EnergyPlus to modify and scale the schedule value during simulation runtime dynamically. Algorithm 1 shows the pseudocode for setting and scaling the equipment schedule fraction based on occupant schedule fraction and the equipment base fraction. At every simulation timestep t, the equipment schedule value $EQUIP_t$ is set equal to the calibration parameter X (base load fraction) when the building is not in operation. During operating hours, $EQUIP_t$ is set to equal the occupant schedule fraction OCC_t and scaled within the range $[X, 1]$.

2.3. Performance evaluation

Fig. 3 shows the framework used to evaluate model performance. Firstly, the dataset is split into a training dataset that is used for the calibration and a testing dataset to validate the model's performance. After removing missing values, the dataset comprises of 6504 samples. We use 200 random samples (i.e., the training data) for the Bayesian calibration. The remaining 6304 samples were then used as a hold-out test dataset for evaluating calibration performance. Since a good match between simulation predictions and measured data on the training dataset (i.e., data used for the calibration) does not necessarily mean that a proper calibration has been carried out, the test dataset that

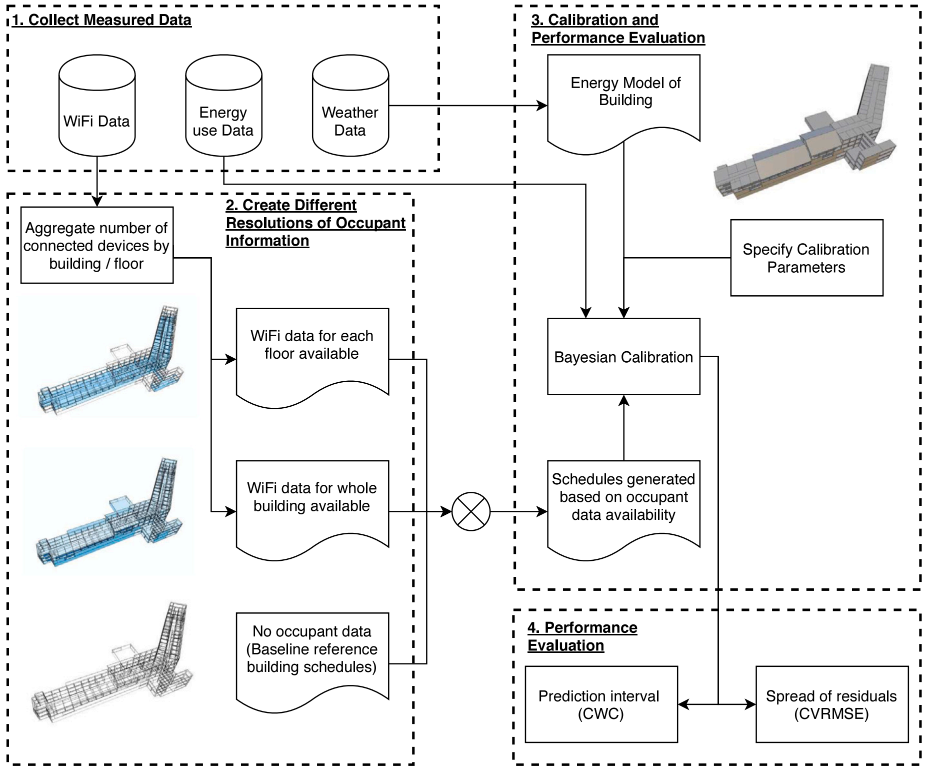


Fig. 1. Framework for using real data to evaluate the impact of occupant presence on building energy simulation. Steps include (1) data acquisition; (2) using WiFi data as a proxy for occupant presence, create different resolution of occupant information; (3) Apply Bayesian calibration; and (4) Evaluate and compare the efficacy of the calibration given different spatial resolution of occupant information.

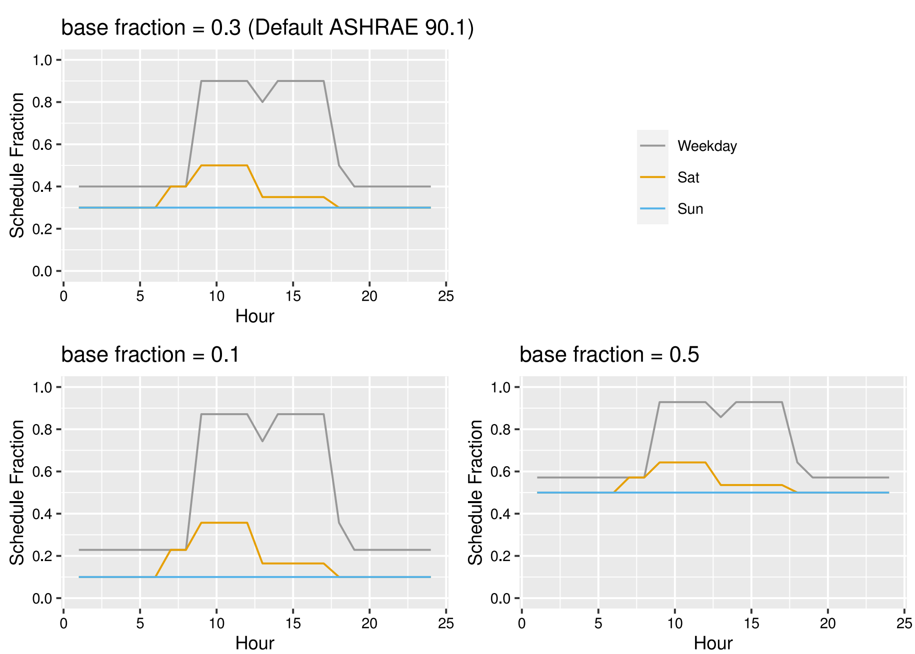


Fig. 2. Modifying the base fraction of the load profile/schedule. Top plot shows the default ASHRAE 90.1 equipment (plug and process) schedule with base fraction of 0.3 (minimum schedule value). Bottom plot shows the same schedule re-scaled with base fraction of 0.1 and 0.5 respective.

was not used during the calibration is used to validate the calibration performance. Section 3.2 and Table 3 summarizes the data, priors, inputs and output used for the Bayesian calibration.

Having completed the Bayesian calibration, the posterior estimates of the calibration parameters were extracted and used as inputs to the EnergyPlus model. We then run these EnergyPlus models and form probabilistic predictions using the simulation output. The probabilistic predictions are then evaluated by comparing them with the 6304 samples of testing data. The comparison is made using two evaluation metrics, the Coverage width-based criterion (CWC) and the

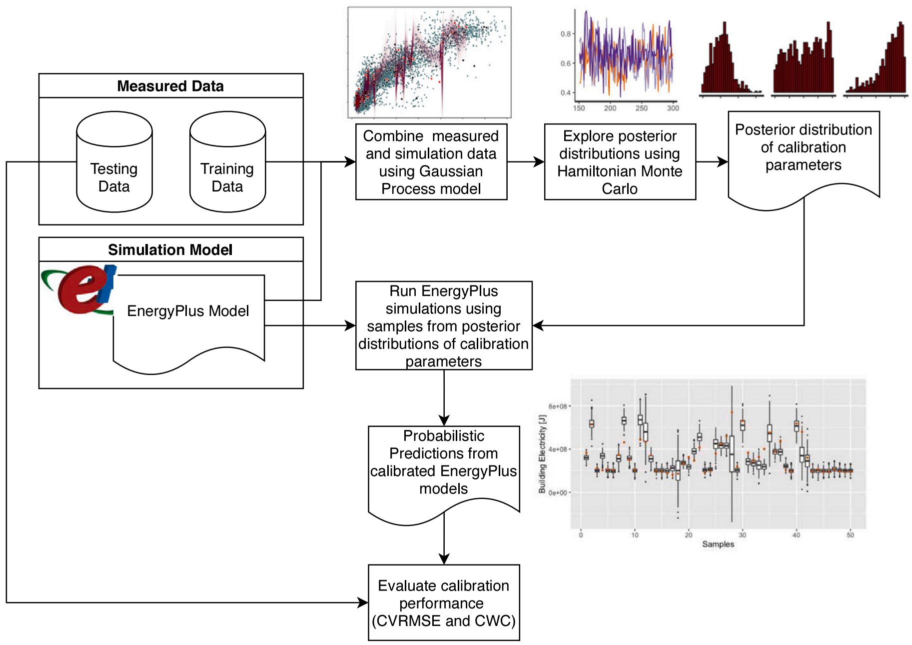


Fig. 3. Framework for evaluating model performance. The posterior distributions of the EnergyPlus parameters generated from the Bayesian calibration is used as inputs to the model. The outputs of the model is then evaluated.

