

# Continuous-time Bayesian calibration of energy models using BIM and energy data



Adrian Chong<sup>a,\*</sup>, Weili Xu<sup>b</sup>, Song Chao<sup>a</sup>, Ngoc-Tri Ngo<sup>a,c</sup>

<sup>a</sup> Department of Building, School of Design and Environment, National University of Singapore, 4 Architecture Drive, Singapore 117566, Singapore

<sup>b</sup> BuildSimHub Inc, 322 North Shore Dr, Suite 200, PA 15212, USA

<sup>c</sup> The University of Danang - University of Science and Technology, 54 Nguyen Luong Bang Street, Danang city, Vietnam

## ARTICLE INFO

### Article history:

Received 28 December 2018

Revised 3 April 2019

Accepted 14 April 2019

Available online 17 April 2019

### Keywords:

Energy simulation

Continuous calibration

Bayesian calibration

Uncertainty analysis

Building information models (BIM)

Green building XML (gbXML)

## ABSTRACT

The present study proposes a framework for the continuous Bayesian calibration of whole building energy simulation (BES) models utilizing data from building information models (BIM) and building energy management systems (BEMS). The ability to import data from BIM and BEMS provides the potential to significantly reduce the time and effort needed for the continuous calibration of BES models. First, five gbXML geometric test cases were used to check the BIM to BES model translation. Translation of the test cases indicates good geometric agreement between the native BIM and the gbXML-based BES model. An actual building calibration case study (with BIM and three years of monthly electrical energy consumption data) was then used to evaluate the proposed continuous calibration method. The results suggest that compared to a non-continuous approach, the continuous Bayesian calibration method showed reduced prediction uncertainty and improved prediction accuracy on a test dataset. The paper also presents information and comparison of the coefficient of variance of the root mean square error (CVRMSE) and the normalized mean biased error (NMBE), recommending looking at their distributions when working with probabilistic BES predictions.

© 2019 Elsevier B.V. All rights reserved.

## 1. Introduction

Buildings and constructions account for 36% of global energy consumption and are responsible for 19% of greenhouse gas emissions [1]. Based on projections, emissions from the building sector could double by 2050 if we carry on business as usual. Furthermore, buildings use energy throughout their life-cycle, with the operating phase being one of the most significant (80–90%) contributors to a building's overall life-cycle energy demand [2]. Therefore, it is essential that buildings are energy efficient and sustainable.

Building energy simulation (BES) is a valuable tool often used to evaluate and identify strategies for improving building energy efficiency and occupant comfort. Traditionally, BES tools were used for code compliance [3,4], as well as to evaluate different design alternatives, and support improved building performance through design optimization [5,6]. However, over the past decade, BES usage has been extended beyond the design phase with applications throughout the building's life-cycle, including measurement and verification (M&V) [7–9], retrofit analysis [10,11], fault detection

[12], and model predictive controls [13,14]. The major difference between BES usage at different phases of a building's life-cycle is in the level of details, namely, more details are needed to validate the energy model during later stages [15].

Despite the potential benefits of BES, it is often preceded by the labor-intensive, time-consuming, and error-prone manual task of creating an energy model, requiring inputs that include (but not limited to) a detailed description of the building's geometry, its associated heating, ventilation, and air conditioning (HVAC) system, the quantification of various internal loads (occupancy, lighting, equipment loads, etc.), as well as weather conditions [15,16]. Furthermore, with the emergence of the Internet of Things (IoT), there is an opportunity to create a digital twin of a building that continuously learns and updates itself using information from various sources. Information needed for BES could include data from building information models (BIM), real-time observations from the building energy management system (BEMS), as well as assumptions from codes and standards.

### 1.1. BIM to BES

BIM is a process for creating and managing information on a building. The resulting building information model that is created

\* Corresponding author.

E-mail address: [adrian.chong@nus.edu.sg](mailto:adrian.chong@nus.edu.sg) (A. Chong).

represents the building and acts as a database of coordinated information [17]. BIM is intended to serve as a shared knowledge resource for information about the building, forming a reliable basis for decision making. This provides an opportunity to leverage data or information using BIM, potentially streamlining several key steps in the traditional manual process that is used to create a BES model.

The Industry Foundation Class (IFC) [18] and the Green Building eXtensible Markup Language (gbXML) [19] are two open standard data schema commonly used for information exchange from BIM to BES [20]. Today, both IFC and gbXML data formats are supported by many BIM authoring tools such as Revit and ArchiCAD. However, accurate modeling of as-built condition still remains a challenge in the translation of BIM to BES [21]. To close this gap, both fully automated and semi-automated processes and frameworks have been proposed to aid the translation process. Several studies proposed methods to provide a more accurate representation of the building's thermal properties, including the use of thermographic images to infer material thermal properties [22,23], the mapping of actual thermal properties inferred from collections of thermal images and simple environmental measurements [24], as well as the manual extension of material properties by leveraging on published libraries of construction and material data [25]. Manual calibration [26] and calibration using optimization [27] have also been applied to BIM-based BES. To address issues of inconsistencies between design and operation, Dong et al. [12] presented an infrastructure to integrate building energy management system data with BIM-based BES for fault detection and diagnostics.

Despite these advances in automation, model discrepancies and uncertainties still exist because it is impractical and often impossible to gather complete information about the actual conditions of the building. Additionally, it has been shown that assuming default values in these instances may result in large model uncertainties and possibly leading to wrong conclusions [28].

### 1.2. Uncertainties related to BIM-based BES

Inputs to BES are often obtained from multiple sources such as drawings, specifications, site visits, or any other sources of information that might be available. Throughout a building's life-cycle, the BES model often needs to be updated to provide an accurate digital representation of the built asset. However, creating and updating the model is a time-consuming process that is filled with uncertainties. Time-consuming because each update typically involves manual translation as well as tedious line and surface drawing if the change is geometry related. Uncertain because complete information is often not available and the modeler often needs to make assumptions about the missing parts. As a result, BES models are often prone to significant amounts of specification and modeling uncertainties [29].

Specification uncertainty refers to uncertainty arising from incomplete or inaccurate modeling that results from a lack of information on the exact properties of the building. For instance, defining lighting and plug loads require either building use schedules and electrical design data or hourly measurements from electrical sub-meters, which are typically not available and thus often estimated or assumed. Similarly, construction layers and material properties about different building elements are usually estimated if the information is not available in any form. In addition, material properties even if they are known, are subject to a certain degree of heterogeneity. Modeling uncertainty refers to the uncertainty that arises from simplifications and assumptions introduced during the development of the model. These simplifications and assumptions could be explicit to the modeler (e.g., defining thermal zones is a subjective process that depends on the modeler's knowledge, experience, and understanding of the building) or im-

PLICIT and hidden within the tool (simplifications made by the simulation engine). Models are representations of the actual physical building systems, and no single model is perfect and beyond dispute even if the best-fitting values of all the inputs to the model are known [30]. Therefore, it is integral that uncertainty analysis is a part of the overall continuous calibration process.

### 1.3. BES calibration

BES calibration is a necessary step in the overall process to increase model confidence and reliability if the model is to be used beyond the design stage in applications such as retrofit analysis, fault detection, and building controls. This is because actual energy consumption can deviate substantially from modeled predictions even when the model has been updated with data from as-built drawings and specifications [31,32]. However, the calibration of BIM-based BES is often neglected or performed manually. Manual calibration is a subjective process that typically requires a deep understanding of the building's systems and its operation, making the calibration process subjective and hard to reproduce. Furthermore, manual methods are time-consuming and labor-intensive due to its iterative nature.

Over the years, many BES calibration methods have been proposed with varying levels of automation and requiring different levels of expertise [33–35]. Among various calibration methods, Bayesian calibration has been increasingly applied to BES because of its ability to naturally incorporate expert knowledge (through the specification of prior probabilities) in the calibration process [36]. Bayesian calibration was first applied to BES for retrofit analysis [10] with an emphasis on the inclusion of model discrepancy using the mathematical formulation developed by Kennedy and O'Hagan [30]. A Gaussian process (GP) model is often used as an emulator during the iterative Bayesian calibration process because of its flexibility and accuracy. The use of GP models have been found to provide the best accuracy but with the highest computational burden as compared to other commonly used meta-models [37]. To alleviate the high computation cost of Bayesian calibration, several methods have been proposed, including the use of linear regression models in place of GP models [38,39], Hamiltonian Monte Carlo (HMC) for more efficient Markov Chain Monte Carlo sampling [40–42], and using a representative subset of the data for the calibration [40].

Bayesian calibration of BES is typically performed after the measurement period, using the data collected in a single calibration exercise. However, the energy model can become out of date if there are changes to the building's operation (e.g., changes to space utilization or type of lighting fixtures). With the proliferation of the Internet of Things (IoT) technology, it is becoming increasingly easier to obtain a building's operating data. Combined with the ability to translate data from BIM to BES, this provides an opportunity to continuously collect and use both BIM and energy data for the calibration of energy models.

### 1.4. Objectives

The aim of the present study is the development of a continuous calibration framework that utilizes different sources of information while accounting for parameter and model uncertainties. To achieve this, the objectives of this study include:

1. Development of a framework for continuous Bayesian calibration;
2. Implementing an open source, object-oriented BIM to BES translator;
3. Demonstrate the use of BIM and energy data for the continuous Bayesian calibration, using an actual building as a calibration case study.

