BAYESIAN CALIBRATION OF BUILDING ENERGY MODELS FOR ENERGY RETROFIT DECISION-MAKING UNDER UNCERTAINTY

A Dissertation
Presented to
The Academic Faculty

by

Yeonsook Heo

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
College of Architecture

Georgia Institute of Technology
December, 2011

BAYESIAN CALIBRATION OF BUILDING ENERGY MODELS FOR ENERGY RETROFIT DECISION-MAKING UNDER UNCERTAINTY

Approved by:

Prof. Godfried Augenbroe, Advisor
College of Architecture
Georgia Institute of Technology

Dr. Serge Guillas
Department of Statistical Science
University College London

Prof. Jeff Wu
School of Industrial & Systems Engineering
Georgia Institute of Technology

Dr. Cheol-Soo Park
Architectural Engineering
SungKyunKwan University

Dr. Ruchi Choudhary
Engineering Department
University of Cambridge

Date Approved: September 16, 2011

for my parents

ACKNOWLEDGMENTS

I could not have accomplished the Ph.D. tenure without the full support of my parents and their values and principles I have learned throughout my life. Also, I thank my sister, Jinsook, and my brother, Jung Hwan, who always believed in me and backed my research efforts.

This work would not have been possible without my advisor Prof. Augenbroe. I admire him for being a great scholar with deep understanding and sharp insights on broad research areas and a great advisor who guided me to gain research perspectives and philosophy. I am indebted to Dr. Ruchi Choudhary for the endless supports and research efforts that provided the foundation for me to grow as a researcher. I have to say that I am one of the luckiest Ph.D. students and no one can expect any more than what I have had with them. I am also grateful to Jeff Wu, Serge Guillas, and Cheol-Soo Park for their supports and constructive remarks on my dissertation work.

My colleagues in the Building Technology Program have been not only great companies to work together for research but also good friends to spare precious time with. Many thanks to Atefe, Huafen, Paola, Sanghoon, Fei, Zhengwei, Jihyun, Yuming, Jaeho, Seanhay, Jason, Vishwadeep, and Clarissa. Thanks also to all friends in Georgia Tech who have been always there for me: Jieun, You Jong, Youngjin, Anna, Miklos, Jinkook, Hugo, Tina, Matt, Jinsol, and all others. Each of you made my long Ph.D. journey in Atlanta the best days of my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Retrofitting of existing buildings is essential to reach reduction targets in energy consumption and greenhouse gas emission. In the current practice of a retrofit decision process, professionals perform energy audits, and construct dynamic simulation models to benchmark the performance of existing buildings and predict the effect of retrofit interventions. In order to enhance the reliability of simulation models, they typically calibrate simulation models based on monitored energy use data. The calibration techniques used for this purpose are manual and expert-driven. The current practice has major drawbacks: (1) the modeling and calibration methods do not scale to large portfolio of buildings due to their high costs and heavy reliance on expertise, and (2) the resulting deterministic models do not provide insight into underperforming risks associated with each retrofit intervention.

This thesis has developed a new retrofit analysis framework that is suitable for large-scale analysis and risk-conscious decision-making. The framework is based on the use of normative models and Bayesian calibration techniques. Normative models are light-weight quasi-steady state energy models that can scale up to large sets of buildings, i.e. to city and regional scale. In addition, they do not require modeling expertise since they follow a set of modeling rules that produce a standard measure for energy performance. The normative models are calibrated under a Bayesian approach such that the resulting calibrated models quantify uncertainties in the energy outcomes of a building. Bayesian calibration models can also incorporate additional uncertainties

associated with retrofit interventions to generate probability distributions of retrofit performance. Probabilistic outputs can be straightforwardly translated into a measure that quantifies underperforming risks of retrofit interventions and thus enable decision making relative to the decision-makers' rational objectives and risk attitude.

This thesis demonstrates the feasibility of the new framework on retrofit applications by verifying the following two hypotheses: (1) normative models supported by Bayesian calibration have sufficient model fidelity to adequately support retrofit decisions, and (2) they can support risk-conscious decision-making by explicitly quantifying risks associated with retrofit options. The first and second hypotheses are examined through case studies that compare outcomes from the calibrated normative model with those from a similarly calibrated transient simulation model and compare decisions derived by the proposed framework with those derived by standard practices respectively. The new framework will enable cost-effective retrofit analysis at urban scale with explicit management of uncertainties.

CHAPTER 1 INTRODUCTION

1.1 Importance of Retrofitting Existing Buildings

In the U. S. and European countries, the building sector accounts for 39% of the total energy consumption (EPA, 2008; DECC, 2010a). While the energy consumption of current buildings is projected to grow annually by 1.7% to 2025 (Ryan, 2004), the total floor area of buildings is projected to increase roughly at the rate of 1- 2% per year. According to Commercial Buildings Energy Consumption Survey (EIA, 2003), in 2003 the U.S. has 4.86 million commercial buildings corresponding to 71.6 billion square feet of floor areas, and adds 1.6 billion square feet of new constructed floor areas every year. Owing to the dominant volume of current buildings, energy retrofits of existing buildings are essential to meet energy and greenhouse gas emission reduction targets. Without enhancing performance of existing buildings, it will be difficult to reach the 2030 challenge of 50% reduction in energy consumption from the building sector.

Energy retrofits of existing buildings have gained interest due to growing awareness of energy inefficiency over the building lifecycle. These inefficiencies can result from degradation of materials and equipment, change in use, and/or unexpected faults. The efficiency of a building degrades even faster if it is not maintained properly. Moreover, building systems underperform when they are not properly installed. Indeed, faults in mechanical and lighting systems in a building can account for between 2% and 11% of the total energy consumption for commercial buildings (Roth, 2005).

Performance deficiencies in existing buildings are also emphasized in a study conducted by the Lawrence Berkeley National Laboratory of 643 existing commercial

buildings (Mills, 2009; Mills, 2004). This study has shown that improving existing buildings will yield median energy savings of 16%. Furthermore, this study projected that if these median energy savings are applied to the U.S. commercial building stock, potential energy-savings will correspond to monetary savings of approximately $30 billion by 2030. According to this projection, energy retrofits of existing buildings can play a significant role in achieving national energy reduction targets cost-effectively.

Federal, state, and local governments have established various goals towards reducing energy consumption and greenhouse gas emissions from the building sector. President Obama launched "Better Buildings Initiative" to improve energy efficiency in buildings; one of the targets is to make commercial buildings in U.S. 20% more energy efficient by 2020 through cost-effective retrofit interventions (White House, 2011). Also, the city of Chicago initiated a Chicago Climate Action Plan to mitigate climate change by reducing greenhouse gas emissions (City of Chicago Climate Action, 2011). One of the five strategies in the plan is energy efficient buildings: retrofitting 50% of commercial buildings and residential buildings in Chicago for 30% energy reduction by 2020.

The advantages of investing in building retrofits for energy and environmental benefits have been long recognized at more local, community levels. The Texas Energy Office commenced the Texas LoanSTAR program in 1988, which had provided 191 loans for public buildings by November 2007 for energy retrofits (SECO, 2007). In 1995, the U.S. Department of Energy developed the Rebuild America Program that assists communities to design and implement retrofit projects through community-based partnerships (Brown, 2004). In this program, a community works with a group of local private and public sector organizations to define its energy-saving goals, select buildings for improvements, and develop financial and action plans for energy efficiency improvements. As of October 2003, the program has supported over 560 communities in 53 states, and renovations have been completed for 610 million square feet of building

floor area (EERE, 2003). Recently in 2009, the U.S. Department of Energy announced a $454 million fund for the Retrofit Ramp-Up program to make building retrofits accessible to existing houses and commercial buildings (Department of Energy, 2009). The program selected 25 communities that proposed innovative business models for whole-neighborhood energy retrofits.