Algorithm 1: Pseudocode for modifying and scaling the equipment schedule based on occupant information and equipment base fraction

X = equipment base fraction;
EQUIP = equipment schedule;
OCC = occupancy schedule;
$\text{OCC}_{\min}$ = minimum occupancy schedule fraction;
for *every simulation timestep t* **do**
  if *building is not in operation* **then**
    $\text{EQUIP}_t = \text{X}$;
  else
    $\text{EQUIP}_t = (\text{OCC}_t - \text{OCC}_{\min})(1 - \text{X})/(1 - \text{OCC}_{\min}) + \text{X}$;
  end
end

coefficient of Variation of the Root-Mean-Square Error (CV[RMSE]). We use CV[RMSE] to provide an indication of the model's prediction accuracy and CWC to evaluate the quality of the probabilistic predictions. Consequently, each metric provides a different measure of calibration effectiveness and is chosen to provide a robust evaluation of the uncertainties associated with the simulation predictions.

2.3.1. Prediction interval (PI)

With the presence and accumulation of different sources of uncertainties [52], probabilistic predictions have often been proposed in the literature to aid decision making [53]. To that end, PI construction is often used to predict the range (with a certain probability) a future observation is likely to fall. Therefore, PI construction provides the capability to quantify the uncertainties associated with the differences between the measured data and the simulation predictions.

Probabilistic predictions are often evaluated based on their coverage probability. An example of such a metric is the PI coverage probability (PICP), which measures the percentage of observations that falls within the PIs (Eq. (2)).

$$PICP = \frac{1}{n_{test}} \sum_{i=1}^{n_{test}} c_i^{(\alpha)} \tag{2}$$

where

$$C_i = \begin{cases} y_i = 1, & y_i \in [L_i, U_i] \\ y_i = 0, & y_i \notin [L_i, U_i] \end{cases} \tag{3}$$

n_{test} is the number of samples in the test dataset; y_i denotes the ith observed value; $L_i^{(\alpha)}$ and $U_i^{(\alpha)}$ are the lower and upper bounds of the ith PI with a confidence level of $(1-\alpha)\%$, respectively.

However, coverage probability by itself is not a good measure because a very high coverage probability can easily be achieved with PIs that have a broad range (i.e., a very small and large lower and upper bound respectively). As shown in Fig. 4a, PIs that are too wide are not useful in practice because they convey little information about the variations in the measurements. Therefore, evaluation of PIs based on coverage alone is subjective and can lead to misleading results [54]. Consequently, another measure that quantifies the width of the PIs is necessary. The Prediction Interval Normalized Mean Width (PINMW) (Eq. (4)) measures this property of the PI.

$$PINMW = \frac{1}{n_{test} \cdot R} \sum_{i=1}^{n_{test}} (U_i^{(\alpha)} - L_i^{(\alpha)})^2 \tag{4}$$

where R is the range of the measured values. Normalizing by R (Eq. (4)) allows us to compare the PIs generated from different datasets and calibration methods.

Although narrow PIs are more informative about measurement variability, the PIs are not reliable if they have poor coverage probability (Fig. 4b). Therefore, probabilistic predictions should have a high coverage probability and a narrow range to be useful (Fig. 4c). Consequently, performance metrics used to evaluate PIs should consider both coverage probability and PI width. An example of such a metric is the coverage width-based criterion (CWC). Using CWC provides an evaluation of the PIs from these two conflicting viewpoints of coverage probability and PI width.

2.3.2. Coverage width-based criterion (CWC)

In this study, we use the coverage width-based criterion (CWC) (Eq. (5)) [54] to evaluate the PIs. The CWC is a measure that assesses the quality of the PIs based on both their coverage and their width

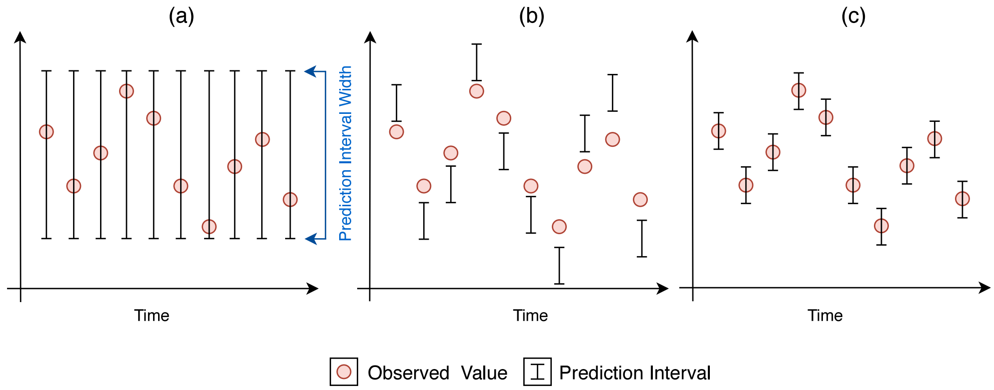


Fig. 4. (a) Prediction intervals with a high coverage probability but a wide range contain little or no information about the variations in the measurements; (b) Narrow PIs with poor coverage probability may be informative about measurement variability but are not reliable; (c) PI should have a high coverage probability and a narrow range to be useful.

through a combination of the PICP (Eq. (2)) and the PINMW (Eq. (4)) respectively.

$$CWC = PINMW + \gamma(PICP)e^{-\eta(PICP-\mu)} \tag{5}$$

where

$$\gamma(PICP) = \begin{cases} 0, & PICP \geq \mu \\ 1, & PICP < \mu \end{cases} \tag{6}$$

μ is the nominal confidence level and its value can be determined based on the confidence level $(1-\alpha)\%$ associated with the PIs. η is a scaling factor for distinguishing PICP violations from μ. Together, μ and η are hyperparameters used to control the location and the amount of CWC jump. In this study, η and μ are set to 50 and 0.65 respectively.

Intuitively, the CWC exponentially penalizes PIs if PICP is less than the nominal confidence interval μ regardless of the width of the PIs. If $PICP \geq \mu$ than CWC value is influenced by PINMW. As illustrated in Fig. 4c, ideally, we want PIs with small CWC values, which is indicative of PIs with narrow widths (small PINMW) and high coverage probability (large PICP).

2.3.3. CV(RMSE)

The coefficient of Variation (CV) of the Root-Mean-Square Error (RMSE) (Eq. (7)) is commonly used to assess how well the calibrated energy simulation describes the measured data, and is used in this study to provide an indication of how well the mean of the posterior predictions matches the measured data. Relative to monthly and hourly calibration data, ASHRAE Guideline 14 [55] specifies a threshold of 15% and 30% respectively. With probabilistic predictions, prediction performance is evaluated using the mean of the CV(RMSE) computed using the MCMC samples from the posterior predictions.

$$CV(RMSE) = 100 \times \frac{\sqrt{\sum_{i=1}^{n_{test}} (y_i - \hat{y}_i)^2 \,/\, (n_{test} - 1)}}{\bar{y}} \tag{7}$$

$$\hat{y}_i = \frac{1}{m}\sum_{j=1}^{m} \hat{y}_{i,j} \tag{8}$$

where n_{test} is the number of samples in the test dataset; $\hat{y}_i$ denotes the mean of the ith predictions (Eq. (8)); $\bar{y}$ denotes the mean of the n_{test} measured values that forms the test dataset; m is the number of MCMC samples in the posterior predictions of $\hat{y}_i$.

3. Case study

3.1. Building description

The building case study is an actual educational building located at the National University of Singapore in Singapore. It is a six-storey building with a total floor area of 5527 m^2. Table 1 shows the percentage floor area breakdown of the building based on space types, with faculty rooms and teaching spaces accounting for the majority of the occupied spaces. The HVAC system is a VAV system served by central district water-cooled chillers. Given Singapore's tropical climate, no heating is required, and cooling is provided throughout the year to maintain thermal comfort.

Table 1
Space type percentage breakdown.