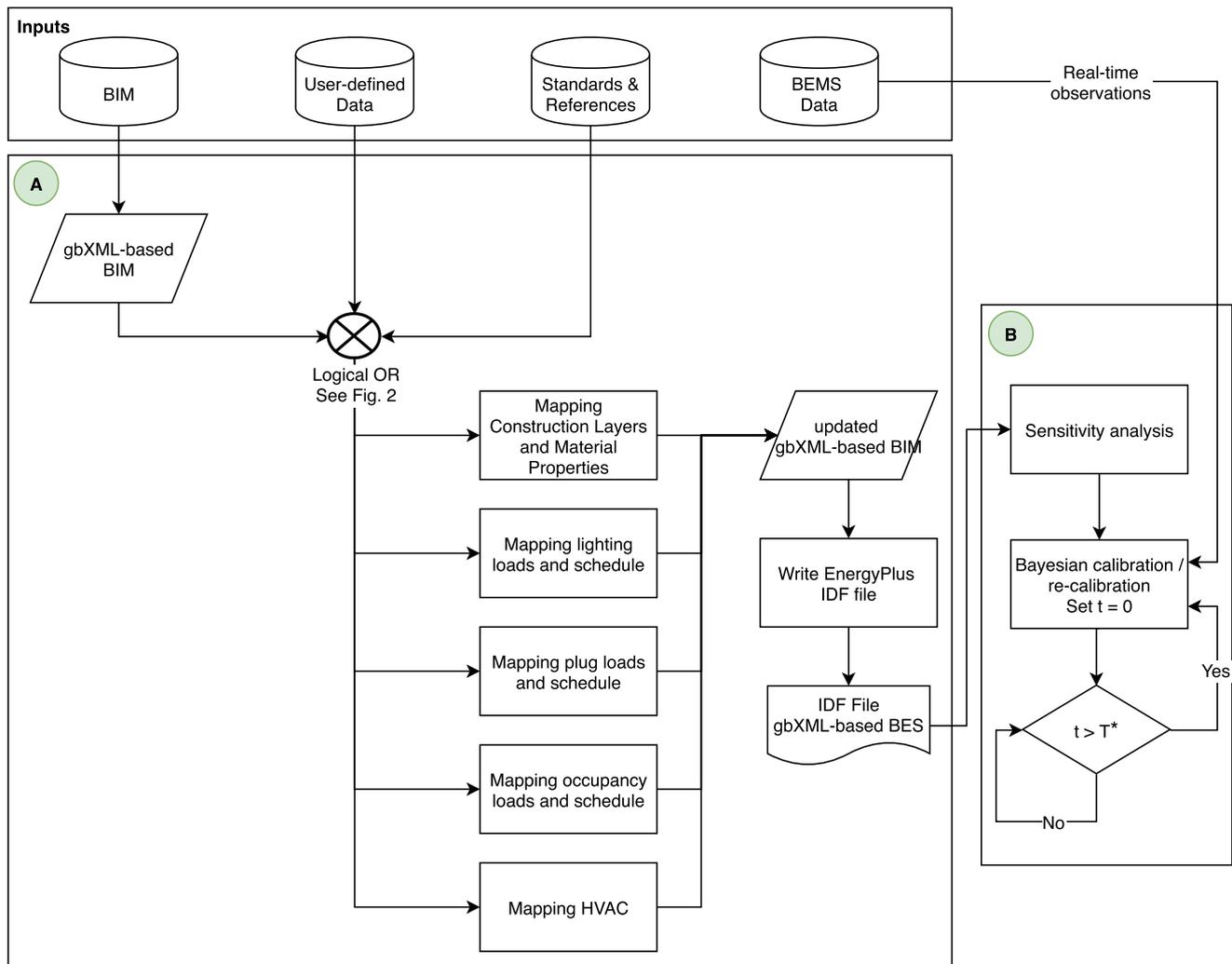


Fig. 1. Overview of proposed framework for continuous Bayesian calibration using data from BIM and BEMS.

## 2. Method

The proposed framework for this study consists of two main parts (Fig. 1):

- (A) BIM to BES translation; and
- (B) Continuous Bayesian calibration.

Part A involves using an object-oriented approach to populate the BIM exported gbXML with information that includes construction layers and material properties, internal loads (lighting, plug, and occupancy), as well as HVAC systems information. This process eliminates time-consuming efforts needed to create a ready-to-simulate EnergyPlus input data file (IDF) because essential information needed to generate a BES model is often missing from a BIM exported gbXML [43]. The updated gbXML-based BIM is then translated into an EnergyPlus IDF. Part B involves the continuous application of Bayesian calibration using the BIM-based BES (EnergyPlus IDF) created in part A and energy data from the BEMS. As “new data” arrives, the “existing observations” are not discarded but instead assimilated to the model through the use of priors.

### 2.1. BIM to BES translation<sup>1</sup>

An object-oriented approach for the BIM to BES translation was adopted in this study. The translator was built with two core

modules: the gbXML-IDF module and the IDDParse module (Fig. 2). The gbXML-IDF module processes the data from the BIM exported gbXML and translates them into an internal data structure. While translating, the gbXML-IDF module is used to continuously validate the translated data to ensure its validity with EnergyPlus. This validation was carried out based on the EnergyPlus input data dictionary (IDD), which provides the syntax and data model for each EnergyPlus input object [16].

As illustrated in Fig. 1, the gbXML-IDF module processes the BIM exported gbXML in a component-based order from construction layers and material properties to internal loads (lighting, plug, and occupancy) to the HVAC systems and controls. At each component, a two layer validation would be performed by the IDDParse module. First, a value validation ensures that the mapped data has the correct data type (alpha or numeric), data range (minimum and maximum values), and references (list of model inputs that references the object). This layer of validation is necessary to prevent severe errors that could result in simulation failure. Second, a format validation restructures the data into EnergyPlus object-based key-value pair data structure. This layer of validation ensures that the simulation engine understands the data being parsed.

During the data mapping process, the gbXML-IDF module uses a plug-in interface to provide a gateway for importing data into the gbXML during run-time. Similar to most of the application programming interfaces (API), the plug-in interface provides methods

<sup>1</sup> Source code available at <https://github.com/weilixu/gbEPlus>.

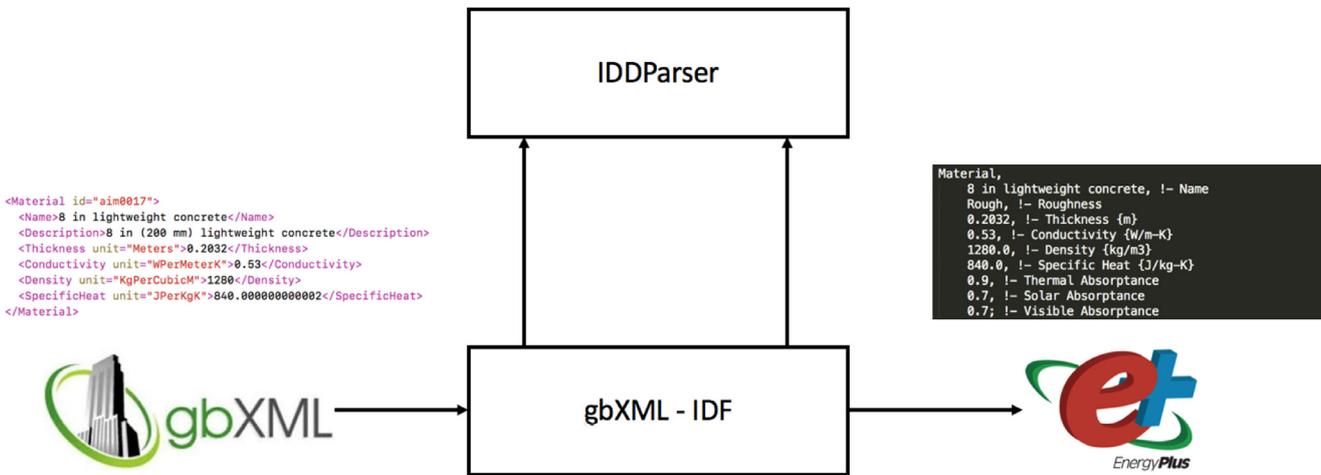


Fig. 2. gbXML-EnergyPlus translator.

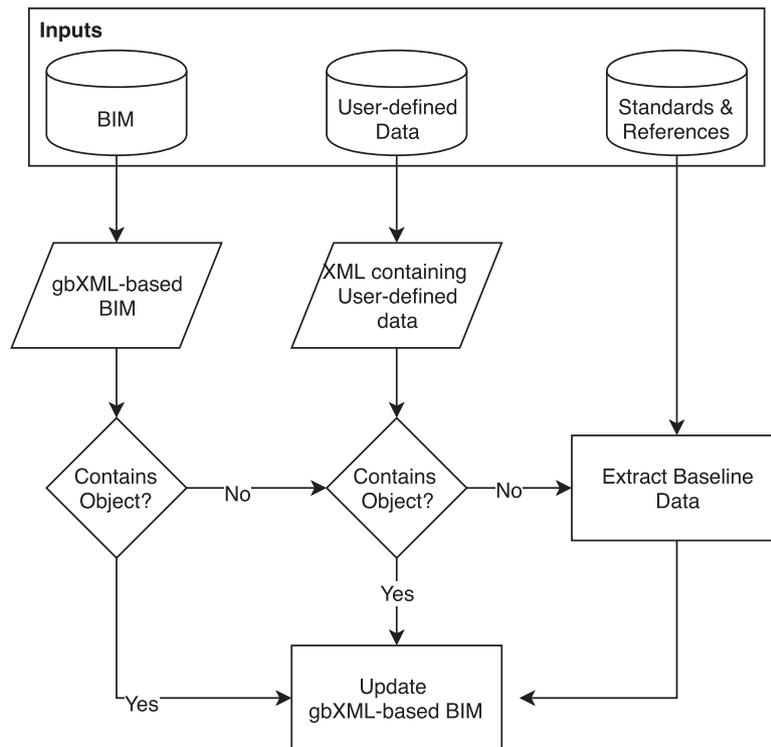


Fig. 3. Logical rules for generating inputs to update the gbXML-based BIM that would subsequently be used to create the BES model.

accessing the translator so that customized datasets can be packaged and plugged into the translation process. Since the plug-in interface is method-driven, it does not require these datasets to be formatted using a predefined schema. Through these methods, the translator is able to receive data from different data sources, providing flexibility in the BIM to BES translation. Put differently, users can interface with the translation by using the methods of the plug-in interface to specify inputs from various data sources and with different priorities.

In this study, inputs to each component of the BES model were generated based on the following order using the plug-in interface (Fig. 3): (1) gbXML-based BIM, (2) user-defined data, and (3) baseline data. In other words, for each EnergyPlus object, inputs would be generated from the gbXML-based BIM first.