In addition to community-level retrofit efforts, large organizations (e.g., campuses, corporate owners, government entities) regard energy retrofits of their facilities as a profitable investment opportunity. Indeed, energy retrofits can be cost-effective for institutions as they reduce the energy costs of large portfolio of buildings while increasing long-term real estate value. Energy-efficiency services for large public-sector facilities yielded $2.8 billion in revenues for the Energy Service Companies in 2008 alone (Satchwell, 2010). Besides, as of May 2011, 25 federal agencies have implemented more than 570 energy retrofit projects under the Federal Energy Management Program (FEMP) to improve energy efficiency of federal government buildings (FEMP, 2011).

1.2 Importance of Risk Analysis in Retrofit Decision-makings

Quantifying risks is important in the case of large-scale/high-cost retrofits since it provides explicit information about underperforming risks associated with each retrofit option. The objective of risk analysis is to quantify the magnitude of savings from retrofit decisions and the likelihood of the savings. Risk analysis typically requires probabilistic analysis, and should be preceded by uncertainty quantification. Many studies have demonstrated the significant role of uncertainty quantification and risk analysis in the context of buildings, for example in the design of HVAC systems (de Wit, 2002), mold risks (Moon, 2007), and energy management in off-grid solar homes (Hu, 2009). These

studies have shown how quantitative knowledge of risk changes the rational choice of the best design option.

Quantitative risk analysis is especially necessary to adequately manage risk in retrofit investments, especially in the context of the Energy Service Companies (ESCOs) industry. ESCOs undertake energy retrofits of existing buildings through energy performance contracts that typically guarantee savings as part of their service. Energy service contracts have been a useful medium for delivering energy-efficiency services to various building sectors (e.g., government entities, schools, universities). The U.S. ESCO industry is expected to grow by an annual growth rate of 26% through 2011 with 75% of the revenues coming from energy performance contracts for building retrofits (Satchwell, 2010). A performance contract guarantees energy-savings or energy-cost savings directly tied to the total cost (service cost) of the improvement contracted by the ESCO. In other words, the service cost is in part determined by the magnitude of the guaranteed cost savings. However, if the savings are overestimated and not realized during the contract period, ESCOs may have to compensate building owners for the shortfall depending on the contract clauses. The expression of a guarantee allows building owners to invest in the retrofits with high confidence, but the structure leads to relatively safe and often less aggressive ambitions towards energy savings.

This means that ESCOs are less likely to recommend high-impact, high-cost technologies, unless the probability of energy savings can be quantified appropriately and associated risks expressed such that comparison between competing technologies is supported adequately. Hence, uncertainty analysis has been emphasized in energy efficiency projects to quantify financial and physical risks in the saving potential from ECMs (Mills, 2006; Mills, 2003; Mathew, 2005). However, there is a lack of sufficient research in developing methods that quantify risks for use in energy performance contracts.

1.3 Current Methods for Evaluating Energy Retrofits

All retrofit projects essentially aim to improve building energy efficiency in a cost effective way by implementing the most optimal mix of technologies and retrofit interventions. In order to achieve this goal for a large portfolio of buildings, retrofitting should be generally preceded by the following steps. First, it is necessary to benchmark each individual building within the portfolio to identify the ones that need energy efficiency improvements most. Second, candidate energy conservation measures (ECMs) must be evaluated in the actual context of identified buildings for selecting the optimal measures. These steps are accomplished by a thorough energy audit of all buildings in the portfolio and using transient simulation models to predict the relative benefits of a set of ECMs. The 'deep' energy audits also serve to calibrate the transient simulation model, so the model accurately reflects the buildings. This thesis will show that this kind of methodology suffers from modeling inefficiencies due to the detailed level of modeling expertise required in the analysis process. It can be reasonably applied only for one or several buildings, but does not scale up to large sets of buildings. Hence, improving the energy efficiency of a large set of buildings will need a new generation of scalable and adaptable modeling methodologies. The modeling methodologies should not only be scalable to evaluate the performance of every building in the portfolio but also be adaptable to represent each building as operated in order to correctly evaluate all feasible ECMs for the particular building.

In the specific context of the ESCO industry, current practice for evaluating the energy saving potential of a building involves Investment Grade Audits (IGA). An IGA involves site surveys and collecting data about actual characteristics to establish 'current status' or 'baseline energy' of the building being considered. This process helps an ESCO identify the distribution of energy use within the building by end use and identify potential areas of improvement. Subsequent to an IGA, an ESCO evaluates candidate

ECMs by engineering analysis methods established by International Performance Measurement and Verification Protocol (IPMVP, 2010). The protocol offers several alternative methods for estimating energy savings from ECMs for a building, but all of them follow a deterministic approach. They compute an absolute value of energy-savings from a set of ECMs without quantifying any risks associated with the investments.

In reality, ESCOs quantify risks associated with ECMs, but do so on the basis of an experts' knowledge and prior beliefs. As an acceptable rule of thumb, an experts' subjective judgment states guaranteed savings to be between 60% and 70% of the deterministic energy-saving estimate (Hansen, 2004). This rule of thumb is applied uniformly to all ECMs. It does not reflect risks pertaining to an individual technology or energy saving measure. Furthermore, the set of ECMs considered by ESCOs tend to be limited to those with proven track records in yielding energy savings. Indeed, the most commonly implemented ECMs by the ESCOs are high-efficiency lighting systems or lighting controls (Goldman, 2002).

So far, despite the increasing recognition of the importance of risk analysis in performance contracts, the deterministic approach in current practice ignores uncertainty quantification in the retrofit analysis process. Moreover, no formal methodology exists to introduce risk analysis in the decision-making process.

1.4 New Methodologies for Retrofit Analysis

The objective of this thesis is to develop a new retrofit analysis framework that can support large-scale retrofit decisions under uncertainty. The new framework introduces three main features, and is applicable for retrofitting individual buildings, and also scalable to large portfolio of buildings. The three main features are as follows:

- ***Normative Energy Models***: The proposed retrofit analysis is based on using normative energy models in lieu of transient building simulation. Normative

energy models define energy flows in a building with a relatively small set of parameters. Hence, they greatly alleviate burdens in data collection, modeling, and computational time. In addition, they do not require modeling expertise since they follow a set of modeling rules that produce a standard measure for energy performance.

- ***Bayesian Calibration of the Normative Energy Model***: A Bayesian approach is used for calibrating uncertain parameters in the normative model and quantifying uncertainties in the parameters. This process improves the reliability of the baseline energy model and naturally enables probabilistic analysis of ECMs.
- ***Probabilistic Analysis****:* Evaluation of ECMs is based on translating the probabilistic outputs from the Bayesian model into risks of underperformance associated with ECMs. Hence, the probabilistic analysis can support risk-conscious decisions that reflect the decision-makers' willingness to accept a certain level of risk in their investments.

1.5 Research Hypotheses and Methodology

The thesis illustrates the proposed retrofit analysis framework for auditing applications through two case studies. The case studies are also used to verify the following two hypotheses:

- Hypothesis 1: Normative models supported with Bayesian calibration can adequately support retrofit decisions without compromising the degree of confidence in decisions.
- Hypothesis 2: Normative models that undergo Bayesian calibration can explicitly provide knowledge about performance uncertainty, and support the choice of rational decisions according to decision-makers' objectives.

The first hypothesis tests the feasibility of the proposed methodology, especially regarding the appropriate model granularity level for energy retrofits. We hypothesize that retrofit analysis at the whole building level does not require transient simulation models and normative models, if calibrated correctly, are adequate for evaluating energy-savings potential of retrofit options. This hypothesis is examined by comparing outcomes from the calibrated normative model with those from a similarly calibrated transient simulation model.

The second hypothesis examines the importance of uncertainty analysis in retrofit decision-making. We hypothesize that probabilistic analysis based on the Bayesian calibrated models can adequately support rational decisions according to decision-makers' risk attitude in retrofit projects. The thesis does not claim that the probabilistic analysis always leads to better decisions. Instead, the thesis illustrates how quantification of risks can potentially influence the choice of ECMs. This is done by comparing decisions derived from the risk analysis with those derived by following the standard practice of deterministic analysis.