Space type	Percentage of total
Faculty rooms	32%
Lecture halls/Seminar rooms	18%
Open offices	7%
Circulation	35%
Mechanical and electrical rooms	5%
Others	4%

Nine months of measured data at an hourly resolution from 1 Apr 2018 to 31 Dec 2018 was collected and consists of the following:

- Total building electricity energy consumption.
- WiFi connection counts for each floor of the building. WiFi data is collected from all WiFi Access Points (AP) located within the building.
- Actual Meteorological Year (AMY) weather data from the Singapore Changi airport weather station (WMO #486980).

3.2. Virtual experiments

Using data from the case study building, four virtual experiments A, B, C, and D are defined to evaluate the impact of occupant presence on the calibration efficacy of building energy simulation and are summarized in Table 2 with descriptions as follows.

Case A Default ASHRAE 90.1 occupancy, lighting and equipment load schedules.

Case B Default ASHRAE 90.1 occupancy, lighting and equipment load schedules. Equipment base load fraction is modeled as an uncertain parameter (Fig. 2) that is calibrated against hourly whole building energy usage data.

Case C Default ASHRAE lighting schedule. Whole building aggregated occupant information is used to derive both building aggregated occupancy and equipment load schedules. Equipment base load schedule fraction is modeled as an uncertain parameter that is calibrated against hourly whole building energy usage data.

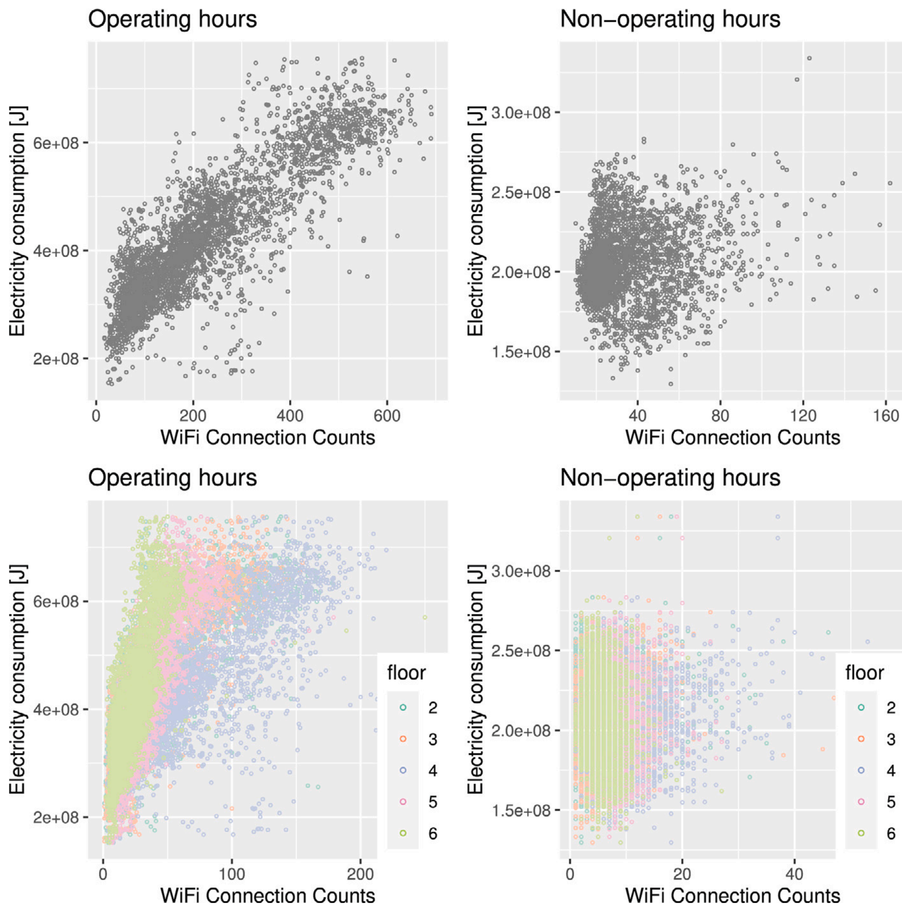


Fig. 5. Scatter plot of building aggregated (top plots) and floor aggregated (bottom plots) WiFi connection counts against building electricity energy consumption during operating and non-operating hours. Building aggregated count equals the sum of the floor aggregated counts.

Case D Default ASHRAE lighting schedule. Floor aggregated occupant information is used to derive the corresponding occupancy and equipment load schedules for each floor. Equipment base load schedule fraction for each floor aggregated schedule is modeled as an uncertain parameter that is calibrated against hourly whole building energy usage data.

Table 3 summarizes the measured data, calibration parameters, GP hyperparameters, and priors used for the Bayesian calibration. To ensure a fair comparison, each case was optimized by conducting a grid search of the Gaussian process hyperparameters on the training dataset and selecting the hyperparameters that gave the lowest CWC (Eq. (5)) values. The calibration is carried out against hourly building energy consumption data.

WiFi connection counts have been shown in previous studies to be a suitable estimate of the hour to hour variations in building occupancy [37] and equipment (plug and process) loads [56]. WiFi connection counts have also been shown to have strong positive correlation with building electricity consumption [37,57,58]. Therefore, WiFi data is used in this study as an implicit estimate of occupancy and equipment profiles or schedules. We reaffirm this by plotting scatterplots of electricity consumption against WiFi connection counts during operating and non-operating hours (Fig. 5) and computing their Pearson correlation. At the building level, a strong correlation (r = 0.86) was observed between electricity consumption and WiFi data. When the WiFi data was dis-aggregated by floors, a slight decrease in correlation was observed (r = 0.78, 0.84, 0.80, 0.82, 0.73 for floors 2, 3, 4, 5, and 6 respectively) but remains strong. By contrast, a weak correlation was observed during non-operating hours with both building aggregated (r = 0.17) and floor aggregated (r = 0.16, 0.26, 0.18, 0.11, 0.02 for floors 2, 3, 4, 5, and 6 respectively) WiFi data. The weak correlation during non-operating hours provides justifications for modeling the base loads as an uncertain parameter to be calibrated. To ensure that the diversity factors generated based on the WiFi connection counts are correctly applied, we sum the hourly total occupant count for the building and checked that they are the same across the different spatial resolutions.

4. Results

Table 4 summarizes the results of the four virtual experiments using the testing data. CWC and CV(RMSE) were computed to provide a quantitative assessment and comparison of the predictive performance at different resolution of occupant information with both synthetic and real data (see Table 2).

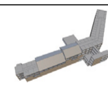
As illustrated in Fig. 6, the quality of the prediction intervals improves when occupant data is included in the calibration. This is indicated by lower CWC and CV(RMSE) values in cases C and D relative to A and B. Specifically, compared to default ASHRAE 90.1 schedules (case A), including occupancy data (cases C and D) shows a 11%–13% reduction in mean CV(RMSE) and an exponential improvement in CWC. The CWC measure (Eq. (5)) is based on PICP (prediction coverage) and PINMW (prediction width). Its hyperparameters (η and μ) were chosen so that PIs with $PICP \leq 65\%$ are heavily penalized. The considerably improvements in the CWC measure is mainly brought about by an increase in prediction coverage (PICP). From Table 4, it can be observed that including occupant information increases PICP by more than 30% (compared to using default ASHRAE 90.1 schedules).

A comparison of the simulation predictions with the measured data reveals that the calibrated model is unable to correctly capture the variability when the default ASHRAE 90.1 schedules is used. Fig. 8

Table 2
Summary of experiments used to investigate the effect that different spatial resolution of occupant presence has on the calibration efficacy of building energy simulation.

Case	Occupant data source	Occupant data resolution	Calibrates equipment base load
A	ASHRAE 90.1	None	No
B	ASHRAE 90.1	None	Yes
C	WiFi	Aggregated by building	Yes
D	WiFi	Aggregated by floor	Yes

Table 3
Data description, priors, inputs and output for Bayesian calibration using the real dataset (Building B).