The system would then check if any user-defined data has been provided. If the gbXML does not contain the data and no user-defined data has been provided, the inputs would be generated using information from standards and references. The suggested data prioritization was meant to emulate an energy modeler's workflow. Currently, a full set of baseline data (including construction layers, material properties, lighting requirements, outdoor air requirements, and HVAC systems) from ASHRAE Standard 90.1 Appendix G [44] and the Singapore Building and Construction Authority (BCA) Green Mark Scheme [3] has been built into the "Standards and References" database. With such a framework (Fig. 3), the translator would not require the BIM to contain all information before a BES model that is ready for simulation could be generated.

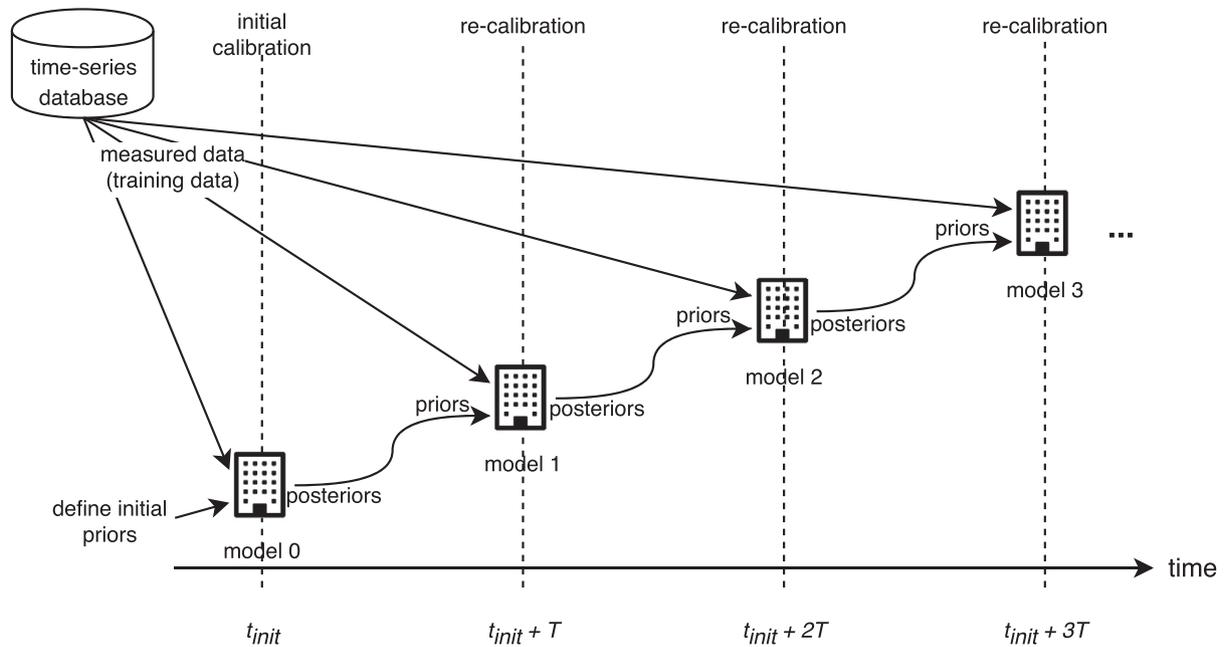


Fig. 4. Framework for continuous Bayesian calibration. The BES model is re-calibrated after every time period  $T$  with priors defined based on the posterior samples of the previously “calibrated” model.

## 2.2. Continuous Bayesian calibration

### 2.2.1. Continuous calibration

Continuous Bayesian calibration is used to improve and maintain confidence in the BIM translated BES. The continuous calibration is carried out following the principles of a receding horizon or model predictive control [45]. Instead of performing the calibration once, the BES model is continuously calibrated after each time interval  $T$  (Fig. 4). The smaller time window also decomposes the calibration process into smaller, computationally tractable calibration problems.

A Bayesian approach fits nicely into a continuous-time calibration framework because it provides a flexible framework for dynamically updating model parameters and their uncertainties while taking into account past information in the form of priors. The continuous Bayesian updating is based on Bayes’ Theorem (Eq. (1)), and the objective is to derive the posterior probability distribution of the calibration parameters given some measured data and prior knowledge of the calibration parameters.

$$p(t | y) \propto p(y | t) \times p(t) \quad (1)$$

During re-calibration, the posterior estimates of the previously calibrated BES model will be used to derive the prior probability distributions. These priors combined with recent measurements will then be used to derive the posterior probability estimates for the updating or re-calibration of the BES model. This is with the exception of the first model since no previously calibrated model exists and the prior probability distributions will be based on the subjective belief of the modeler, which is typically derived from multiple sources of information such as drawings, specifications, site surveys and audits, standards and references.

The continuous Bayesian calibration framework works as follows (Fig. 4). At time  $t$ , we consider a time interval extending  $T$  time-steps into the future:  $t, t + 1, \dots, t + T$ . We then carry out the following steps:

1. Define priors. Posteriors of the previous model are used to derive prior probability distributions through the use of maximum

likelihood estimation (MLE) and Akaike information criterion (AIC).

2. Form a predictive model. Replace all unknown quantities over the time interval with their current estimates, using data available at time  $t$ .
3. Execute. Generate samples from the posterior distributions using the model from step 2.
4. Repeat. Continuously perform steps 1–3 after every  $T$  time-steps.

We use the Akaike information criterion (AIC) (Eq. (2)) to select the probability distribution that gives the “best” fit to the data (posterior samples of the previous model) [46]. For this study, the data is fitted to 5 different continuous probability distributions and they include the Beta, Gamma, Lognormal, Normal, and the Weibull distribution.

$$AIC = -2 \log \mathcal{L}(\theta | x) + 2k \quad (2)$$

where  $\theta$  denotes the parameters of the probability distribution and is determined using maximum likelihood estimation;  $x$  denotes the posterior samples of the previous model and  $k$  is the number of parameters to be estimated. Given the posterior samples  $x$ , maximum likelihood estimation maximizes  $\mathcal{L}(\theta | x)$  over all possible  $\theta$ .

### 2.2.2. Kennedy and O’Hagan formulation<sup>2</sup>

Bayesian calibration is carried out using the formulation proposed by Kennedy and O’Hagan [30]. The formulation explicitly models three types of uncertainties in the BES model (Eq. (3)):

1. parameter uncertainty;
2. model inadequacy or the discrepancy between the energy model and the true building behavior; and
3. Observation errors.

$$y(x) = \eta(x, t) + \delta(x) + \epsilon(x) \quad (3)$$

where,  $y(x)$  is the observed field measurement,  $\eta(x, t)$  is the output of the BES given observable inputs  $x$  and calibration parameters  $t$ ,  $\delta(x)$  is the model inadequacy, and  $\epsilon(x)$  is the observation errors.

<sup>2</sup> Source code available at <https://github.com/adChong/bc-stan>.

Since the iterative calibration process can be computationally intensive, a Gaussian process (GP) model is used to emulate the energy model. We use a GP model because of its flexibility and accuracy [37,47]. Measured data is combined with simulation data using the approach proposed by Hidgon et al. [47]. Due to the complexity of the BES, the posterior distributions cannot be derived analytically and Markov chain Monte Carlo (MCMC) is used to sample from the posterior probability distributions. As suggested by Chong et al. [40], we use Hamiltonian Monte Carlo (HMC) for more efficient MCMC sampling and better convergence [41,42]. Details and source code for the Bayesian calibration method employed in this study can be found in Chong and Menberg [48].

### 2.3. Calibration performance

The performance of the calibrated BES models was analyzed based on their predictive performance. Following ASHRAE Guideline 14 [8], the two indices used for evaluating predictive performance are the coefficient of variance of the root mean square error (CVRMSE) (Eq. (4)) and the normalized mean biased error (NMBE) (Eq. (5)):

$$CVRMSE[\%] = 100 \times \frac{\sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2 / (n-1)}}{\bar{y}} \quad (4)$$

$$NMBE[\%] = 100 \times \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{(n-1) \times \bar{y}} \quad (5)$$

where  $y_i$  = observed value at hour  $i$ ,  $\hat{y}_i$  = predicted value at hour  $i$ ,  $\bar{y}$  = mean energy consumption of  $n$  observations; and  $n$  = number data points.

Given that Bayesian calibration produces probabilistic predictions, we analyze the distributions of CVRMSE and NMBE computed using the MCMC samples from the posterior predictions. We also analyze the CVRMSE and NMBE computed using the mean posterior predictions:

$$\hat{y}_i = \frac{1}{m} \sum_{j=1}^m \hat{y}_{i,j} \quad (6)$$

where  $m$  is the number of MCMC samples in the posterior predictions of  $\hat{y}_i$  in Eqs. (4) and (5) above.

Another important consideration in MCMC simulations and thus Bayesian calibration is that of convergence. Since a continuous calibration approach was adopted, newly calibrated models would continue to be generated over time. Convergence for every calibrated model is assessed by checking that the Gelman–Rubin statistics  $\hat{R}$  (ratio of between-chain variance to within-chain variance) for the posterior samples are within  $1 \pm 0.1$  [48,49].