1.6 Organization of Thesis

This thesis is outlined as follows;

- Chapter 1 has presented motivations for large-scale retrofits and quantitative risk analysis in energy retrofit decisions, and proposed a new framework that can support large-scale retrofit decisions with risk analysis.
- Chapter 2 describes the limitations of existing modeling and calibration methods, details the three new features of the proposed retrofit framework, and outlines how the two hypotheses of this thesis were tested.
- Chapter 3 covers uncertainty quantification in the context of energy analysis of buildings. It presents a process for quantifying uncertain parameters in building energy models.

- Chapter 4 presents two case studies for demonstrating the feasibility of the proposed retrofit analysis framework in energy retrofit applications (hypothesis 1).
- Chapter 5 shows the limitations of current methods used for accounting uncertainty and risk in energy-savings contracts. It further demonstrates how the proposed framework is more appropriate to support risk-conscious decision-makings (hypothesis 2).
- Chapter 6 summarizes the thesis with conclusions and suggestions for future research.

CHAPTER 2 A NEW FRAMEWORK FOR RETROFIT ANALYSIS

2.1 Limitations on Current Methods for Retrofit Analysis

Energy retrofit projects follow the International Performance Measurement and Verification Protocol for the determination of energy-savings from energy conservation measures (ECMs) (Hansen, 2004; IPMVP, 2010). ASHRAE Guideline 14-2002 (ASHRAE, 2002) also provides guidelines and calculation methods for retrofit analysis. Both these guidelines recommend deriving energy-savings by subtracting projected (calculated) energy use during the post-retrofit period from baseline energy use during the pre-retrofit period. It is recommended that energy use is normalized to reflect energy-savings solely due to ECMs and excluding the effects of other factors such as weather conditions and changes in building usage patterns.

For the estimation of energy savings, the guidelines provide three methods: (1) retrofit isolation, (2) whole building metering, and (3) calibrated simulation. The retrofit isolation method evaluates the savings from an upgraded building component (e.g., boiler, lighting system) by metering its energy efficiency during pre-retrofit and post-retrofit periods. The whole-building metering method is based on monitoring and comparing total energy consumption of a building during pre-retrofit and post-retrofit periods. Since these two methods rely on measurement data for energy-saving estimation, they provide information only after retrofit options are implemented in a building, and cannot thus evaluate retrofit options during the decision-making stage. The calibrated simulation method involves the use of energy simulation models; a simulation model is calibrated based on pre-retrofit data to represent actual building behavior in use and

predict energy-savings of all considered retrofit options. Hence, among the three methods, only the calibrated simulation method can serve as a method to support retrofit decision-making by projecting energy-saving impacts from retrofit options.

Energy simulation models thus play a key role in computing potential energy savings from retrofits. In order to reliably predict energy-savings from energy conservation measures (ECMs), a simulation model should represent a building as operated; that is, models should capture building systems as-installed, as-operated, and as-used. For reliable predictions, calibration of energy models has hence been emphasized (Ahmad, 2006; Reddy, 2006; Yoon, 2003). Calibration requires building audits and monitored energy consumption. The audits help determine observable model parameters. Then, monitored energy consumption enables tuning of unobservable model parameters so the baseline model represents the actual building accurately. If the baseline model can generate outcomes that closely match monitored energy consumption of a building, then it is more likely to predict reliable estimates of energy-savings from planned retrofit options for that building. This is widely an accepted approach for analyzing existing buildings (Pan, 2007; Zhu, 2006; Reddy, 2005; Yoon, 2003; Pedrini, 2002).

ASHRAE Guideline 14-2002 provides a standard procedure for the whole-building calibrated simulation approach (ASHRAE, 2002). First, a modeler should plan the calibration exercise by specifying a simulation software, the unit of monitored data for calibration (i.e., monthly, hourly), and acceptable tolerances for model validation. Second, one audits the building (e.g., building dimensions, construction specifications, system nameplates information, occupancy and operation schedules, and whole-building utility data). Third, based on collected information, one builds a simulation model of the building. In this step, one needs to make assumptions to represent the actual building reasonably: reducing the number of zones, simplifying HVAC systems, and defining

operation and occupancy schedules. Fourth, one compares simulation model outcomes to measured data, and refines the model until discrepancy between predicted energy uses and measured energy uses satisfies acceptable tolerances. ASHRAE Guideline 14 stipulates acceptable tolerances in terms of statistical indices: Normalized Mean Bias Error (*NMBE*) and Coefficient of Variation of Root Mean Square Error (*CVRMSE*). *NMBE* and *CVRMSE* should be within 5% and 15% with the use of monthly data for a model to be deemed valid. After calibration, one obtains a baseline model which can be applied to evaluate a given set of ECMs.

In general, the calibration procedure can be summarized with two steps: (1) *operational adjustments* and (2) *parameter estimation*. The term *operational adjustments* refers to the process of auditing a building to determine appropriate values for the observable parameters of a building simulation model. It typically includes site visits, interviews with building managers, field measurements to determine physical properties of the building, occupancy patterns, plug-in loads, and control settings. This is an important part of the calibration process since actual building operation often deviates from specifications assumed and documented during design and construction. The next step *parameter estimation* determines appropriate values for non-observable simulation parameters. Most simulation exercises on retrofit analysis employ a heuristic method for the parameter estimation process. For example, Pedrini (2002) manually calibrated internal loads, equipment operation, and occupancy schedules based on the building's monitored energy consumption and metering power demand of specific equipment in the building. Pan (2007) adjusted values of infiltration rates in the energy model until discrepancy between simulated and monitored energy was reasonably small. Yoon (2003) uses a stepwise calibration procedure for energy simulation models: (a) building a base case model, (b) analyzing differences between simulations and measured data through scatter plots, (c) tuning model parameters for internal loads with measured energy uses

during intermittent seasons, (d) refining internal gain levels and operating schedules through additional site visits, (e) tuning model parameters for HVAC systems with measured energy uses during heating and cooling seasons, and (f) validating a calibrated model. In these studies, each step involves experts' manual interventions for selection of calibration parameters and their values under test.

The quality of calibrated models thus relies heavily on the subjective judgment of experts. Moreover, current methods result in one single set of parameter values that results in a good fit between monitored and computed energy consumptions. In the parameter estimation process, uncertainties regarding parameter values are left un-quantified although they always exist in any model. Indeed, even after the most rigorous calibration, the model cannot perfectly represent the reality since it is still the abstraction of the reality. Hence, without quantifying uncertainties in the calibrated model, one cannot be aware of the reliability of model outcomes, and accordingly cannot evaluate relative cost-benefits of different retrofit options with confidence.

While employing expert-driven deterministic methods for calibration, the protocols have stipulated transient simulation models as the standard modeling approach (IPMVP, 2010; ASHRAE, 2002). They stipulate that energy analysis should be based on commercially available computer simulation models that compute dynamic energy consumption with the use of hourly weather data. No doubt, several commercial transient simulation software have been well-used over the last decade (e.g., eQuest, Energyplus, Blast, Trnsys), and they have earned the confidence of the simulation user community in the industry. These models are regarded as high-fidelity models that accurately approximate the actual building behavior if all parameter values are correct.

Transient simulation models emulate performances of systems in a building by solving the full set of dynamic heat balance equations using numerical methods. They discretize the whole building and systems into nodes that are connected by elements.

Each element corresponds to one heat transfer phenomenon, and all elements linked to each node are translated into one dynamic heat balance equation. The entire nodal network with all dynamic heat balance equations is solved simultaneously at each time step during the entire simulation period. Transient simulation models can thus be used to model a building and its control systems to a high degree of detail, which is quite beneficial when detailed design and sizing of specific systems need to be evaluated within the context of overall energy consumption of the building. However, this level of detail is often not necessary to compare cost-benefits of competitive retrofit technologies at the macro level, and it tends to burden the modeling process with excessive details. Due to the high cost of the model, the use of transient simulation models is generally prohibitive for large-scale retrofit analysis for a portfolio of buildings. However, because the protocols dictate the use of transient simulation models for retrofit analysis, other types of simplified methods have not been investigated as potentially feasible.