	Description
Building	 Educational building located in Singapore
Data collection period	1 Apr 2018 to 31 Dec 2018
Weather	Singapore Changi Airport WMO #486980 (AMY)
No. of samples	training: 200 testing: 6304
Observed output $y(x)$	Total building energy consumption (excluding plant side cooling consumption)
Observed inputs x	Outdoor dry-bulb air temperature [°C] Outdoor relative humidity [%] Direct solar radiation [W/m^2] Occupant schedules [–] (see Table 2)
Calibration parameters t[a]	Occupant density [m^2/person] $\sim \mathcal{U}(15,23)$ Lighting power density (LPD) [W/m^2] $\sim \mathcal{U}(3,20)$ Equipment power density (EPD) [W/m^2] $\sim \mathcal{U}(3,30)$ Base load ratio (EPD) [–] $\sim \mathcal{U}(0.05,0.6)$
Gaussian process hyperparameters	$\lambda_\eta \sim Gamma\ (a=10, b=10)$ $\lambda_\delta \sim Gamma\ (a=10, b=0.1)$ $\rho^\eta, \rho^\delta \sim Beta\ (a=1, b=x)$ $\lambda_\epsilon \sim Gamma\ (a=10, b=y)$ $\operatorname{argmin}_{(x,y)} CWC$[b] $x \in (0.1, 0.3, 0.5)$ and $y \in (0.01, 0.05, 0.1)$

[a]For case D (total number of occupants per floor is approximated using WiFi connection count data), the power density and base load of each floor is modeled as a random variable that is calibrated. For a five storey building, this means a total of ten calibration parameters for EPD and each corresponding base load. Note that we model occupant density as a calibration parameter. This is because, although WiFi data acts as a good proxy for occupant hourly variation/schedule, it is not a good estimate for actual occupant count.
[b]Optimized by conducting a grid search on the training dataset and selecting the x and y values that gave the lowest CWC (Eq. (5)) values. A detailed explanation of the GP hyperparameters for Bayesian calibration can be found in Chong and Menberg [33].

Table 4
Performance of prediction intervals for the four case studies to 2 significant figures.

Metrics	Case			
	A	B	C	D
CV(RMSE)	37	29	24	26
CWC	9.7×10^6	1.3	0.16	0.20
PICP	0.33	0.65	0.67	0.71
PINMW	0.11	0.17	0.16	0.20

illustrates this observation when comparing case A simulation predictions with the measured data. There are two distinct observations, (1) the measurements tend to fall outside the range of Case A simulation predictions, and (2) the simulation predictions overestimates and underestimates the building's electricity consumption during operating (7am to 7pm on weekdays and 7am to 5pm on Saturdays) and non-operating hours respectively. Evidently, a higher degree of freedom when tuning the ASHRAE 90.1 schedules is necessary. We test this hypothesis with case B where the base load schedule fraction is modeled as a calibration parameter (see Section 2.2 for the methodology).

By calibrating the base load schedule fraction (Case A to Case B), CWC values were substantially reduced (Fig. 6) due to significant improvements in PI coverage of the measurements (approximately 30% increase in PICP as shown in Table 4). The increase in PICP is further reinforced by Fig. 8 that shows a better match between case B's simulation predictions with the measured data than case A's. What

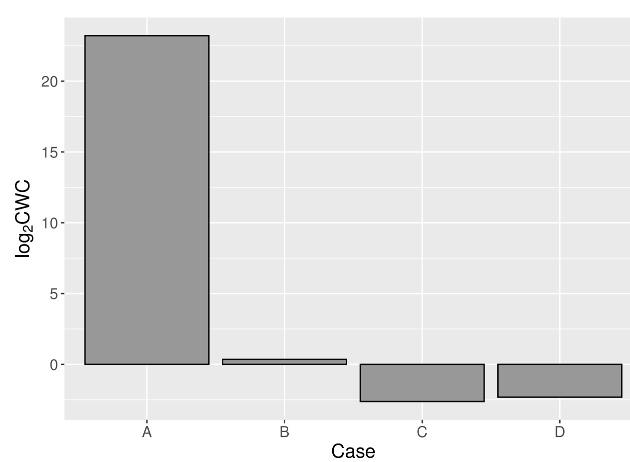


Fig. 6. Coverage Width-based Criterion (CWC) values for the four cases A–D (described in Table 2) for the test dataset. Plotted on the logarithmic scale to help visualize the differences between the four cases.

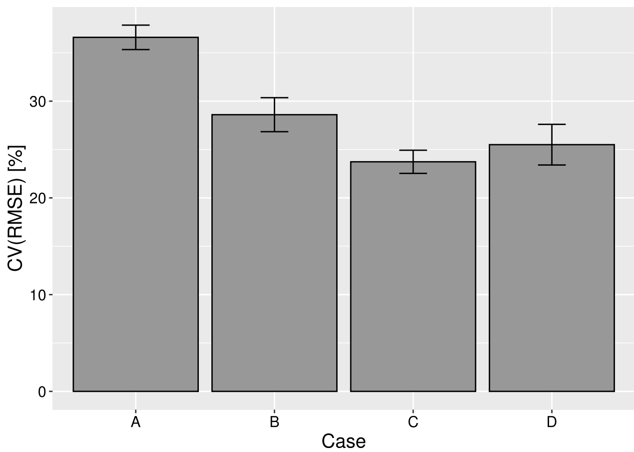


Fig. 7. Coefficient of Variation of the Root-Mean-Square Error (CV[RMSE]) values for the four cases A–D (described in Table 2) for the test dataset.

can be seen in this figure is the improved prediction coverage of the measured data, especially during non-operating hours (i.e., the base loads). Case B's mean CV(RMSE) is also 8% lower than case A (Fig. 7). By contrast, the discrepancies between cases B and C is not as large as that between cases A and B. Including building aggregated occupancy shows an additional 5% reduction in CV(RMSE) (Fig. 7) and slight improvement in CWC from 1.3 to 0.16 (Fig. 6). Nonetheless, Fig. 8 shows that including building aggregated occupancy (case C) leads to simulation predictions that are able to match the measured data better during operating hours. Interestingly, Fig. 8 shows that despite having a lower CWC value (Fig. 6), case C has comparable variance with cases A and B. Table 4 elucidates that the lower CWC is due to significant improvements in PI coverage (PICP) compared to case A. By contrast, the slight improvement in CWC between cases B and C can be attributed to a slight improvement in both PI coverage (PICP) and width (PINMW).

Providing floor aggregated occupancy (case D) does not improve the quality and accuracy of the simulation PIs. A distinct improvement in matching the measured data could not be observed between cases C and D (Fig. 8). Likewise, Figs. 6 and 7 show that compared to case C, case D has a slightly higher CV(RMSE) and CWC. A possible explanation for the worsened CV(RMSE) and CWC is that the diversity factors from occupancy data provided at the floor level do not add more information (as compared to building aggregated occupancy information) towards the prediction of total building energy consumption. Additionally, calibrating the power densities and base load schedule fraction of each floor may result in too many calibration parameters and a overly complex model. Surprisingly, CV(RMSE) of case D is between that of cases B and C. Following the principle of parsimony, case B that does not rely on any occupant data and has fewer lesser calibration parameters might be preferred in scenarios where a CV(RMSE) of 29% is deemed acceptable.

5. Discussion

5.1. Comparing ASHRAE schedule to using occupancy data

By substituting ASHRAE 90.1 reference schedules with building aggregated occupancy data, the deviation between predicted energy consumption and measured data were reduced by 13%. The improvements in prediction accuracy comes as no surprise since occupancy diversity factors was shown to differ by as much as 46% in average day profile peaks for private offices and about 12% reduction for open plan office spaces as compared to ASHRAE 90.1 reference schedules [17]. Variations in occupant related loads and schedules were also shown to have an significant impact on total building predicted energy usage [11]. Using an actual building, we showed that even with an adequate calibration methodology, CV(RMSE) stayed above the hourly threshold of 30% set by ASHRAE Guideline 14 [55] (Fig. 7). By including occupancy information, this deviation was reduced to 29% and within ASHRAE Guideline 14 [55] definition of a calibrated model.

Compared to the existing literature on occupant behavior, a noticeable difference in this study lies in applying Bayesian calibration to quantify the uncertainties in the simulation predictions. Additionally, the quality of the constructed prediction intervals (PIs) were evaluated using the coverage width-based criterion (CWC) that evaluates the PIs for its correctness (coverage of PIs) and its informativeness (width of PIs). As shown in Fig. 6, including occupancy information exponentially improves the quality of the constructed prediction intervals (PIs). Specifically, the PIs from a calibrated energy model where the schedules were based on occupancy data were more theoretically correct, with more than 30% of the measured data being covered by the constructed PIs. One would have also expected that including occupancy data would result in smaller PI widths representing a gain in knowledge about model uncertainty. However, a comparable PI width was observed across cases A, B, and C (Fig. 8). These observations suggest that including occupancy data improves the correctness of the constructed PI but does not have a significant impact on reducing prediction uncertainty.