### 3. Geometric test cases and calibration case study

Five geometric test cases and one actual building calibration case study were used in this study. The five geometric benchmark test-cases were created based on ASHRAE Research Project 1468 [50], and are used to ensure compliance of the BIM to BES translation with gbXML's standards for geometry accuracy and information integrity [51]. The five test cases are used to ensure that the requirements for geometric data exchange from BIM to BES are met, while the calibration case study is used to demonstrate and evaluate the application of the proposed framework (using BIM and energy data for continuous Bayesian calibration). Since the purpose of the geometric test cases is to validate geometry data mapping, their BIM was created to contain only geometric information. The five test cases and one case study are summarized as follows (Table 1):

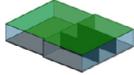
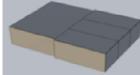
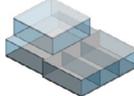
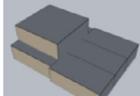
- Five geometric benchmark gbXML test-cases:
  - Test case 1: Double height space sharing an adjacent interior wall with four single height spaces.
  - Test case 2: One of the spaces shares an adjacent interior wall with three identical single height spaces.
  - Test case 3: A Simple two-zone model with complex sloping roof
  - Test case 4: Large space that consists of a sloped slab on grade and an unusual roof geometry
  - Test case 5: This test case is a slight variant of test case 2 and includes an additional second story volume.
- One actual building calibration case study:
  - The case study building is an office building located on the campus of the National University of Singapore (NUS) in Singapore. The building consists of 3 blocks each 3-story high with a total gross floor area of 5445 m<sup>2</sup>. The air-conditioning and mechanical ventilation (ACMV) system is a variable air volume system served by a central district cooling system. Given Singapore's tropical climate, no heating is required and the ACMV system is required to operate throughout the year to maintain thermal comfort within the building. The BIM was created using Revit and contains both the actual construction layers and material properties. 3 years of electrical energy consumption data at a monthly resolution from 1 January 2014 to 31 December 2016 were collected and used to evaluate the proposed framework as illustrated in Fig. 1. The weather file used for calibration is the Actual Meteorological Year (AMY) weather data from the Singapore Changi airport weather station (WMO #486980).

## 4. Results

### 4.1. BIM to BES

Table 1 provides a 3D visual comparison of the model in the native BIM (Revit) and the BES (EnergyPlus). The images show that

**Table 1**  
Comparison of 3D view for revit model as compared to BIM-based EnergyPlus model.

	Native BIM (Revit)	Energy model (EnergyPlus)
Test case 1		
Test case 2		
Test case 3		
Test case 4		
Test case 5		
Actual case study		

**Table 2**

Weighted absolute percentage error (WAPE) for gross wall area, window opening area, and zone volume of the BIM translated BES.

	Gross wall area [%]	Window opening area [%]	Roof area [%]	Zone volume [%]
Test case 1	0.004	0.015	0.001	0.000
Test case 2	0.0146	NA*	0.000	0.000
Test case 3	3.720	0.018	0.002	0.001
Test case 4	1.716	0.011	0.001	0.000
Test case 5	0.907	NA*	0.000	0.000
Actual case study	0.309	1.549	0.035	0.000

\* Test cases 2 and 5 have no windows.

the translator was able to correctly identify the correct number of spaces/zones. Table 2 presents the weighted absolute percentage error (WAPE) (Eq. (7)) for the gross wall area, window opening area, roof area and zone volumes for the BIM-based BES in EnergyPlus as compared to the original native BIM in Revit. WAPE is used as a measure of the BIM to BES translation accuracy in percentage terms and is calculated as the sum of absolute deviation divided by the total area or volume. Therefore, the lower the WAPE, the better the performance of the translation. For instance, a WAPE of 5% in gross wall area can be interpreted as a 5% deviation in wall surface area between the BIM and the translated BES model.

$$WAPE[\%] = 100 \cdot \frac{\sum_{i=1}^n |A_i - F_i|}{\sum_{i=1}^n A_i} \quad (7)$$

where  $A_i$  is the area of the surface or volume of the zone in the BIM, and  $F_i$  is the area of the surface or volume of the zone in the BES.

WAPE values for all test cases (including the actual case study) are very small ( $\leq 1\%$ ) for most surface areas and all zone volumes. This is with the exception of the gross wall surface area of test case 3 (WAPE = 3.720%) and test case 4 (WAPE = 1.716%), as well as the window opening surface area for the actual case study (WAPE = 1.549%). A more detailed analysis revealed that the difference in gross wall surface area of test cases 3 and 4 is due to the roofs tilting at an angle. This is because walls in the BIM are translated as surfaces in the BES model. During the translation, the inner surface of the wall object is used to represent this surface in the BES model, resulting in walls with a slightly larger wall area since the wall would be extended to the roof (Fig. 5). Nonetheless, the BIM translated BES model maintains a correct thermal view since the zone volume remains the same. It should be noted that this error is due to translation from the native BIM to gbXML. This was verified by computing the wall surface areas of the model in the BIM exported gbXML and comparing it with their counterparts in the translated BES model. The WAPE was also found to be equal to 0.000% in both test cases 3 and 4 when the comparison was made between gbXML and BES instead of between native BIM and BES.

The slightly larger WAPE (1.549%) for the window opening area of the actual case study is because windows are modeled as curtain wall objects in Revit. However, when exported to gbXML, the frames of these curtain wall objects are not exported resulting in a discrepancy between the native BIM and the BIM exported gbXML. A further check also showed that the WAPE between the BIM exported gbXML and the translated BES to be equal to 0.000%.

#### 4.2. Parameter screening

Before conducting the continuous BES calibration, Morris method [52] was used to screen-out non-influential parameters for the actual case study. Twenty-eight  $t_1, t_2, \dots, t_{28}$  parameters were initially identified as uncertain after the BIM to BES translation. The twenty-eight parameters include parameters related to envelope thermal properties, internal loads, and the HVAC system.

Table A.4 lists the twenty-eight parameters that were considered in the sensitivity analysis as well as the initial value and range assigned to each parameter. The initial values of each parameter were assigned based on measured data, as-built drawings, specifications, and expert opinion.

Fig. 6 shows the results of the sensitivity analysis using the modified mean  $\mu^*$  and standard deviation  $\sigma$  [53]. From Fig. 6, five parameters (equipment power density  $t_{27}$ , lighting power density  $t_{26}$ , AHU fan pressure rise  $t_{28}$ , AHU fan efficiency  $t_{21}$ , cooling setpoint  $t_{20}$ ) were selected and subsequently used for the continuous Bayesian calibration. Non-influential parameters were fixed at their initial values (Table A.4) to mitigate issues of non-identifiability during the BES model calibration [48].

#### 4.3. Calibrated simulation performance (actual calibration case study)

##### 4.3.1. Prior probability specification

In the proposed continuous Bayesian calibration method, priors for the calibration parameters would be derived based on the corresponding posteriors from the previously calibrated model. However, the first BES to be calibrated does not have a previous model that can be used for the derivation of the priors. Therefore, the following choices for the priors were made to initialize the continuous Bayesian calibration process:

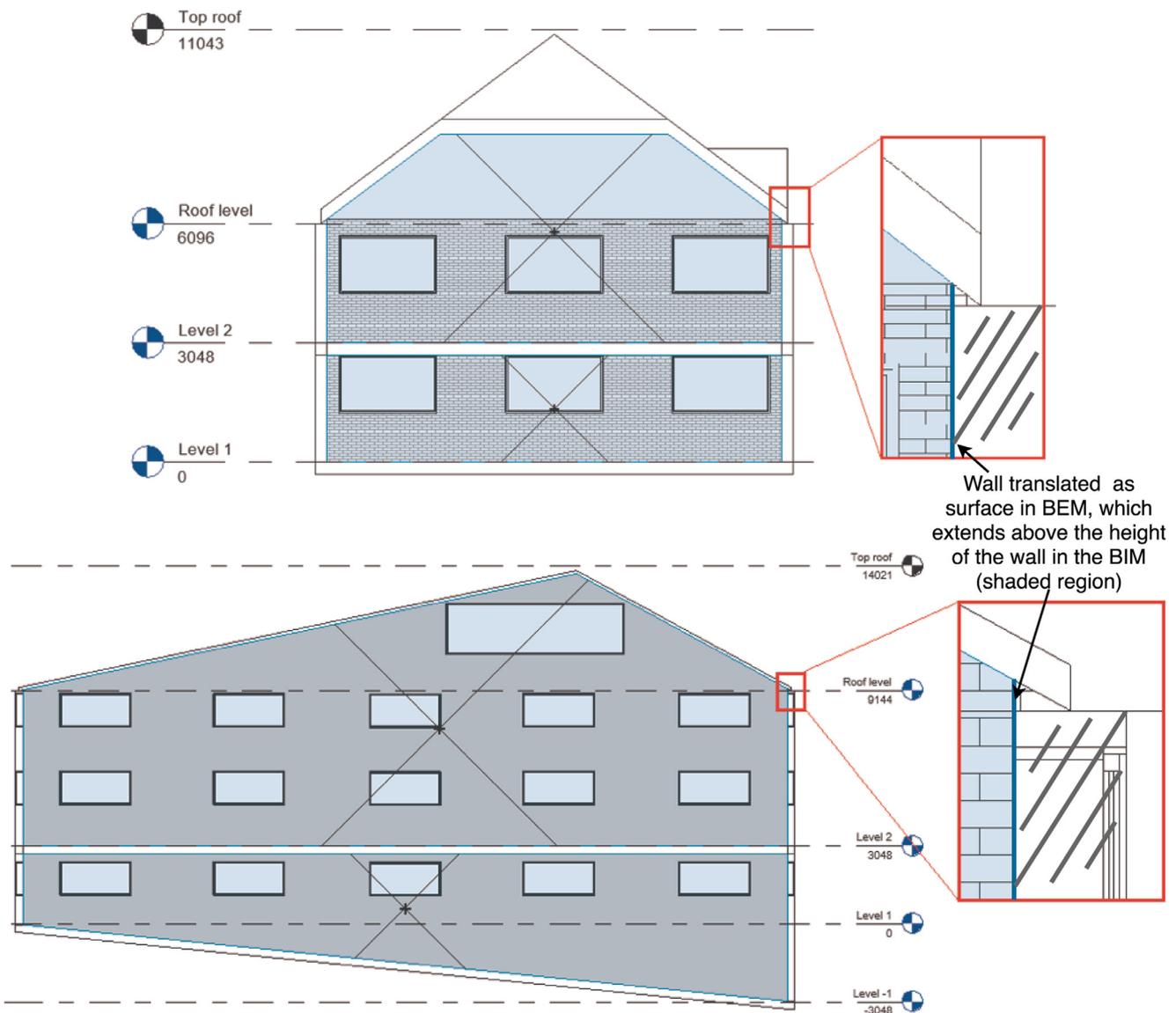
- Equipment power density [ $W/m^2$ ]  $\sim \mathcal{U}(2, 15)$
- Lighting power density [ $W/m^2$ ]  $\sim \mathcal{U}(2, 12)$
- Cooling setpoint temperature [ $^{\circ}C$ ]  $\sim \mathcal{U}(19, 26)$
- AHU fan efficiency [–]  $\sim \mathcal{U}(0.2, 0.9)$
- AHU fan pressure rise [ $m^3/s$ ]  $\sim \mathcal{U}(100, 500)$

These calibration parameters were assigned flat or uniform prior probability distributions to represent the subjective belief that there is insufficient prior knowledge to assign specific informative priors. The upper and lower bounds were defined using a wider range of possible values than expected to avoid eliminating possible parameter values based on subjective beliefs apriori [48].