2.2 Main Features of the Proposed Retrofit Analysis Framework

2.2.1 Normative Energy Model

This thesis proposes that normative building energy models can adequately support retrofit analysis with the added advantage of being feasible for evaluating a large-portfolio of buildings. Normative models are quasi-steady state models designed to calculate the energy consumption by main end-uses in a building. They approximate energy flows in a building at the macro level with a simplified description of a building. A well-accepted normative method is defined in the CEN-ISO standards for energy performance calculation (ISO 13790, 2008; CEN, prEN 15203/15315, 2006). The normative model used in this thesis is the Energy Performance Standard Calculation

Toolkit (EPSCT) developed by Georgia Institute of Technology based on the CEN-ISO standards (Lee, 2011).

CEN-ISO standards are initially developed to evaluate energy performances of buildings in a standardized way, in particular to support design benchmarking. They were developed under the Energy Performance of Buildings Directive (EPBD) as the standardized calculation methodology to benchmark new buildings against a reference case for building energy certification. The CEN-ISO standards provide a set of modeling rules that produce a standard measure for energy performance hence assuring the objectivity of model outcomes. More importantly, the standards define a calculation model with normatively defined parameters that capture all the major characteristics of a building and its components. As a result, the standards can make the modeling process much faster by enormously reducing the level of information required from building audits. The level of information required from building audits is much less. This is an extremely useful advantage since gathering detailed specifications can be extremely time-consuming, if not impossible. Moreover, modeling effort and computational run-time is significantly reduced.

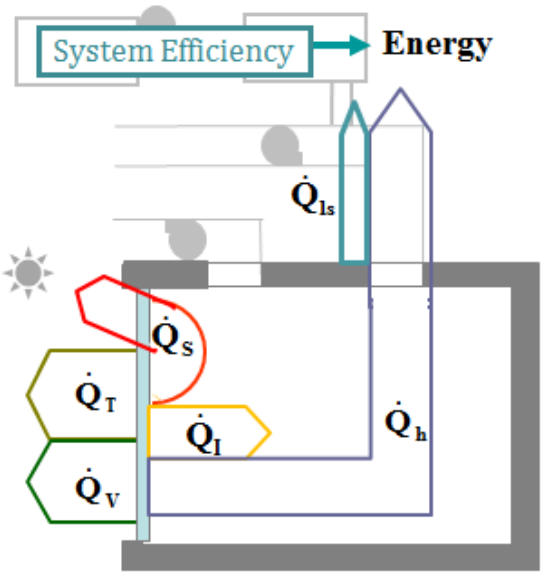


Figure 2.1 Schematic of the normative energy model

Figure 2.1 illustrates approximations of energy flows in the normative energy model for the energy performance calculation. The calculations of heat gains and losses are aggregated by transmittance, ventilation, solar radiation, and internal gains at the boundaries of the building envelope. Equation 2.1 calculates transmittance heat transfer Q_T as a function of the overall transmittance heat transfer coefficient H_{tr}, the set-point temperature $\theta_{int,set}$, monthly-average exterior temperature θ_e, and the duration of the calculation step (month) t. The overall transmittance coefficient refers to an area-weighted U value of the entire building envelope. Equation 2.2 computes ventilation heat transfer Q_V from the difference between exterior and interior temperature, the time duration, and overall ventilation heat transfer coefficient H_{ve}. The coefficient is the sum of coefficients, each of which corresponds to outside airflow rates due to infiltration, natural ventilation, and mechanical ventilation.

For transmittance heat losses: $Q_T = H_{tr}\left(\theta_{int,set} - \theta_e\right) \cdot t$ (2.1)

For ventilation heat losses: $Q_V = H_{ve}\left(\theta_{int,set} - \theta_e\right) \cdot t$ (2.2)

For solar heat gains: $Q_S = \left(\sum_k (F_{sh,k}\, A_{sol,k} I_{sol,k} - F_{r,k}\Phi_{r,k})\right) \cdot t$ (2.3)

For internal heat gains: $Q_I = (f_{oc} q_{oc} A_{oc} + f_{ap} q_{ap} A_{ap} + f_{li} q_{li} A_{li}) \cdot t$ (2.4)

Equation 2.3 computes solar heat gains that arise from shortwave and longwave radiations on the building fabric. The first part in the equation calculates the amount of global solar radiation as a function of envelope shading reduction factor F_{sh}, effective solar collecting area A_{sol}, and monthly-average global solar irradiance per area for each orientation. The second part computes the amount of longwave radiation from sky view factor F_r and heat flow due to thermal radiation Φ_r. Equation 2.4 calculates internal heat gains from occupants, appliances, and lighting devices. Heat gains from each type of internal heat sources are determined in terms of the fraction of the time heat sources produce heat f, heat production rate from sources per area q, and the total floor area A.

From these aggregated heat gains and losses, the normative model calculates heating and cooling needs with the use of utilization factors that approximate thermal inertia effects due to the building's thermal capacity, following Equations 2.5 and 2.6. The utilization factor is derived by an empirically driven equation in terms of the gain-to-loss ratio and the time constant of a building. The time constant is also empirically defined as the entire envelope heat capacity divided by heat loss coefficient ($H_{tr} + H_{ve}$). In the equations, Q_{ht} is the total heat loss due to transmittance and ventilation, Q_{gn} is the total heat gains due to solar and internal gains, and η_H and η_C are a utilization factor for heating and cooling respectively.

$$\text{For heating need: } Q_{H,nd} = Q_{ht} - \eta_H Q_{gn} \tag{2.5}$$

$$\text{For cooling need: } Q_{C,nd} = Q_{gn} - \eta_C Q_{ht} \tag{2.6}$$

Following the calculation of various energy demands of a building in a similar manner, the model utilizes overall efficiency of the energy generation and the distribution system to calculate the energy consumption for heating and cooling. Equations 2.7 and 2.8 derive energy consumption from energy demands with the two macro-level parameters; $\eta_{H,sys}$ and $\eta_{C,sys}$ refer to the generation system seasonal efficiency, and $Q_{h,ls,sys}$ and $Q_{c,ls,sys}$ refer to the distribution system losses for heating and cooling respectively. The model defines a simplified procedure to derive the seasonal efficiency value with the use of normative factors that reflect the effects of part loads and control settings. In the same manner, the model also determines the total energy losses during delivery in terms of a weight factor explaining losses in pipes or ducts and a waste factor expressing losses due to simultaneous heating and cooling. In short, the normative model derives overall system performance from the brief description about system types, configuration, and control settings instead of the detailed level of HVAC system modeling typically required in transient simulation models.

For heating energy consumption: $E_{heat} = \frac{\sum_i (Q_{H,nd} + Q_{h,ls,sys})}{\eta_{H,sys}}$ (2.7)

For cooling energy consumption: $E_{cool} = \frac{\sum_i (Q_{C,nd} + Q_{c,ls,sys})}{\eta_{C,sys}}$ (2.8)

Regarding the other end-use energy consumptions, the normative model also defines normative factors that account for systems efficiency, their configuration, and their control settings. For the lighting energy consumption, Equation 2.9 requires the installed lighting power value P_n and normative factors that take into account day-lighting utilization (F_D), occupancy sensor (F_O), and dimming control (F_c). Also, the model provides a normative value for each type of building to derive the domestic hot water (DHW) demand $Q_{nd,dhw}$, and utilize the generation system efficiency $\eta_{gen,dhw}$ and the distribution system efficiency $\eta_{dis,dhw}$ to calculate the DHW energy consumption. Equation 2.11 calculates the fan energy consumption as considering ventilation system efficiency and system operation settings. e_{vent} refers to the system efficiency by indicating specific electricity consumption defined per ventilation type. The system operation settings are parameterized by forced airflow rates ($\dot{V}_H$ and $\dot{V}_C$ for heating and cooling) and operation schedules: f_{vent} is the fraction of the time the system is on and $f_H$and f_C are monthly variations in the time fraction for heating and cooling. For the pump energy consumption, the model utilizes correction factors that express the effects of system design, control, and operation features particularly for water-based heating systems: $P_{hydr,des}$ is the designed hydraulic power, β_{dis} is the mean part load normatively defined, and f_{NET}, f_{HB}, $f_{G,PM}$ are correction factors for hydraulic networks, hydraulic balance, and integrated pump management (EN 15136-2-3, 2007).