Interestingly, with case A, it was shown that a CV(RMSE) of 37% can be achieved even though the constructed PIs only covered about 30% of the measured data. This observation provides evidence that despite calibrating for internal load (people, lighting, and equipment) densities, using fixed *a priori* reference schedules may result in simulation predictions in which the correctness of the constructed PIs is questionable.

5.2. Modifying ASHRAE schedule

Although we show that using ASHRAE reference schedules provide PIs with inadequate coverage of measured building energy usage, the question is whether these reference schedules can be adjusted to ensure sufficient accuracy and correctness. Accordingly, we show that increasing the calibration degree of freedom by modeling the schedule base load fraction as a calibration parameter reduces the calibration performance gap with comparable correctness in the constructed PIs than if occupancy data was included. However, the mean CV(RMSE) remains 5% larger. This suggests that if the peak and base loads can be reasonably estimated, the ASHRAE reference schedules might not be as detrimental as purported when used in BPS for predicting building energy usage.

To date, fixed *a priori* reference schedules represent the lowest level of complexity [15]. Schedules are also easy to include with implementation-ready interfaces in existing energy simulation software. Studies have shown that reference schedules may differ substantially from actual diversity profiles [17–19]. However, the question remains whether a reference schedule with its base and peak loads adjusted serves as a good enough abstraction of actual occupant presence and their interaction with the building systems. Specifically, would adding additional degrees of freedom to the reference schedule suffice to achieve the predictive performance needed to meet the simulation objectives, or is it necessary to introduce more complex probabilistic and agent-based occupancy models? In future work, investigating this research question might prove important.

5.3. Model complexity, inputs, outputs, and simulation objectives

Table 5 summarizes the spatial resolution investigated in this study as compared to BPS literature. Out of the studies listed in Table 5, only Sangogboye et al. [42] and Kim et al. [59] evaluated the impact of schedules derived from occupancy information on the energy model's

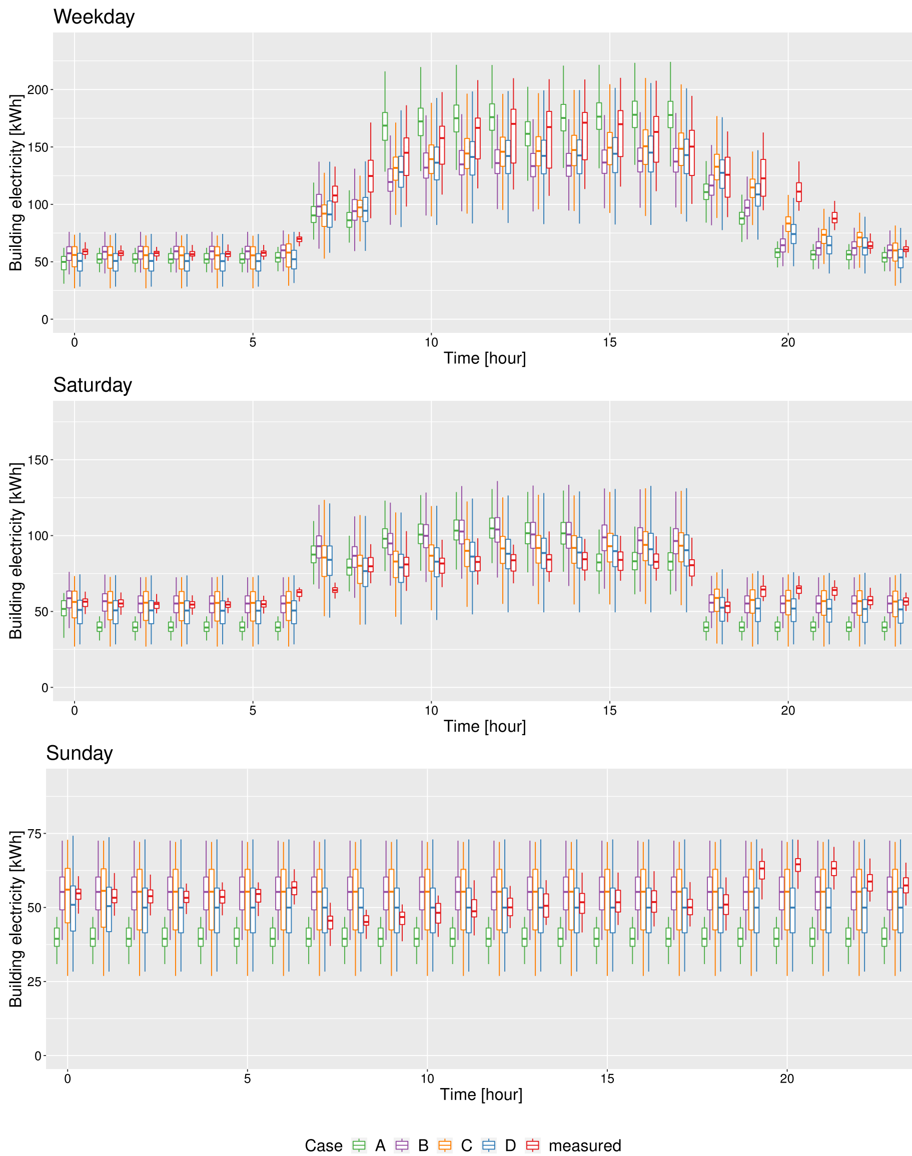


Fig. 8. Box-plots of hourly building energy consumption measurements compared with simulation predictions (for cases A, B, C, and D) separated by weekdays, Saturday, and Sunday. Descriptions of the cases can be found in Table 2.

Table 5
Summary of spatial resolution of measured data used as input to BPS investigated in this study as compared to BPS literature.

Spatial resolution				References
Building aggregated	Floor aggregated	Space type aggregated	Individual room	
✓	✓			Current study
✓			✓	[42]
✓				[20,19]
	✓			[60]
		✓		[17]
			✓	[18,28,21,30,59,24]

prediction accuracy. The remaining studies focused on the evaluation of occupancy prediction models against observational data.

The study by Sangogboye et al. [42] is the only other study that examined the effects of more than one spatial resolution on predictive performance. They reported similar findings of an increase in prediction error from 19.72% to 21.03% when switching from average building aggregated occupancy data to detailed individual room occupancy data. In this study, increasing the resolution from building aggregated occupancy data to floor aggregated occupancy data showed a 2% increase in prediction error. Additionally, we showed that increasing the spatial resolution increased prediction uncertainties. It is probable that using higher resolution occupancy data increases the number of model parameters and thus the model's complexity as compared to an abstracted version where similar parameters are lumped into a single building aggregated parameter. Depending on the outputs or responses that the simulation model is calibrated against, increasing the number of calibration parameters can result in issues of identifiability as a result of overparameterization [33,34].

The study by Kim et al. [59] showed a drastic improvement in prediction accuracy using schedules derived from sub-metered plug-load

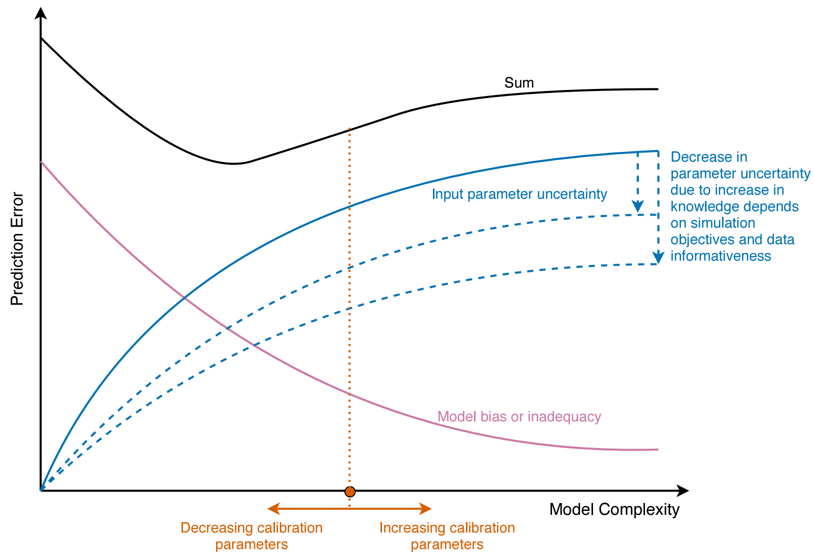


Fig. 9. Trade-offs between parameter uncertainty and model inadequacy with increasing model complexity, and the corresponding effects due to the simulation objectives, model inputs and outputs, and calibration parameters.
Source: Adapted from Trčka and Hensen [32].

data and individual room occupancy data as compared with default schedules. Through three case studies, they showed that CV(RMSE) could be reduced from 67% to 21%, 128% to 31%, and 156% to 16% respectively. Comparatively, we demonstrated in the current study a reduction in CV(RMSE) from 37% (default schedule) to 24% (building level occupancy data). The considerable improvements in prediction accuracy in Kim et al. were because their baseline models that use the default schedules had substantially larger errors. A probable explanation is that the baseline default schedule model has been Bayesian calibrated in our study but not in Kim et al.. Although not directly comparable, it is also interesting to note that the CV(RMSE) obtained in our study with building aggregated occupancy data is in a similar range to that observed in Kim et al. with individual room occupancy data.