##### 4.3.2. Training and testing data

In this study, prediction accuracy was evaluated using hold-out testing data that was not used for the calibration. For the non-continuous calibration, of the 3 years of monthly electricity energy consumption data used for this case study, Bayesian calibration was carried out once using the first 12 months of data (Fig. 7). The calibrated BES model is then used to obtain posterior predictions over the following 24 months (hold-out testing data).

For the continuous calibration, prediction accuracy was evaluated by treating future data that was not used for the calibration of the current model as the hold-out testing data. Fig. 7 illustrates this process within the context of this case study. As shown in the figure, the BES model is re-calibrated every  $T = 1$  month using the most recent 12 months of historical data. This means that for each calibrated model, the testing data used to evaluate its prediction accuracy is the most imminent  $T = 1$  month of future data that has not been used for its calibration.



**Fig. 5.** Differences in wall area in BIM to BES translation because the wall is translated as a surface in the BES model, resulting in the wall being slightly taller. Top: Cross section of test case 3; Bottom: Cross section of test case 4.

#### 4.3.3. Prediction accuracy

Fig. 8 shows the resulting posterior predictions over the training data (top plot) and the testing data using both continuous and non-continuous Bayesian calibration (lower plot). A quick glance at Fig. 8 shows a significant difference between predictions over the training data (top plot) and the testing data (bottom plot). Predictions over the training data have significantly lower posterior uncertainty and show an exceptional match between the predictions and the measured values.

Predictions over the test dataset using non-continuous Bayesian calibration shows larger prediction uncertainties (yellow box-plots in Fig. 8) as compared to predictions with continuous Bayesian calibration (blue box-plots in Fig. 8). Additionally, non-continuous calibration tends to underestimate actual values, with most measurements falling above the third quartile of the box-plots. On the contrary, bias in the predictions is significantly reduced when continuous Bayesian calibration is used, with most measurements falling within the inter-quartile range of the box plots. This observation of reduced bias can be attributed to the use of more recent data in continuous Bayesian calibration because the model is updated

or re-calibrated every  $T = 1$  month, keeping the model up-to-date with the latest information about the building's operation. These results (Fig. 8) suggest a significant difference in posterior uncertainty when using training or testing data, and a better overall fit and lower bias with continuous Bayesian calibration relative to non-continuous Bayesian calibration.

We further evaluate the proposed continuous calibration method using varying amounts of historical data (12 months, 6 months and 1 month) to calibrate or train the model. Table 3 shows the CVRMSE and NMBE computed using the mean posterior predictions (Eq. (6)) while Fig. 9 shows their corresponding distributions (computed using the MCMC samples from the posterior predictions). Table 3 shows that when we calibrate the model with less training data, CVRMSE decreases. On the contrary, Fig. 9 shows that using 6 months of training data gives the best performance. Comparing Table 3 and Fig. 9 also shows that CVRMSE computed with the mean posterior predictions tends to be underestimated (i.e., accuracy is overestimated). Therefore, given probabilistic outputs, the analysis should be carried out on the whole distribution of CVRMSE and NMBE.

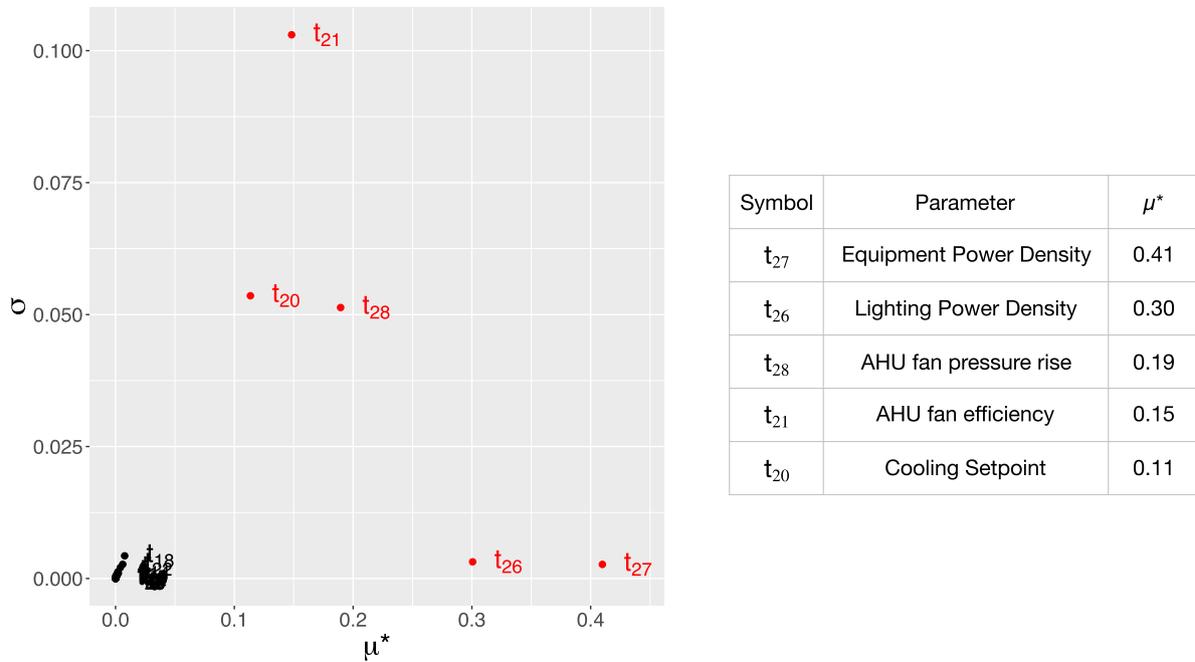


Fig. 6. Graphical plot of sensitive measures  $\mu^*$  and  $\sigma$ . The closer the parameters are to the upper right the more influential the parameter.

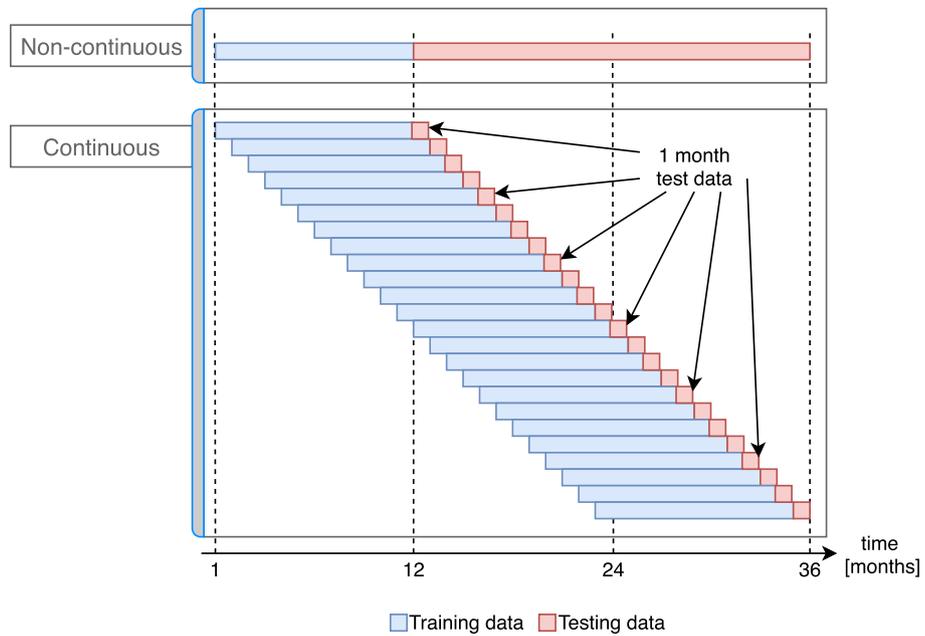
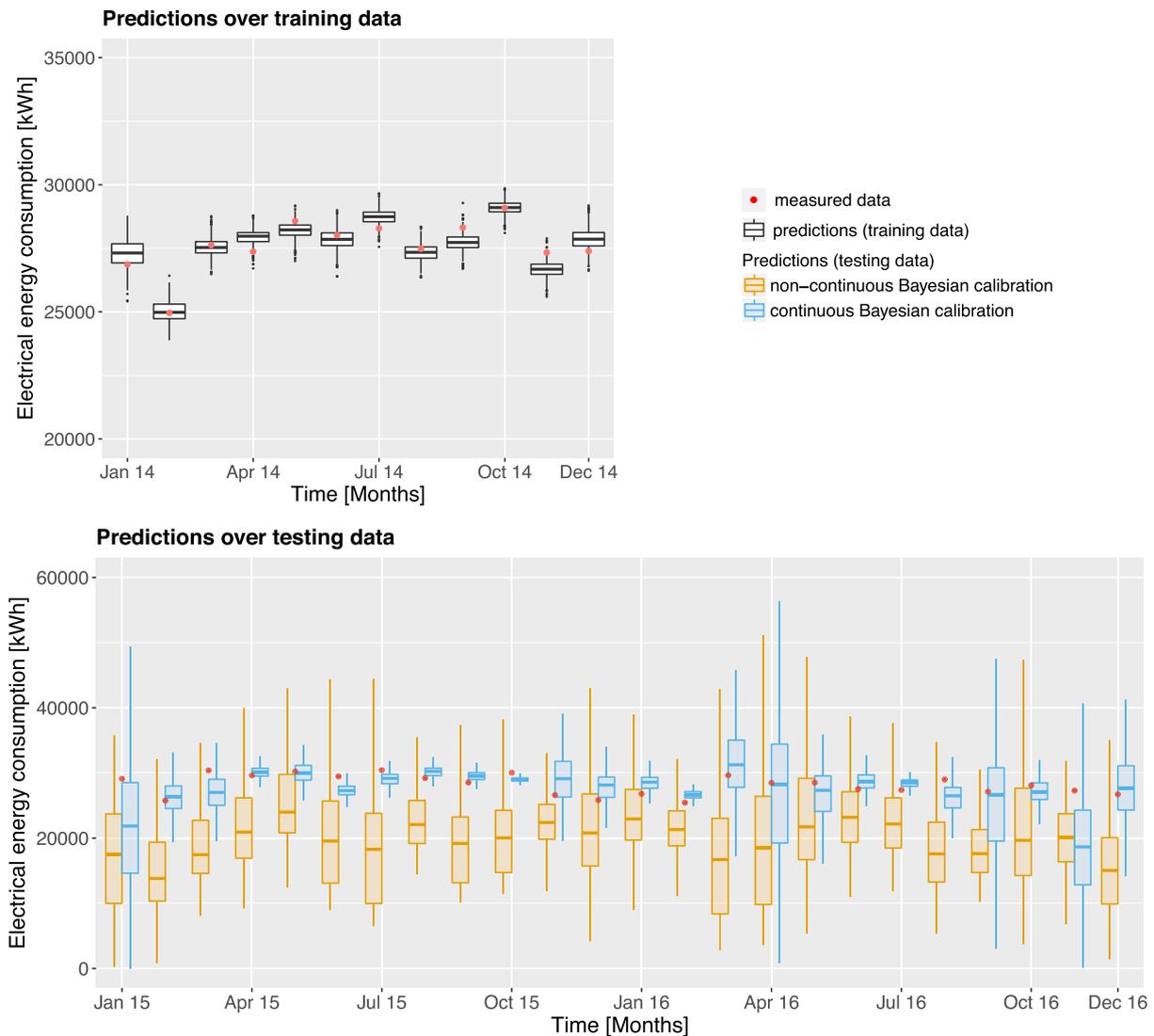