For lighting energy consumption: $E_{light} = \frac{\sum_i (P_n F_c (t_D F_O F_D + t_N F_O))}{1000}$ (2.9)

For DHW energy consumption: $E_{dhw} = \frac{Q_{nd,dhw}/\eta_{dis,dhw}}{\eta_{gen,dhw}}$ (2.10)

For fan energy consumption: $E_{fan} = e_{vent}(f_H \dot{V}_H + f_C \dot{V}_C) f_{vent} t$ (2.11)

For pump energy consumption: $E_{pump} = \frac{P_{hydr,des}}{1000} \beta_{dis} t_{op} f_{NET} f_{HB} f_{G,PM} e_{dis}$ (2.12)

Normative models approximately represent energy performances of the aggregate-level building systems with a small number of macro-level inputs based on a simplified description of a building and its systems. Accordingly, normative models can drastically improve the cost-effectiveness of the modeling process. Furthermore, the normative nature of the models can make the modeling process transparent as it implies that no modeling expertise is needed. Hence, normative models promise to be good candidates for building audits as they replace the complex and expensive transient simulation models. Although they are widely used and verified for design benchmarking, their suitability for auditing purposes is limited and untested. Since normative models are based on a relatively small set of macro-level parameters, testing is needed to determine whether macro-level parameters for sub-system characteristics are able to capture interactive, cumulative effects of its components through calibration. Therefore, this thesis will examine whether normative models can be suitably calibrated for retrofit analysis of buildings, and calibrated normative models can support retrofit decisions without compromising the degree-of-confidence in decisions.

2.2.2 *Bayesian Calibration of the Normative Energy Model*

This thesis proposes a Bayesian approach as the core of calibration as it could quantify uncertainty in the estimates of calibration parameters in a form of probability distributions. The Bayesian paradigm treats a probability as a numerical estimate of the degree-of-belief in a hypothesis. The Bayesian paradigm updates our prior belief on true values of uncertain parameters in a computer model given monitored data on building performance. Bayesian calibration is based on Bayes' theorem expressed in Equation 2.13; where *p(θ)* are prior distributions assigned for uncertain parameters based on expert

knowledge from a pool of sources (e.g., experiments, surveys, industry standards, etc); $p(y|\theta)$ is a likelihood function that measures how closely computer results with testing parameter values match observations. Prior distributions are updated using observations in which the likelihood of obtaining observations from the computer model drives the updating process. As a result, Bayesian calibration results in plausible distributions of calibration parameters, referred to as posterior distributions $p(\theta|y)$.

$$p(\theta|y) \propto p(\theta) \times p(y|\theta) \tag{2.13}$$

Bayesian calibration has been widely adopted in environment and earth sciences to enhance reliability in model predictions. Models emulate the complex dynamics of systems with a large number of parameters, many of which are generally unknown due to limited empirical data to estimate parameter values. However, the reliability of model predictions depends on not only model fidelity but also accuracy of input values. Therefore, in order to estimate parameter values with measures of uncertainty from observed data on model outputs, Bayesian approach has been employed for ecological models (van Oijen, 2005), hydrologic models (Qian, 2005; Liu, 2008), and atmospheric models (Guillas, 2009). Notwithstanding the popularity and benefits of Bayesian techniques, building energy models have been calibrated only in a deterministic manner without accounting for uncertainties. Hence, this thesis attempts to extend Bayesian techniques to the domain of building energy simulations so that calibrated energy models can explicitly project their parameter uncertainty in model outputs.

The Bayesian calibration module requires the three major steps: (1) specification of prior probability distributions for calibration parameters, (2) formulation of the likelihood function, and (3) Markov Chain Monte Carlo method for posterior simulation. For the likelihood function, the thesis follows the Kennedy and O'Hagan formulation of Bayesian calibration developed by (Kennedy and O'Hagan, 2001). The statistical formula captures three types of uncertainties: (1) parameter uncertainty in the building energy

model, (2) discrepancy between the model and the true behavior of the building, and (3) observation errors. As accounting for these uncertainties, Equation 2.14 defines the relationship between model outputs and observations.

$$y(x) = \eta(x, \theta) + \delta(x) + \varepsilon(x) \tag{2.14}$$

$y(x)$ denotes observations at known conditions x (e.g., external climate conditions, occupancy schedules, etc); $\eta(x, \theta)$ denotes building energy model outcomes at known conditions x and calibration parameters θ. The formula also introduces an additional stochastic term $\delta(x)$ that captures the discrepancy between the model and the true physical behavior. In fact, building energy models are based on approximations of the heat transfer processes occurring in a building, and they may not therefore capture the actual consumption of the building even with true values of the calibration parameters. The discrepancy term prevents over-estimation of calibration values, and indicates where the energy model falls short. The formula also includes a stochastic term $\varepsilon(x)$ that expresses errors in collecting observations.

The Kennedy and O'Hagan formulation requires three sets of data as input: (1) monthly utility data as observations $y(x)$, (2) prior probability density functions of calibration parameters $p(\theta)$, and (3) model outcomes from exploring the space of calibration parameters $\eta(x, \theta)$. Given these input, both the model outputs $\eta(x, \theta)$ and the discrepancy term $\delta(x)$ are modeled as Gaussian processes (Rasmussen, 2006). A Gaussian process is a generalization of a multivariate normal vector to the case where the index set is infinite. The energy model output under specific known conditions and for a chosen set of calibration parameters, despite being deterministic, is assumed to follow a normal distribution. Jointly, with several outputs under different sets of known conditions, they form a multivariate normal vector with a specific covariance structure. With such distributional assumptions, we can obtain probabilistic distributions of outputs at unknown set of conditions and parameter values from the given dataset of input and

output values. Ideally, one wants to evaluate the energy model outputs over a very dense set of conditions and parameter values. Since this exhaustive exploration is infeasible, one uses design of experiments to explore the parameter space as much as possible with a manageable computational burden.

A Gaussian process model is specified by a mean function and a covariance function. A mean function is a matrix of mean output values for the given set of input values. A covariance function is a matrix, each element of which indicates proximity between the two sets of input values with respect to their outputs. Equation 2.15 expresses the (i, j)-th element of the covariance matrix for the Gaussian process model of $\eta(x, \theta)$. The covariance function contains two hyper-parameters to control the predictive power of a Gaussian process model; λ controls the precision of a Gaussian process model and β controls correlation strength in each input parameter. Equation 2.16 defines the covariance function for the Gaussian process model of $\delta(x)$. It should be noted is that $\delta(x)$ depends only on known parameters x whereas $\eta(x, \theta)$ depends on both known parameters x and calibration parameters θ. In addition to the two Gaussian process models, observation errors are defined as a Gaussian distribution $N(0, \sigma^2)$, assuming that they are normally distributed and uncorrelated.

$$\Sigma_{\eta(i,j)} = \frac{1}{\lambda_\eta} exp\left(-\sum_{k=1}^{p} \beta_{\eta,k}(x_{ik} - x_{jk})^2 - \sum_{k'=1}^{q} \beta_{\eta,p+k'}(\theta_{ik'} - \theta_{jk'})^2\right) \quad (2.15)$$

$$\Sigma_{\delta(i,j)} = \frac{1}{\lambda_\delta} exp\left(-\sum_{k=1}^{p} \beta_{\delta,k}(x_{ik} - x_{jk})^2\right) \quad (2.16)$$