Fig. 9 represents the trade-offs between model bias and parameter uncertainty, drawing from the results of this study; ongoing discussions to develop parsimonious building energy models [32,15,61]; and the relevant literature that delves into the effects on model performance with varying model complexity, data quality, and the simulation objectives [62–65]. The findings from this study suggest that although complex models provide a better representation of the actual physical building systems and sub-systems, they may not be useful if a comprehensive dataset is not available to constrain each pathway. This is because the model's performance depends on (1) model complexity, (2) simulation objectives, and (3) data informativeness. A more accurate representation of building systems and sub-systems requires more knowledge that could come from informative data streams, detailed drawings, and specifications because of the increasing number of model parameters. Additional complexity inadvertently leads to an increase in the number of uncertain parameters that need to be calibrated against an often limited amount of data streams. Therefore, it is crucial that modelers identify the appropriate level of complexity for a given dataset and simulation objective (question of interest). As suggested by [66], simulation is no longer the art of performing high fidelity simulations but rather performing the right type of virtual experiment with the right model. The effects of model complexity and structure, simulation objectives, and data informativeness on building performance simulation remain unanswered at present and is an important issue for future research.

6. Conclusion

The present study was designed to determine the effect of occupancy data on the building energy simulation model's predictive performance. To robustly support claims about improvements in predictive performance, Bayesian calibration is used to construct prediction intervals (PIs) and quantify any associated uncertainties. We also propose using the coverage width-based criterion to better evaluate the quality of the PIs.

The more obvious finding to emerge from this study is that compared to ASHRAE 90.1 reference schedules, including occupancy information (in the form of hourly profiles) improves the quality and accuracy of the PIs.

One of the more significant findings revealed in this study is that increasing the calibration degree of freedom by calibrating the schedule base load fraction reduces the gap between using ASHRAE reference schedules and occupancy information substantially. The finding suggests that with a reasonable estimation of the peak and base loads during design, the ASHRAE reference schedules might not be as detrimental as purported when the simulation objective is to predict total building energy usage.

The second significant finding was that a higher spatial resolution of occupancy data might result in poorer predictions with larger uncertainties of total building energy usage than whole building aggregated occupancy data. The finding supports the idea that additional model complexity leads to an increase in parameter uncertainty. If data that is informative about the simulation objective is not available to constrain

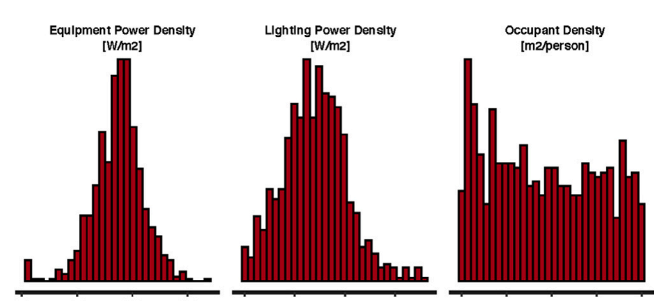


Fig. A.10. Posterior distribution of calibration parameters for Case A.

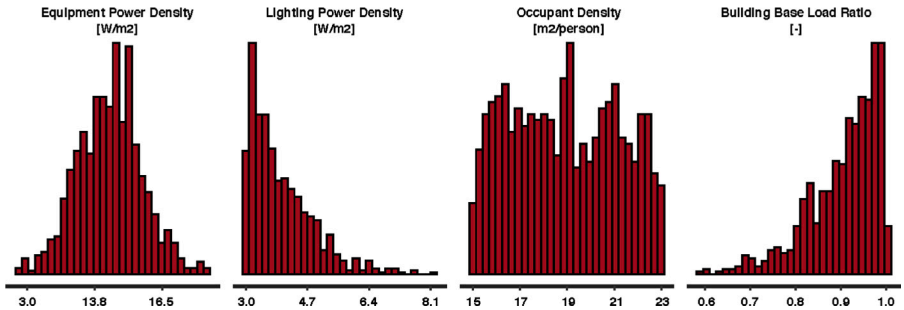


Fig. A.11. Posterior distribution of calibration parameters for Case B.

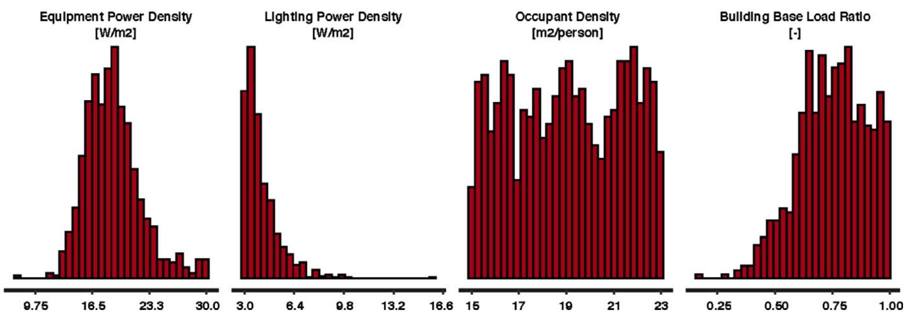


Fig. A.12. Posterior distribution of calibration parameters for Case C.

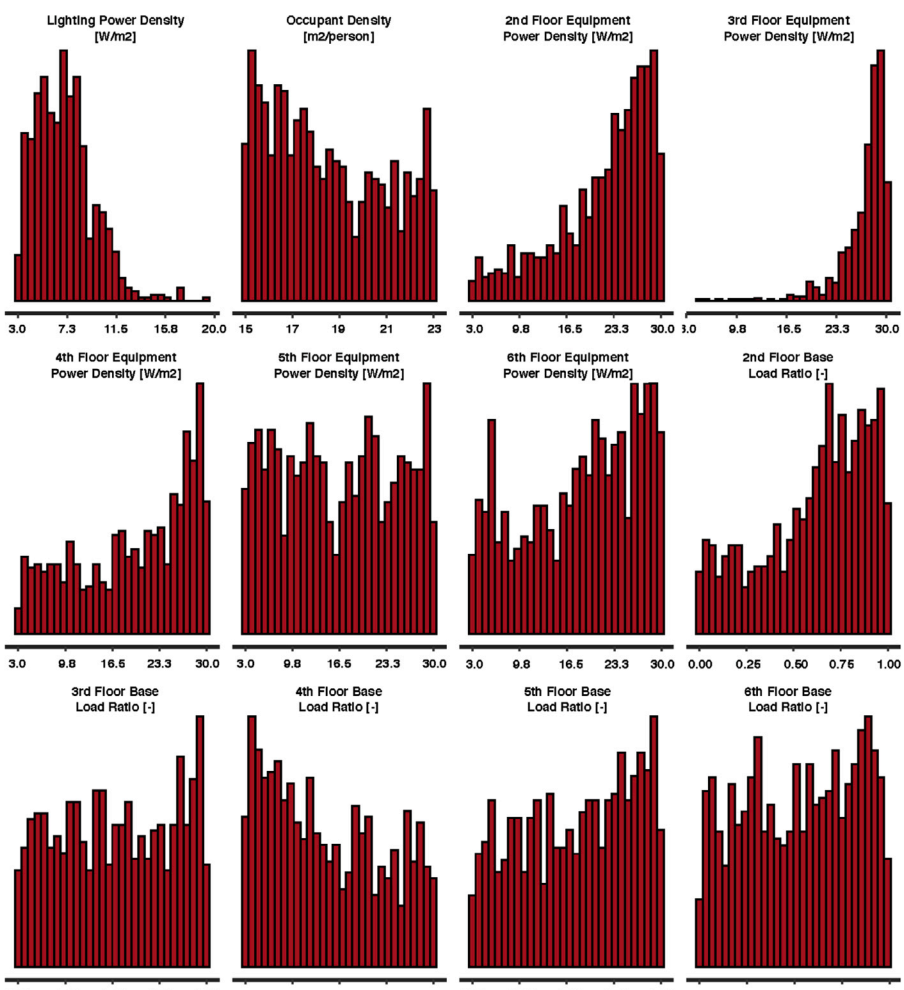


Fig. A.13. Posterior distribution of calibration parameters for Case D.

each pathway, it would inadvertently lead to an increase in overall uncertainty despite a reduction in model bias or inadequacy arising from more detailed modeling.