Fig. 7. Partitioning of the training and hold-out testing data for the non-continuous and continuous calibration.

**Table 3**  
CVRMSE and NMBE over testing data computed using mean value of the posterior predictions.

Metric	Predictions over Training data	Predictions over testing data			
		Non-continuous (training = 12 months)	Continuous (training = 12 months)	Continuous (training = 6 months)	Continuous (training = 1 month)
CVRMSE [%]	1.04	29.78	10.12	9.85	7.25
NMBE [%]	-0.00	-28.41	-2.78	0.52	2.33



**Fig. 8.** Box and whisker plots showing posterior predictions over training data (top); and over testing data using non-continuous and continuous Bayesian calibration (bottom).

An examination of the CVRMSE in Fig. 9 reveals that the 75th percentile using 6 months of training data is lower than the 25th percentile using 1 month of training data, indicating with confidence that CVRMSE is lower with 6 months data. A feasible explanation is that as more data points are added, the mean function of the Gaussian process model used to calibrate the BES model would adjust itself to pass through these points, leading to lower posterior uncertainty closer to the observations [54]. Consequently, larger posterior uncertainties are expected with a smaller training dataset since there is a higher chance that the testing data is different or far away from the observations used to train the GP model.

However, it can be observed in Fig. 9 that using 12 months of data to train the model led to an increase in median CVRMSE and absolute NMBE, as well as a higher variability in both CVRMSE and NMBE. This is because the 6 months of training data might provide a better representation of the future one month of electricity consumption that we are predicting. Using 12 months of data for the calibration might result in noisy inputs and the resulting quality of the estimated GP model suffers due to the increased uncertainties in the training data since the 12 months of data might contain information that might not be representative of the current build-

ing's operation. Additionally, the present building case study is located in Singapore, which is characterized by a hot and humid climate with uniform temperature and high humidity throughout the year. Therefore, the absence of seasonal variation might be another reason for 6 months of training data having better CVRMSE and NMBE performance. Nonetheless, within the context of this case study, we can conclude that 6 months of training data should be used for the continuous Bayesian calibration.

#### 4.3.4. Parameter posterior uncertainty

The posterior distributions of the calibration parameters over time is shown in Fig. 10. Going from left to right, the figure shows the posterior distributions from the initial BES model that was calibrated to the posterior distributions of the BES model that was continuously re-calibrated after 6, 12, 18 and 24 months respectively.

Visually, it can be observed that there is a significant reduction in posterior uncertainty in all five calibration parameters over time. Quantitatively, the reduction in posterior uncertainty can also be observed from the decreasing standard deviation over time (Fig. 10). The reduction in posterior uncertainties can be explained

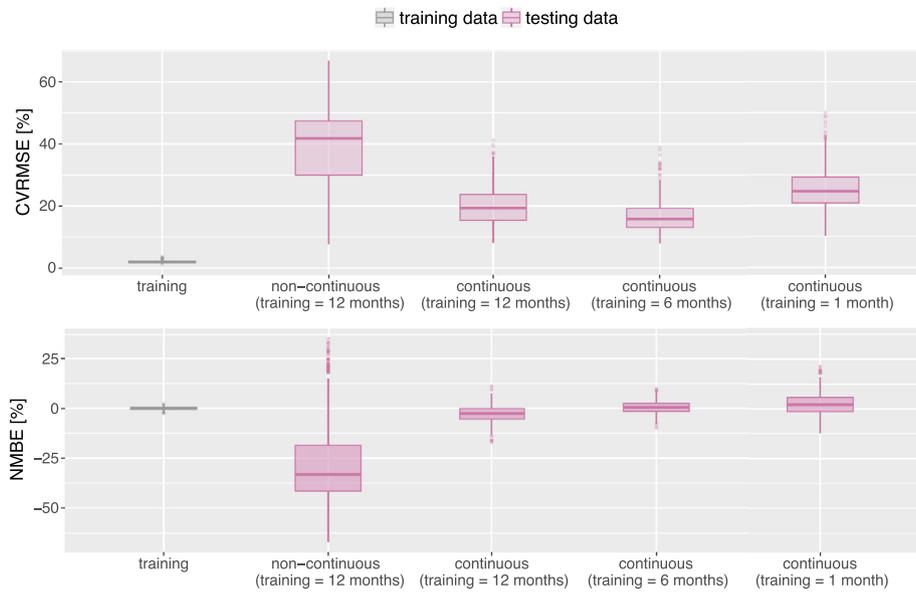


Fig. 9. Box and whisker plots showing the distribution of CVMRSE (top) and NMBE (bottom) computed using samples from the posterior predictions generated by MCMC.

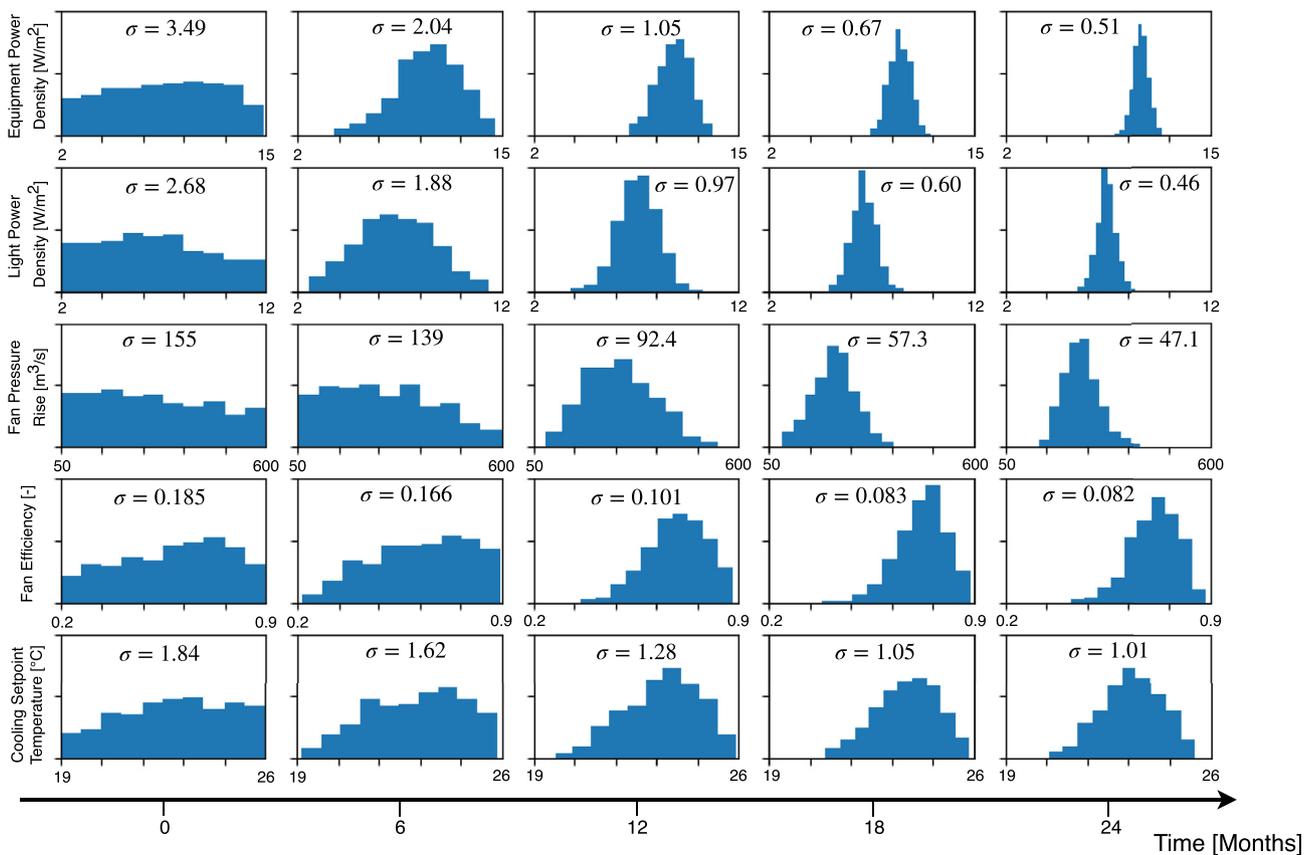


Fig. 10. Evolution of posterior distribution of calibration parameters with time using 6 months of training data for the continuous Bayesian calibration. The x-axis shows the posterior distribution of the calibration parameters as the BES model is calibrated or re-calibrated within the proposed continuous Bayesian calibration framework.

by the specification of more informative priors over time. More informative priors were specified as the model is re-calibrated because trends in the measured data remain fairly consistent. This, in turn, leads to an increase in model confidence as more data or information telling the same thing gets integrated into the model.