This Gaussian process formulation enables us to compute likelihoods of observations $p(y|\theta)$ given model parameters. In the Kennedy and O'Hagan framework, the joint vector for the likelihood function consists of observations and computer results, denoted as $z = (y^T, \eta^T)^T$. The vector is modeled as a Gaussian process model with mean function μ_z and covariance function Σ_z. Equation 2.17 expresses the covariance function

for the data-likelihood while encompassing parameter uncertainty, model discrepancy, and observation errors. The data-likelihood function is defined in Equation 2.18.

$$\Sigma_z = \Sigma_\eta + \begin{pmatrix} \Sigma_\delta & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \Sigma_\varepsilon & 0 \\ 0 & 0 \end{pmatrix} \quad (2.17)$$

$$f(z) \propto |\Sigma_z|^{-\frac{1}{2}} \times \exp\left(-\frac{1}{2}(z-\mu_z)^T {\Sigma_z}^{-1}(z-\mu_z)\right) \quad (2.18)$$

Following Bayes' theorem, we attain posterior distributions by multiplying the likelihood function and all prior density functions for uncertain parameters. Uncertain parameters include not only calibration parameters θ but also hyper-parameters for the Gaussian process models. Prior distributions for the hyper-parameters are assigned such that the building energy model can explain most of the variation in the observations with a relatively smaller bias and even smaller observation errors. During the posterior simulation, the calibration module corrects the predictive power of the energy model by updating prior distributions for the hyper-parameters while simultaneously updating prior distributions for calibration parameters.

$$p(\theta, \lambda_\eta, \beta_\eta, \lambda_\delta, \beta_\delta, \sigma|\, z) \propto p(z|\theta, \lambda_\eta, \beta_\eta, \lambda_\delta, \beta_\delta, \sigma) p(\theta) p(\lambda_\eta) p(\beta_\eta) p(\lambda_\delta) p(\beta_\delta) p(\sigma) \quad (2.19)$$

A Markov Chain Monte Carlo (MCMC) method, specifically the Metropolis-Hastings algorithm, is used to draw from the joint multivariate posterior distribution. MCMC method generates a random walk through the parameter space such that the collection of sample points can approximate theoretical posterior density functions. Similar to Markov chains, the method draws a proposed point based on the current point in an iterative manner, and accepts the proposed point when it satisfies an acceptance criterion. The Metropolis-Hastings algorithm defines the criterion by the ratio of a posterior density at the proposed point to that at the current point (Gelman, 2004). The algorithm accepts the proposed point with the probability equal to $min(ratio, 1)$ only when the ratio is larger than a number randomly generated from the uniform distribution

$U(0,1)$. As a result, we obtain a sample of accepted values $[\theta^{(0)},\cdots,\theta^{(n)}]$ as posterior distributions.

Figure 2.2 summarizes the entire calibration process. The procedure starts with a two-step pre-process targeted to objectively quantify uncertainty in model parameters and select calibration parameters. The first step is to quantify uncertainty in all uncertain parameters in the energy model based on expert knowledge collected from site surveys, industry reports, standards, and technical papers. The second step is to objectively select a smaller set of dominant parameters for calibration through a parameter screening technique called the Morris method. The Morris method enables the ranking of uncertain parameters with respect to the effects of their uncertainty on energy consumption.

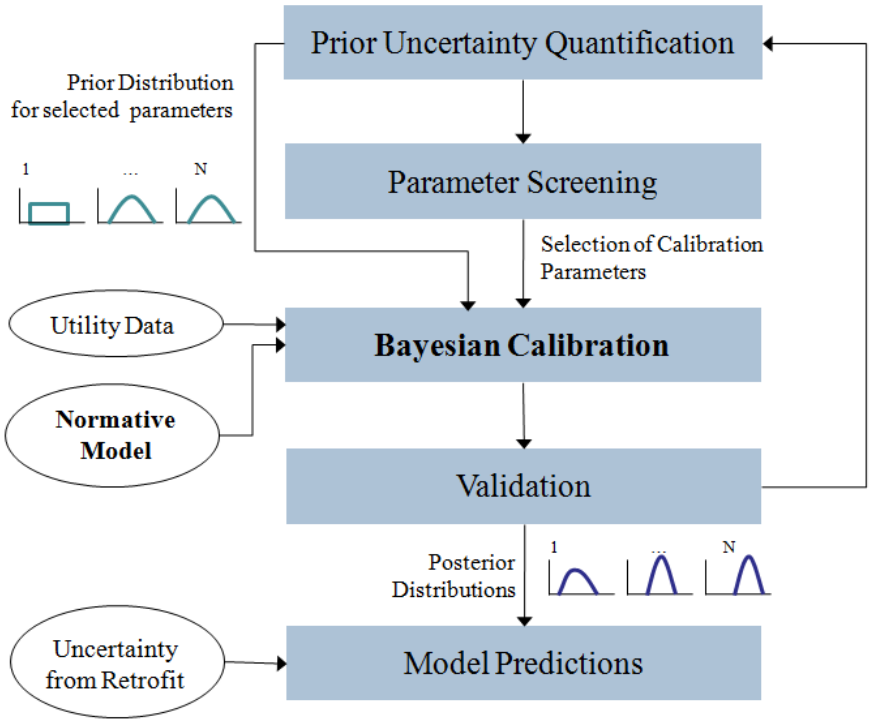


Figure 2.2 Systematic procedure for retrofit analysis

The Bayesian calibration module calibrates selected parameters given prior distributions for the parameters, monthly utility data, and the energy model of a building

being investigated. With the resulting posterior distributions, the calibrated model is validated by posterior predictive checking that measures agreements between the utility data and predicted model outcomes. This validation step employs statistical measures that quantify the fit between the monitored and the predicted energy uses. Last, the validated model propagates uncertainty quantified by the calibration and additional uncertainty from evaluating ECMs to compute probabilistic outcomes of their potential energy-savings.

2.2.3 *Probabilistic Analysis*

The new framework is designed for probabilistic analysis that defines uncertainty and translates effects of uncertainty on outcomes of interest in order to support retrofit decision-making under uncertainty. Figure 2.3 shows the overall probabilistic analysis process in the retrofit decision-making stage. The first step is uncertainty quantification that identifies sources of uncertainty and quantifies uncertainty in the identified sources in a form of statistical distributions. This step is extensively investigated in Chapter 3. The second step propagates uncertainty through the normative energy model to obtain a probability distribution of outcomes. The stage of uncertainty propagation requires a statistical technique that efficiently draws samples from statistical distributions. Sampling techniques are explained in detail in Chapter 3. The resulting probabilistic outcomes can be translated into any measure that reflects decision-makers' rational objectives and their risk attitude for rational decision-making.