If applied to more buildings, the outcome of this study would provide a basis for quantifying occupancy uncertainty during design, thus preempting the performance gap through better quantification of the variability in the predictions.

CRediT authorship contribution statement

Adrian Chong: Conceptualization, Methodology, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization. **Godfried Augenbroe:** Methodology, Writing - review & editing, Supervision. **Da Yan:** Writing - review & editing, Supervision.

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.

Acknowledgments

This research is supported by the National University of Singapore [grant number R-296-000-190-133].

Appendix. Posterior distribution

See Figs. A.10–A.13.

References

[1] IEA EBC annex 53: Total energy use in buildings—Analysis and evaluation methods, Energy Build, 2017;152:124–36. http://dx.doi.org/10.1016/j.enbuild.2017.07.038.

[2] Yan D, O'Brien W, Hong T, Feng X, Gunay HB, Tahmasebi F, et al. Occupant behavior modeling for building performance simulation: Current state and future challenges. Energy Build 2015;107:264–78. http://dx.doi.org/10.1016/j.enbuild.2015.08.032.

[3] Heydarian A, McIlvennie C, Arpan L, Yousefi S, Syndicus M, Schweiker M, et al. What drives our behaviors in buildings? A review on occupant interactions with building systems from the lens of behavioral theories. Build Environ 2020;106928. http://dx.doi.org/10.1016/j.buildenv.2020.106928.

[4] Webber CA. Field surveys of office equipment operating patterns. 2001.

[5] Turner C, Frankel M, et al. Energy performance of LEED for new construction buildings. New Build Inst 2008;4:1–42.

[6] O'Brien W, Gaetani I, Gilani S, Carlucci S, Hoes P-J, Hensen J. International survey on current occupant modelling approaches in building performance simulation. J Build Perform Simul 2017;10(5–6):653–71. http://dx.doi.org/10.1080/19401493.2016.1243731.

[7] Jia M, Srinivasan RS, Raheem AA. From occupancy to occupant behavior: An analytical survey of data acquisition technologies, modeling methodologies and simulation coupling mechanisms for building energy efficiency. Renew Sustain Energy Rev 2017;68:525–40.

[8] Chen S, Yang W, Yoshino H, Levine MD, Newhouse K, Hinge A. Definition of occupant behavior in residential buildings and its application to behavior analysis in case studies. Energy Build 2015;104:1–13. http://dx.doi.org/10.1016/j.enbuild.2015.06.075.

[9] Peng C, Yan D, Wu R, Wang C, Zhou X, Jiang Y. Quantitative description and simulation of human behavior in residential buildings. Build Simul 2012;5(2):85–94. http://dx.doi.org/10.1007/s12273-011-0049-0.

[10] IEA. Project summary report - annex 53. AECOM Ltd; 2016, URL http://www.ecbcs.org/Data/publications/EBC_PSR_Annex53.pdf,

[11] Clevenger CM, Haymaker J. The impact of the building occupant on energy modeling simulations. In: Joint international conference on computing and decision making in civil and building engineering, Montreal, Canada. Citeseer; 2006, p. 1–10.

[12] Sun K, Hong T. A framework for quantifying the impact of occupant behavior on energy savings of energy conservation measures. Energy Build 2017;146:383–96. http://dx.doi.org/10.1016/j.enbuild.2017.04.065.

[13] Azar E, Menassa CC. A comprehensive analysis of the impact of occupancy parameters in energy simulation of office buildings. Energy Build 2012;55:841–53. http://dx.doi.org/10.1016/j.enbuild.2012.10.002.

[14] O'Neill Z, Niu F. Uncertainty and sensitivity analysis of spatio-temporal occupant behaviors on residential building energy usage utilizing Karhunen-Loève expansion. Build Environ 2017;115:157–72. http://dx.doi.org/10.1016/j.buildenv.2017.01.025.

[15] Gaetani I, Hoes P-J, Hensen JL. Occupant behavior in building energy simulation: Towards a fit-for-purpose modeling strategy. Energy Build 2016;121:188–204. http://dx.doi.org/10.1016/j.enbuild.2016.03.038.

[16] ASHRAE. ASHRAE 90.1, energy efficient design of new buildings except low-rise residential buildings. Atlanta, Georgia: American Society of Heating, Refrigerating, and Air Conditioning Engineers; 2019.

[17] Duarte C, Van Den Wymelenberg K, Rieger C. Revealing occupancy patterns in an office building through the use of occupancy sensor data. Energy Build 2013;67:587–95. http://dx.doi.org/10.1016/j.enbuild.2013.08.062.

[18] Ahn K-U, Kim D-W, Park C-S, De Wilde P. Predictability of occupant presence and performance gap in building energy simulation. Appl Energy 2017;208:1639–52. http://dx.doi.org/10.1016/j.apenergy.2017.04.083.

[19] Wu W, Dong B, Wang QR, Kong M, Yan D, An J, et al. A novel mobility-based approach to derive urban-scale building occupant profiles and analyze impacts on building energy consumption. Appl Energy 2020;278:115656. http://dx.doi.org/10.1016/j.apenergy.2020.115656.

[20] Davis III JA, Nutter DW. Occupancy diversity factors for common university building types. Energy Build 2010;42(9):1543–51. http://dx.doi.org/10.1016/j.enbuild.2010.03.025.

[21] D'Oca S, Hong T. Occupancy schedules learning process through a data mining framework. Energy Build 2015;88:395–408. http://dx.doi.org/10.1016/j.enbuild.2014.11.065.

[22] Yu T. Modeling occupancy behavior for energy efficiency and occupants comfort management in intelligent buildings. In: 2010 ninth international conference on machine learning and applications. IEEE; 2010, p. 726–31. http://dx.doi.org/10.1109/ICMLA.2010.111.

[23] Reinhart CF. Daylight availability and manual lighting control in office buildings: simulation studies and analysis of measurement [Ph.D. Thesis], Fraunhofer-IRB-Verlag; 2001.

[24] Wang D, Federspiel CC, Rubinstein F. Modeling occupancy in single person offices. Energy Build 2005;37(2):121–6. http://dx.doi.org/10.1016/j.enbuild.2004.06.015.

[25] Page J, Robinson D, Morel N, Scartezzini J-L. A generalised stochastic model for the simulation of occupant presence. Energy Build 2008;40(2):83–98. http://dx.doi.org/10.1016/j.enbuild.2007.01.018.

[26] Sun K, Yan D, Hong T, Guo S. Stochastic modeling of overtime occupancy and its application in building energy simulation and calibration. Build Environ 2014;79:1–12. http://dx.doi.org/10.1016/j.buildenv.2014.04.030.

[27] Wang C, Yan D, Jiang Y. A novel approach for building occupancy simulation. In: Building simulation, vol. 4, Springer; 2011, p. 149–67. http://dx.doi.org/10.1007/s12273-011-0044-5.

[28] Dong B, Andrews B, Lam KP, Höynck M, Zhang R, Chiou Y-S, et al. An information technology enabled sustainability test-bed (ITEST) for occupancy detection through an environmental sensing network. Energy Build 2010;42(7):1038–46. http://dx.doi.org/10.1016/j.enbuild.2010.01.016.

[29] Berger C, Mahdavi A. Review of current trends in agent-based modeling of building occupants for energy and indoor-environmental performance analysis. Build Environ 2020;173:106726. http://dx.doi.org/10.1016/j.buildenv.2020.106726.

[30] Mahdavi A, Tahmasebi F. Predicting people's presence in buildings: An empirically based model performance analysis. Energy Build 2015;86:349–55. http://dx.doi.org/10.1016/j.enbuild.2014.10.027.