### 5. Discussion

Firstly, results of the geometric conversion (using the 5 geometric benchmark test cases) indicate good agreement between the native BIM and the BES model and that errors were due to discrepancies between the native BIM and the exported gbXML

rather than discrepancies between the exported gbXML and the BES. These findings suggest that more has to be done to improve the interoperability between BIM authoring tools and open data standard data schemas such as gbXML. The development of well-established guidelines and BIM integration workflow for energy modeling would also alleviate time-consuming efforts needed to adjust and configure the native-BIM so that it is suitable for export to an open data BIM schema.

Secondly, through an actual building case study, this study showed that compared to non-continuous Bayesian calibration, the proposed continuous Bayesian calibration method provides better accuracy with reduced uncertainty in the resulting posterior predictions and calibration parameters. This is because the proposed method continuously updates the model with more recent data that might be more representative of the building's operation. Through this study, it was also demonstrated that small amounts of training data may result in large uncertainty in the posterior predictions due to lack of observations at similar conditions. On the contrary, large amounts of training data might lead to noisy inputs and reduced accuracy due to increased uncertainties in the training data. In the present case study, six months of training data were found to provide the best fit and the lowest bias with the test dataset. The limitations of the present study naturally include the use of a case study building that is located in Singapore, which is characterized by a hot and humid climate with uniform temperature and high humidity throughout the year. Therefore, it may not generalize to climates that experience annual seasonal variability. To generalize its applicability, future research is needed to test the proposed continuous Bayesian calibration framework to case study buildings located in temperate climates. Additionally, the current study uses monthly calibration data making it difficult to assess if the model is accurately representing the real behavior of the building at smaller time intervals. In a detailed calibration of an office building located in Ireland, it was highlighted that it is possible for a calibrated model to meet the most stringent monthly error criteria without accurately representing the building at an hourly interval [32]. However, data resolution should also depend on the aim of the simulation because increasing the size of the dataset increases the cost<sup>3</sup> of the calibration. Ideally, the calibrated model should be of the lowest complexity while preserving its validity with respect to the intended purpose of the simulation [55]. Therefore, future work includes an investigation on the amount of data and model fidelity needed to achieve the intended purpose of a calibrated BES model.

Lastly, it was demonstrated in this study that the uncertainty and accuracy of the posterior predictions on the training and testing data can be significantly different. Using training data to evaluate calibration performance might instill false confidence in the model's predictive performance on unobserved data. Currently, building energy models are considered "calibrated" when they meet the CVRMSE and NMBE threshold specified by standards [7–9]. The thresholds set out vary depending on whether the models are calibrated with monthly or hourly data but there are no requirements for the use of a test dataset (data that was not used for the calibration). However, in most applications, what we are interested in is not modeling the structure in the training data to achieve high prediction accuracy on historical data. In contrast, the purpose of calibration is typically to produce a BES model that is representative of actual building performance. Therefore, the emphasis of calibration should be on the generalization of the model and its ability to meet the aims of the simulation. In addition, provisions for CVRMSE and NMBE thresholds stipulated in ASHRAE

Guideline 14 [8] remains designed for use with deterministic predictions. Although the most recent version of ASHRAE Guideline 14 [8] includes new provisions to account for uncertainty, they were limited to savings uncertainty. To overcome this, CVRMSE and NMBE are typically computed using the mean value of each time-step's prediction [10,37,40]. We showed that this may overestimate prediction accuracy and propose looking at their distributions instead. In this study, this was done through the use of box-plots to compare the distributions of CVRMSE and NMBE. Alternatively, Kristensen et al. [56] used mean  $\pm 95\%$  confidence interval to compare CVRMSE and NMBE. In another study, the continuous rank probability score (CRPS) [57] was employed as a test statistic for the evaluation of the calibration performance of models with uncertainty [38].

## 6. Conclusion

The present study proposed a framework for the continuous-time Bayesian calibration of building energy simulation (BES) models using data from building information models (BIM), as well as energy data from the building energy management system (BEMS). To extract useful information from a BIM, a BIM to "ready-to-simulate" BES framework was proposed by integrating user-defined information and data from standards and references into the translation process. The proposed continuous Bayesian calibration method extends Kennedy and O'Hagan's [30] Bayesian calibration approach into a continuous calibration framework by integrating the principles of a receding horizon or model predictive control. We use a Bayesian approach because it provides a flexible framework to dynamically update the energy model while accounting for past information in the form of priors. Prior distributions used for re-calibrating the model are defined using the posterior estimates from the previously calibrated model. Combined with recent measurements these priors will be used for the re-calibration.

To check for compliance when mapping geometry data from BIM to BES, five geometric test cases from gbXML [51] were used to test the translation. The test cases showed that discrepancies in the BIM to BES translation were primarily caused by exporting from the native BIM to gbXML, suggesting a need for the development of well-established guidelines and BIM integration workflow for energy modeling.

Through an actual building calibration case study, its BIM and three years of its monthly electrical energy consumption, we showed that the proposed framework could maintain prediction accuracy while reducing parameter posterior uncertainties. We also propose evaluating prediction performance using a hold-out test dataset for better generality and looking at whole distributions of CVRMSE and NMBE when working with probabilistic predictions

## Conflict of interest

None.

## Acknowledgments

This research is funded by the Republic of Singapore's [National Research Foundation](#) (NRF) [grant number [BCA RID 94.17.2.6](#)] through Building and Construction Authority (BCA)'s Green Buildings Innovation Cluster (GBIC) R&D Grant.

<sup>3</sup> large dataset requires more processing and are more expensive to acquire.

## Appendix A. Details of uncertain parameters

**Table A.4**  
List of uncertain parameters and their range.

Model parameter	Symbol	Initial value	Min	Max
Envelope thermal properties:				
Concrete conductivity [W/m – k]	$t_1$	1.442	1.319	1.612
Concrete density [kg/m <sup>3</sup> ]	$t_2$	2400	2352	2448
Concrete specific heat [J/kg – K]	$t_3$	1000	755	1245
Rockwool conductivity [W/m – k]	$t_4$	0.035	0.033	0.037
Rockwool density [kg/m <sup>3</sup> ]	$t_5$	32.0	30.4	33.6
Rockwool specific heat [J/kg – K]	$t_6$	840	798	882
Gypsum board conductivity [W/m – k]	$t_7$	0.167	0.159	0.175
Gypsum board density [kg/m <sup>3</sup> ]	$t_8$	880	836	924
Gypsum board specific heat [J/kg – K]	$t_9$	840	798	882
Precast concrete conductivity [W/m – k]	$t_{10}$	0.660	0.627	0.693
Precast concrete density [kg/m <sup>3</sup> ]	$t_{11}$	2400	2280	2520
Precast concrete specific heat [J/kg – K]	$t_{12}$	1000	950	1050
Window U-value [W/m <sup>2</sup> – k]	$t_{13}$	1.280	1.216	1.344
Window shading coefficient [–]	$t_{14}$	0.311	0.295	0.327
Hollow core slab (roof) conductivity [W/m – k]	$t_{15}$	1.254	1.147	1.393
Hollow core slab (roof) density [kg/m <sup>3</sup> ]	$t_{16}$	2081	2039	2123
Hollow core slab (roof) specific heat [J/kg – K]	$t_{17}$	1000	755	1245
Internal loads:				
People density [m <sup>2</sup> /person]	$t_{18}$	10	8	15
People activity level [W]	$t_{19}$	110.7	70.0	130.0
Light power density [W/m <sup>2</sup> ]	$t_{20}$	5.5	4	12
Light fraction radiant [–]	$t_{21}$	0.32	0.20	0.60
Equipment power density [W/m <sup>2</sup> ]	$t_{22}$	10	4	15
Equipment fraction radiant [–]	$t_{23}$	0.30	0.10	0.35
Infiltration rate [ACH]	$t_{24}$	0.2	0.1	0.4
HVAC:				
Cooling setpoint [°C]	$t_{25}$	24	21	25
Outdoor air flow [m <sup>3</sup> /s – person]	$t_{26}$	0.009	0.006	0.011
AHU fan efficiency [–]	$t_{27}$	0.61	0.3	0.8
AHU fan pressure rise [m <sup>3</sup> /s]	$t_{28}$	170	100	500