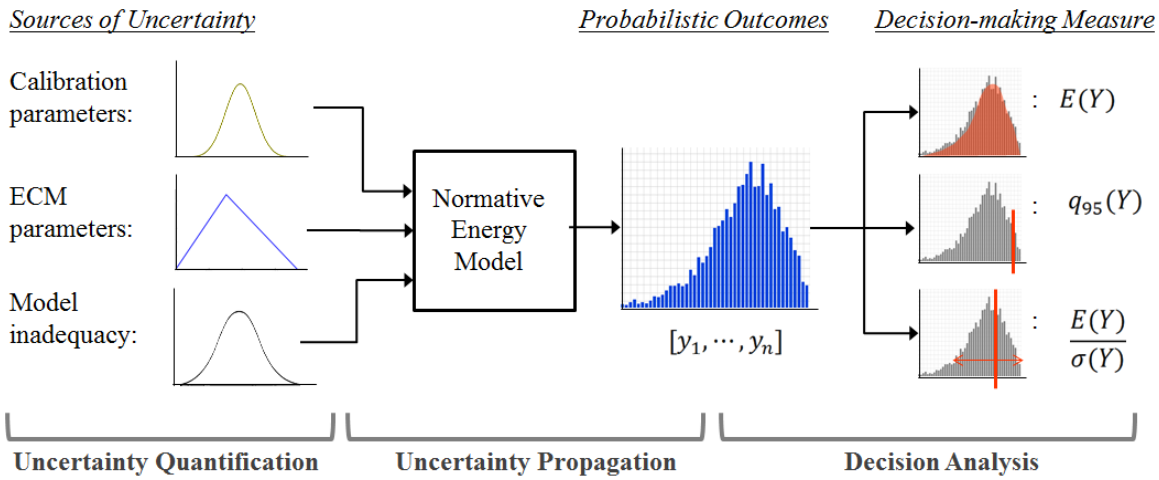


Figure 2.3 Probabilistic analysis process in the retrofit decision-making

2.3 Evaluation of the Hypotheses

The major hypothesis in the thesis is that retrofit analysis at the whole building level does not require the most advanced simulation models and normative models with added calibration can adequately predict retrofit energy-savings for rational retrofit decision-making. The thesis evaluates the feasibility of normative models with respect to their intended uses (retrofit purposes). When models are deployed for problem-solving or decision-making, their credibility to compute "correct" results has been always a major concern to users. However, it is too time-consuming and expensive to build a model that is absolutely valid over the complete domain of its applications. Instead, users should determine the validity of a model for a specific application (Sargent, 2005). Figure 2.4 depicts the relationships between the cost of validating a model, the level of the model confidence, and the value of the model to the user.

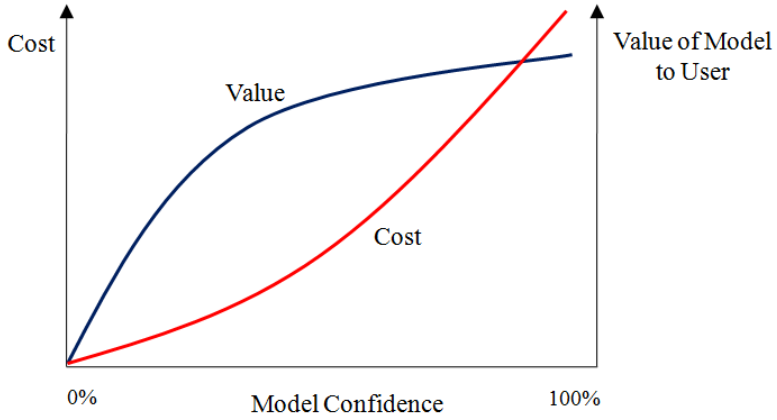


Figure 2.4 Relationships between a model confidence, its cost, and its value to user (from Sargent, 2005)

These relationships can be equally applicable to all modeling exercises. Model confidence always increases at the expense of the modeling cost and time. For instance, transient simulation models require detailed description of the building. At the same time, they significantly increase the modeling cost (in terms of time and expertise). However, increase in the modeling cost may not necessarily result in adding value to the analysis process. Indeed, if a normative model is well calibrated, it may be as accurate as a similarly calibrated transient simulation model and equally applicable (unless specific dynamic effects associated with equipment control or operation need to be evaluated).

The thesis follows the same reasoning to evaluate a model resolution level appropriate for energy retrofit applications. The thesis evaluates whether the normative model has sufficient model fidelity to correctly evaluate ECMs and lead to reasonable retrofit decision-making in comparison to transient simulation models. The feasibility of the normative model is inspected by pair-wise comparisons between the normative model and the transient simulation model. The two models are compared under the following criteria:

- Criterion 1: Accuracy of Calibrated Models
- Criterion 2: Accuracy of Predictions
- Criterion 3: Effects of Prediction Accuracy on Decisions

The first criterion evaluates the accuracy of the calibrated normative model in terms of how it replicates the historical utility data of the building being modeled. This criterion can evaluate only the reliability of the baseline model, but cannot tell that the calibrated normative model can reliably predict energy-savings of ECMs. The normative model can potentially compromise the degree of confidence in predictions because it approximates the mathematical representation of physical systems. Therefore, the second criterion is introduced to test the prediction accuracy of the normative model when supported by Bayesian calibration. If the normative model generates the same predictions as the transient simulation model, the normative model will lead to the same retrofit decisions as the transient simulation model. However, even if their outputs are not the same, the normative model can still be a good candidate if it does not bias decisions in the retrofit analysis process. Hence, the third criterion examines the effects of prediction accuracy on decisions by comparing decisions supported by the two models in the context of plausible decision-making scenarios.

2.3.1 Criterion 1: Accuracy of Calibrated Models

This criterion evaluates the accuracy of calibrated models with respect to agreement between model predictions and monitored data. The criterion uses standard validation metrics such as the index of agreement d and coefficient of variation of the root mean square error $CVRMSE$ for comparing the outputs from the calibrated model with observed values of energy consumption. d and $CVRMSE$ are expressed in Equations

2.20 and 2.21; P_i denotes a predicted energy use for period i, O_i an observed energy use for period i, and $\overline{O}$ the mean of all observed energy uses.

$$d = 1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(|P_i - \overline{O}| + |O_i - \overline{O}|)^2} \quad (2.20)$$

$$CVRMSE = \frac{\sqrt{\sum_{i=1}^{n}(O_i - P_i)^2 / n}}{\overline{O}} \quad (2.21)$$

ASHRAE Guideline 14 (ASHRAE, 2002) stipulates the tolerance limits of calibrated simulation models in terms of $CVRMSE$; $CVRMSE$ should be less than 15% with the use of monthly utility data when models are deemed valid. The index of agreement d is also widely used to assess the efficiency of models in comparison to measured data (Legates, 1999; Krause, 2005). The index of agreement ranges between 0 and 1 with higher values indicating better fit between model outcomes and observed data. Based on these tests, if the calibrated model is deemed satisfactory, it can be exercised for computing energy-saving potential of different retrofit options.

2.3.2 Criterion 2: Accuracy of Model Predictions

This criterion examines the disparity of predictions between the calibrated normative model and the calibrated transient simulation model. In the retrofit investment decision-making process, retrofit interventions are typically evaluated by cost-effectiveness such as cost/benefit ratio and simple payback time (Goldman, 2002). We employ Simple Payback Time (SPT), defined as investment costs divided by annual energy-saving costs, for decision-makings. The criterion compares SPT predictions of candidate ECMs derived by the calibrated normative model with those derived by the calibrated transient model.

The criterion employs a two-sample Kolmogorov-Smirnov test (K-S test) to quantitatively evaluate whether probabilistic outcomes projected by the two models are identical. The two-sample K-S test checks the location and the shape of the two samples to confirm whether the two samples come from the same cumulative distribution. The test hypothesizes that the two samples are from the same distribution, denoted as H=0, and rejects the hypothesis when they are different at 5% significant level. In addition, the criterion qualitatively evaluates the disparity by visually comparing the two histograms of predictions.

2.3.3 *Criterion 3: Effects of Prediction Accuracy on Decisions*

This criterion evaluates if the two calibrated models derive consistent results for supporting decisions. The criterion compares the ranking of candidate ECMs driven by the two models in the context of plausible decision-making scenarios. We evaluate candidate retrofit options under the three scenarios that express different levels of decision-makers' risk consciousness. Table 2.1 summarizes the three scenarios with performance measures applied for the final evaluation of the calibrated models. Scenario 1 represents conventional practice that does not concern risks but overall performance. We use expected values of SPT for this scenario. Scenario 2 represents guaranteed savings in performance contracts commonly used in the energy service companies. The guarantee is translated into 95-quantile of the SPT distribution. Scenario 3 represents one of the existing risk measures, saving curve score proposed for actuarial pricing of retrofit projects (Mathew, 2005). The score is defined as the mean savings divided by risk, and the magnitude of risk is computed as the standard deviation of mean savings estimate. Since we use the payback time inverse to mean savings, we use 1/SPT instead of mean savings. If the two calibrated models lead to the same ranking of testing ECMs under

these three scenarios, it can be concluded that the normative model with Bayesian calibration can adequately support retrofit decision-makings under uncertainty.