[31] Gilani S, O'Brien W, Gunay HB. Simulating occupants' impact on building energy performance at different spatial scales. Build Environ 2018;132:327–37. http://dx.doi.org/10.1016/j.buildenv.2018.01.040.

[32] Trčka M, Hensen JL. Overview of HVAC system simulation. Autom Constr 2010;19(2):93–9. http://dx.doi.org/10.1016/j.autcon.2009.11.019.

[33] Chong A, Menberg K. Guidelines for the Bayesian calibration of building energy models. Energy Build 2018;174:527–47. http://dx.doi.org/10.1016/j.enbuild.2018.06.028.

[34] Yi DH, Kim DW, Park CS. Parameter identifiability in Bayesian inference for building energy models. Energy Build 2019;198:318–28. http://dx.doi.org/10.1016/j.enbuild.2019.06.012.

[35] Yan D, Jin Y, Sun H, Dong B, Ye Z, Li Z, et al. Household appliance recognition through a Bayes classification model. Sustainable Cities Soc 2019;46:101393. http://dx.doi.org/10.1016/j.scs.2018.12.021.

[36] Ouf MM, Issa MH, Azzouz A, Sadick A-M. Effectiveness of using WiFi technologies to detect and predict building occupancy. Sustain Build 2017;2:1–10. http://dx.doi.org/10.1051/sbuild/2017005.

[37] Melfi R, Rosenblum B, Nordman B, Christensen K. Measuring building occupancy using existing network infrastructure. In: 2011 international green computing conference and workshops. IEEE; 2011, p. 1–8.

[38] Balaji B, Xu J, Nwokafor A, Gupta R, Agarwal Y. Sentinel: Occupancy based HVAC actuation using existing WiFi infrastructure within commercial buildings. In: Proceedings of the 11th ACM conference on embedded networked sensor systems. 2013, p. 1–14. http://dx.doi.org/10.1145/2517351.2517370.

[39] Salimi S, Hammad A. Critical review and research roadmap of office building energy management based on occupancy monitoring. Energy Build 2019;182:214–41. http://dx.doi.org/10.1016/j.enbuild.2018.10.007.
[40] Yang J, Santamouris M, Lee SE. Review of occupancy sensing systems and occupancy modeling methodologies for the application in institutional buildings. Energy Build 2016;121:344–9. http://dx.doi.org/10.1016/j.enbuild.2015.12.019.
[41] Yu C, Du J, Pan W. Improving accuracy in building energy simulation via evaluating occupant behaviors: A case study in Hong Kong. Energy Build 2019;202:109373. http://dx.doi.org/10.1016/j.enbuild.2019.109373.
[42] Sangogboye FC, Arendt K, Jradi M, Veje C, Kjærgaard MB, Jørgensen BN. The impact of occupancy resolution on the accuracy of building energy performance simulation. In: Proceedings of the 5th conference on systems for built environments. 2018, p. 103–6. http://dx.doi.org/10.1145/3276774.3276784.
[43] O'Brien W, Gunay HB, Tahmasebi F, Mahdavi A. A preliminary study of representing the inter-occupant diversity in occupant modelling. J Build Perform Simul 2017;10(5–6):509–26. http://dx.doi.org/10.1080/19401493.2016.1261943.
[44] Chong A, Xu W, Chao S, Ngo N-T. Continuous-time Bayesian calibration of energy models using BIM and energy data. Energy Build 2019;194:177–90. http://dx.doi.org/10.1016/j.enbuild.2019.04.017.
[45] Kennedy MC, O'Hagan A. Bayesian calibration of computer models. J R Stat Soc Ser B Stat Methodol 2001;63(3):425–64. http://dx.doi.org/10.1111/1467-9868.00294.
[46] Higdon D, Kennedy M, Cavendish JC, Cafeo JA, Ryne RD. Combining field data and computer simulations for calibration and prediction. SIAM J Sci Comput 2004;26(2):448–66. http://dx.doi.org/10.1137/S1064827503426693.
[47] Lim H, Zhai ZJ. Comprehensive evaluation of the influence of meta-models on Bayesian calibration. Energy Build 2017;155:66–75. http://dx.doi.org/10.1016/j.enbuild.2017.09.009.
[48] Stein M. Large sample properties of simulations using Latin hypercube sampling. Technometrics 1987;29(2):143–51. http://dx.doi.org/10.1080/00401706.1987.10488205.
[49] Chong A, Lam KP. A comparison of MCMC algorithms for the Bayesian calibration of building energy models. In: Proceedings of the 15th IBPSA building simulation conference, vol. 4, 2017.
[50] Menberg K, Heo Y, Choudhary R. Efficiency and reliability of Bayesian calibration of energy supply system models. In: Proceedings of the 15th IBPSA building simulation conference. 2017.
[51] Chong A, Lam KP, Pozzi M, Yang J. Bayesian calibration of building energy models with large datasets. Energy Build 2017;154:343–55. http://dx.doi.org/10.1016/j.enbuild.2017.08.069.
[52] De Wit S, Augenbroe G. Analysis of uncertainty in building design evaluations and its implications. Energy Build 2002;34(9):951–8. http://dx.doi.org/10.1016/S0378-7788(02)00070-1.
[53] Heo Y, Choudhary R, Augenbroe G. Calibration of building energy models for retrofit analysis under uncertainty. Energy Build 2012;47:550–60. http://dx.doi.org/10.1016/j.enbuild.2011.12.029.
[54] Khosravi A, Nahavandi S, Creighton D, Atiya AF. Lower upper bound estimation method for construction of neural network-based prediction intervals. IEEE Trans Neural Netw 2010;22(3):337–46. http://dx.doi.org/10.1109/TNN.2010.2096824.
[55] ASHRAE. ASHRAE guideline 14, measurement of energy, demand, and water savings. Atlanta, Georgia: American Society of Heating, Refrigerating, and Air Conditioning Engineers; 2014.
[56] Rafsanjani HN, Ghahramani A. Extracting occupants' energy-use patterns from Wi-Fi networks in office buildings. J Build Eng 2019;26:100864. http://dx.doi.org/10.1016/j.jobe.2019.100864.
[57] Zhan S, Chong A. Building occupancy and energy consumption: Case studies across building types. Energy Built Environ 2020. http://dx.doi.org/10.1016/j.enbenv.2020.08.001.
[58] Martani C, Lee D, Robinson P, Britter R, Ratti C. ENERNET: Studying the dynamic relationship between building occupancy and energy consumption. Energy Build 2012;47:584–91. http://dx.doi.org/10.1016/j.enbuild.2011.12.037.
[59] Kim Y-S, Heidarinejad M, Dahlhausen M, Srebric J. Building energy model calibration with schedules derived from electricity use data. Appl Energy 2017;190:997–1007. http://dx.doi.org/10.1016/j.apenergy.2016.12.167.
[60] Wang Z, Hong T, Piette MA, Pritoni M. Inferring occupant counts from Wi-Fi data in buildings through machine learning. Build Environ 2019;158:281–94. http://dx.doi.org/10.1016/j.buildenv.2019.05.015.
[61] Gaetani I, Hoes P-J, Hensen JL. A stepwise approach for assessing the appropriate occupant behaviour modelling in building performance simulation. J Build Perform Simul 2020;13(3):362–77. http://dx.doi.org/10.1080/19401493.2020.1734660.
[62] Gan TY, Dlamini EM, Biftu GF. Effects of model complexity and structure, data quality, and objective functions on hydrologic modeling. J Hydrol 1997;192(1–4):81–103.
[63] Paudel R, Jawitz JW. Does increased model complexity improve description of phosphorus dynamics in a large treatment wetland? Ecol Eng 2012;42:283–94. http://dx.doi.org/10.1016/j.ecoleng.2012.02.014.
[64] Del Giudice D, Reichert P, Bareš V, Albert C, Rieckermann J. Model bias and complexity–understanding the effects of structural deficits and input errors on runoff predictions. Environ Model Softw 2015;64:205–14. http://dx.doi.org/10.1016/j.envsoft.2014.11.006.
[65] Robinson S. Conceptual modelling for simulation Part I: Definition and requirements. J Oper Res Soc 2008;59(3):278–90.
[66] Augenbroe G. The role of simulation in performance-based building. In: Hensen JL, Lamberts R, editors. Building performance simulation for design and operation. Routledge; 2012.