## References

- [1] IEA, Global status report 2017, Glob. Alliance Build. Constr. (2017).
- [2] T. Ramesh, R. Prakash, K. Shukla, Life cycle energy analysis of buildings: an overview, *Energy Build.* 42 (10) (2010) 1592–1600.
- [3] BCA, BCA Green Mark for Non-Residential Buildings nrb: 2015, Singapore Building and Construction Authority, 2016.
- [4] L. USGBC, Reference Guide for Building Design and Construction v4, US Green Building Council, Washington, 2013.
- [5] W. Xu, A. Chong, O.T. Karaguzel, K.P. Lam, Improving evolutionary algorithm performance for integer type multi-objective building system design optimization, *Energy Build.* 127 (2016) 714–729.
- [6] K. Sun, T. Hong, S.C. Taylor-Lange, M.A. Piette, A pattern-based automated approach to building energy model calibration, *Appl. Energy* 165 (2016) 214–224.
- [7] EVO, International performance measurement and verification protocol: concepts and options for determining energy and water savings volume 1, *Effic. Valuat. Org.* (2012).
- [8] ASHRAE, Guideline 14–2014, Measurement of Energy, Demand, and Water Savings, American Society of Heating, Refrigerating, and Air Conditioning Engineers, Atlanta, Georgia, 2014.
- [9] US DOE FEMP, M&V guidelines: measurement and verification for performance-based contracts, version 4.0, *Energy Effic. Renew. Energy* (2015).
- [10] Y. Heo, R. Choudhary, G. Augenbroe, Calibration of building energy models for retrofit analysis under uncertainty, *Energy Build.* 47 (2012) 550–560.
- [11] Y. Heo, G. Augenbroe, D. Graziano, R.T. Muehleisen, L. Guzowski, Scalable methodology for large scale building energy improvement: relevance of calibration in model-based retrofit analysis, *Build. Environ.* 87 (2015) 342–350.
- [12] B. Dong, Z. O'Neill, Z. Li, D. Luo, S. Madhusudana, S. Ahuja, T. Bailey, An integrated infrastructure for real-time building energy modeling and fault detection and diagnostics, *Proc. SimBuild.* 5 (1) (2012) 448–455.
- [13] J. Zhao, K.P. Lam, B.E. Ydstie, O.T. Karaguzel, Energyplus model-based predictive control within design–build–operate energy information modelling infrastructure, *J. Build. Perform. Simul.* 8 (3) (2015) 121–134.
- [14] Z. Zhang, A. Chong, Y. Pan, C. Zhang, S. Lu, K.P. Lam, A deep reinforcement learning approach to using whole building energy model for hvac optimal control, in: ASHRAE and IBPSA-USA SimBuild Building Performance Modeling Conference, 2018.
- [15] T. Maile, M. Fischer, V. Bazjanac, Building Energy Performance Simulation Tool-s-a Life-Cycle and Interoperable Perspective, 2007. Center for Integrated Facility Engineering (CIFE) Working Paper.
- [16] LBNL, Energyplus Input Output Reference: The Encyclopedic Reference to Energyplus Input and Output, US Department of Energy, 2016.
- [17] C. Eastman, P. Teicholz, R. Sacks, K. Liston, BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors, John Wiley & Sons, 2011.
- [18] buildingSMART, IFC, 2018, (<http://www.buildingsmart-tech.org>) Accessed: 2019-09-09.
- [19] gbXML, Green Building XML (gbXML) Schema, 2018, ([http://www.gbxml.org/Schema\\_Current\\_GreenBuildingXML\\_gbXML](http://www.gbxml.org/Schema_Current_GreenBuildingXML_gbXML)) Accessed: 2019-09-09.
- [20] B. Dong, K. Lam, Y. Huang, G. Dobbs, A comparative study of the IFC and gbXML informational infrastructures for data exchange in computational design support environments, in: Tenth International IBPSA Conference, 2007, pp. 1530–1537.
- [21] L. Sanhudo, N.M. Ramos, J.P. Martins, R.M. Almeida, E. Barreira, M.L. Simões, V. Cardoso, Building information modeling for energy retrofitting—a review, *Renew. Sustain. Energy Rev.* 89 (2018) 249–260.
- [22] S. Lagüela, L. Díaz-Vilariño, J. Martínez, J. Armesto, Automatic thermographic and rgb texture of as-built bim for energy rehabilitation purposes, *Autom. Constr.* 31 (2013) 230–240.
- [23] S. Lagüela, L. Díaz-Vilariño, J. Armesto, P. Arias, Non-destructive approach for the generation and thermal characterization of an as-built bim, *Constr. Build. Mater.* 51 (2014) 55–61.
- [24] Y. Ham, M. Golparvar-Fard, Mapping actual thermal properties to building elements in gbxml-based BIM for reliable building energy performance modeling, *Autom. Constr.* 49 (2015) 214–224.
- [25] H. Kim, Z. Shen, I. Kim, K. Kim, A. Stumpf, J. Yu, BIM IFC Information mapping to building energy analysis (BEA) model with manually extended material information, *Autom. Constr.* 68 (2016) 183–193.

- [26] C. Miller, D. Thomas, S.D. Irigoyen, C. Hersberger, Z. Nagy, D. Rossi, A. Schlueter, BIM-Extracted energyplus model calibration for retrofit analysis of a historically listed building in Switzerland, *Proc. SimBuild*. 2014 (2014).
- [27] A. Abdelalim, W. O'Brien, Z. Shi, Data visualization and analysis of energy flow on a multi-zone building scale, *Autom. Constr.* 84 (2017) 258–273.
- [28] K.-U. Ahn, Y.-J. Kim, C.-S. Park, I. Kim, K. Lee, BIM Interface for full vs. semi-automated building energy simulation, *Energy Build.* 68 (2014) 671–678.
- [29] S. De Wit, G. Augenbroe, Analysis of uncertainty in building design evaluations and its implications, *Energy Build.* 34 (9) (2002) 951–958.
- [30] M.C. Kennedy, A. O'Hagan, Bayesian calibration of computer models, *J. R. Stat. Soc.* 63 (3) (2001) 425–464.
- [31] A. Pedrini, F.S. Westphal, R. Lamberts, A methodology for building energy modelling and calibration in warm climates, *Build. Environ.* 37 (8) (2002) 903–912.
- [32] P. Raftery, M. Keane, A. Costa, Calibrating whole building energy models: detailed case study using hourly measured data, *Energy Build.* 43 (12) (2011) 3666–3679.
- [33] J. Clarke, P. Strachan, C. Pernot, An approach to the calibration of building energy simulation models, *ASHRAE Trans.* 99 (1993) 917–917.
- [34] T.A. Reddy, Literature review on calibration of building energy simulation programs: uses, problems, procedures, uncertainty, and tools., *ASHRAE Trans.* 112 (1) (2006).
- [35] D. Coakley, P. Raftery, M. Keane, A review of methods to match building energy simulation models to measured data, *Renew. Sustain. Energy Rev.* 37 (2014) 123–141.
- [36] W. Tian, Y. Heo, P. de Wilde, Z. Li, D. Yan, C.S. Park, X. Feng, G. Augenbroe, A review of uncertainty analysis in building energy assessment, *Renew. Sustain. Energy Rev.* 93 (2018) 285–301.
- [37] H. Lim, Z.J. Zhai, Comprehensive evaluation of the influence of meta-models on Bayesian calibration, *Energy Build.* 155 (2017) 66–75.
- [38] Q. Li, G. Augenbroe, J. Brown, Assessment of linear emulators in lightweight Bayesian calibration of dynamic building energy models for parameter estimation and performance prediction, *Energy Build.* 124 (2016) 194–202.
- [39] W. Tian, R. Choudhary, A probabilistic energy model for non-domestic building sectors applied to analysis of school buildings in Greater London, *Energy Build.* 54 (2012) 1–11.
- [40] A. Chong, K.P. Lam, M. Pozzi, J. Yang, Bayesian calibration of building energy models with large datasets, *Energy Build.* 154 (2017) 343–355.
- [41] A. Chong, K.P. Lam, A comparison of MCMC algorithms for the Bayesian calibration of building energy models, in: *Proceedings of the 15th IBPSA Building Simulation Conference*, 2017.
- [42] K. Menberg, Y. Heo, R. Choudhary, Efficiency and reliability of Bayesian calibration of energy supply system models, in: *Proceedings of the 15th IBPSA Building Simulation Conference*, 2017.
- [43] G. Gourlis, I. Kovacic, Building information modelling for analysis of energy efficient industrial buildings—a case study, *Renew. Sustain. Energy Rev.* 68 (2017) 953–963.
- [44] ASHRAE, ANSI/ASHRAE/IES Standard 90.1 Energy Standard for Buildings Except Low-Rise Residential Buildings, American Society of Heating, Ventilating, and Air Conditioning Engineers, Atlanta, Georgia, 2016.
- [45] J. Mattingley, Y. Wang, S. Boyd, Code generation for receding horizon control, in: *2010 IEEE International Symposium on Computer-Aided Control System Design*, IEEE, 2010, pp. 985–992.
- [46] K.P. Burnham, D.R. Anderson, Multimodel inference: understanding AIC and BIC in model selection, *Sociol. Methods Res.* 33 (2) (2004) 261–304.
- [47] D. Higdon, M. Kennedy, J.C. Cavendish, J.A. Cafoe, R.D. Ryne, Combining field data and computer simulations for calibration and prediction, *SIAM J. Sci. Comput.* 26 (2) (2004) 448–466.
- [48] A. Chong, K. Menberg, Guidelines for the Bayesian calibration of building energy models, *Energy Build.* 174 (2018) 527–547.
- [49] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Rubin, *Bayesian Data Analysis*, 2, Taylor & Francis, 2014.
- [50] M. Clayton, J. Haberl, W. Yan, S. Kota, F. Farias, W. Jeong, J. Kim, J.B. Alcocer, Development of a reference building information model (BIM) for thermal model compliance testing, RP-1468, ASHRAE, 2013.
- [51] gbXML, Test case documentation: test cases for software vendors to become gbXML-compliant, 2018. [http://www.gbxml.org/TestCases\\_for\\_GreenBuildingXML\\_gbXML](http://www.gbxml.org/TestCases_for_GreenBuildingXML_gbXML).
- [52] M.D. Morris, Factorial sampling plans for preliminary computational experiments, *Technometrics* 33 (2) (1991) 161–174.
- [53] F. Campolongo, J. Cariboni, A. Saltelli, An effective screening design for sensitivity analysis of large models, *Environ. Model. Softw.* 22 (10) (2007) 1509–1518.
- [54] C.E. Rasmussen, C.K. Williams, *Gaussian Processes for Machine Learning*, 1, MIT Press Cambridge, 2006.
- [55] M. Trčka, J.L. Hensen, Overview of HVAC system simulation, *Autom. Constr.* 19 (2) (2010) 93–99.
- [56] M.H. Kristensen, R.E. Hedegaard, S. Petersen, Hierarchical calibration of archetypes for urban building energy modeling, *Energy Build.* 175 (2018) 219–234.
- [57] T. Gneiting, A.E. Raftery, Strictly proper scoring rules, prediction, and estimation, *J. Am. Stat. Assoc.* 102 (477) (2007) 359–378.