Table 2.1 Decision-making scenarios with probabilistic measures

	Description	Measure
Scenario 1	Conventional practice	$E(SPT)$
Scenario 2	Guaranteed savings (in ESCO projects)	$q_{95}(SPT)$
Scenario 3	Saving curve score (existing risk measure)	$\frac{1}{E(SPT) \times \sigma(SPT)}$

CHAPTER 3 UNCERTAINTY IN ENERGY RETROFIT ANALYSIS

3.1 Introduction

A building energy model is typically used to predict the energy-savings of energy conservation measures (ECMs) and ultimately assist decision-makers to select the optimal mix of ECMs according to their objectives. However, this model is not capable to exactly predict how much savings ECMs will achieve due to the following reasons. First, even after the model is built based on deep energy audits, the model may not correctly correspond to existing building conditions because the audits cannot provide full information. Second, the model cannot perfectly capture the system behavior. Hence, the model may not predict actual energy consumption even with the best possible values of the model parameters. Third, performance of ECMs in reality can differ from the expected because the actual properties of the system may not be the same as those documented in standards or specifications under dynamic and stochastic building operational conditions. Owing to these uncertainties the energy model often yields energy-saving predictions that unavoidably deviate from actual energy-savings.

Uncertainty analysis helps overcome the lack of knowledge that may bias retrofit decisions by explicitly capturing the effects of incomplete knowledge on outcomes of interest. Hence, the level of rigor in uncertainty analysis depends on the quality of uncertainty quantification. Uncertainty quantification involves identifying sources of uncertainty that potentially impact the outcomes and quantifying uncertainty in a form of the probability density function. The following sections summarize sources of uncertainty in the context of energy retrofit analysis, and quantify uncertainty in the identified

sources. Quantified uncertainty is propagated through the energy model with the use of sampling methods, which results in probability distributions of model outcomes. The probabilistic outcomes can support retrofit decision-makings under uncertainty since they can be naturally translated into a single value according to decision-makers' objective and risk-awareness.

3.2 Sources of Uncertainty in Energy Retrofit Analysis

Risk in the assessment of ECMs can be categorized into two groups: (a) physical risk pertaining to the energy-savings of ECMs and (b) financial risk pertaining to the cost-effectiveness of ECMs. Table 3.1 lists sources of uncertainty that influence the physical and financial performance of candidate ECMs. Physical risk captures the risk of ECMs not resulting in expected energy savings as the result of scenario uncertainty, building physical and operational uncertainty, model inadequacy, and observation error. In addition to the physical risk, uncertainty in investment costs and utility costs impacts the financial risk of ECMs not recovering initial investment costs within an expected period.

Table 3.1 Sources of uncertainty in the evaluation of ECMs

	Category	Factors
Physical	Scenario uncertainty	- Outdoor weather conditions - Building usage / occupancy schedule
	Building physical / operational uncertainty	- Building envelope properties - Internal gains - HVAC systems - Operation and control settings
	Model inadequacy	- Modeling assumptions - Simplification in the model algorithm - Ignored phenomena in the algorithm
	Observation error	- Metered data accuracy
Financial	Investment cost	- Equipment cost - Labor cost - Discount rate
	Utility cost	- Energy source price - Utility energy contracts

The scenario refers to the external environment and the use of the building (e.g., building usage, occupancy, and operation schedules). Actual weather conditions around the building (e.g., local ambient temperature, cloud cover, local wind speed) differ from the TMY weather data used in the model that capture average weather conditions by statistically collating 30-year weather data. Also, actual in-use scenarios fluctuate from average fixed schedules used in the model. Although scenarios are inherently uncertain, we ignore deviations from "average" scenarios, and assume that scenarios are known conditions. This assumption is reasonable since we use monthly average weather data and schedules in the normative models that are much less variable across years than daily variations.

Building physical and operational uncertainty refers to parameter uncertainty in the energy model. The model parameters that fall in this category specify material thermal properties, internal gains, HVAC system properties, and their operation and control settings. The behavior of these parameters in real buildings often deviates from their specifications because the nominal conditions used for performance testing cannot capture dynamic and stochastic building operation conditions. In addition, systems degrade over their life cycle, which amplifies the magnitude of uncertainty in the system performance. Moreover, design and specification documents often lack full description of system properties, but provide information about system types. Therefore, we investigate these fundamental factors that cause uncertainty in model parameters in order to quantify parameter uncertainty in the models.

Another type of uncertainty arises from the inability of the energy models to exactly represent reality. Model inadequacy differs depending on the choice of specific energy models; a higher resolution model is known to represent the reality more accurately than a lower resolution model. Nevertheless, all energy models approximate physical heat transfer phenomena occurring in a building by abstracting complex phenomena into simplified models. In this process, some physical phenomena are ignored if they are regarded insignificant with respect to their effects on system energy performance. We quantify model inadequacy by identifying specific modules that compute the operation state of specific systems from empirically-driven equations and quantifying an uncertainty range of model coefficients in the equations.

Observation error refers to the quality of metered data used in the retrofit analysis process. Since the monitored data is used to calibrate a building energy model and validate the model, the precision of the data can influence the accuracy of the resulting calibrated model. Hence, we capture this potential error stemming from observations by accounting for observation error in the Bayesian calibration formulation. The formulation

also accounts for model inadequacy in order to avoid overestimating parameter values in the calibration process.

In addition to these physical uncertain factors, financial sources of uncertainty are an important factor that determines the cost-effectiveness of retrofit scenarios. The first category is investment costs that consist of equipment and labor costs. The investment costs for implementing an ECM are uncertain because an ECM can be realized by various costs associated with different commercial products and the actual retrofitting period that often differs from the expected period for completion. Furthermore, the value of investment costs can vary depending on the discount rate that changes over time. In addition, utility cost reduction from ECMs depend on not only their energy-savings but also energy source prices and types of utility contracts. Energy source prices constantly change due to variation in demand, commodity costs, and pricing regulation and structure (EIA, 2011).

3.3 Quantification of Uncertainty

This section focuses on quantifying parameter uncertainty in the context of normative models. Uncertainty in model parameters depends on the model granularity level. For instance, model parameters in transient simulation models describe the physical behaviour of an individual component while those in normative models describe the characteristics of systems at an aggregate level. Accordingly, uncertainty associated with different levels of model parameters should be separately investigated. Many studies have extensively investigated quantification of uncertainty in detailed simulation models (de Wit, 2001; Macdonald, 2002; Moon, 2005; Hu, 2009). Yet, uncertainty in normative model parameters has not been properly investigated.

In normative models, we quantify uncertainty in macro-level parameters by investigating a set of detailed-level parameters that a macro-level parameter accounts for

and their relationships. Hence, quantification of aggregate-level parameter uncertainty is accomplished by the three steps: (a) investigating physics-based equations that parameterize the behavior of aggregate-level parameters with a set of detailed model parameters, (b) quantifying uncertainty in the detailed model parameters from the literature review, (c) propagating quantified uncertainty through selected equations to derive a probability distribution for one aggregate-level parameter.

3.3.1 Thermophysical Properties

Thermophysical properties define the physical characteristics of construction materials used in a building fabric that impact energy demands of a building. Normative models utilize four parameters to characterize the thermal behavior of the whole construction assembly: thermal transmittance, solar absorptance, emissivity, and envelope heat capacity. Uncertainty in thermal properties of materials largely arises from variations in how they are measured in laboratories and differences in manufacturing, rather than due to differences in how a building is specifically used or constructed. Thus, we are able to use uncertainties as quantified in Macdonald (2002) for thermal properties of wall and roof materials. For impermeable materials, the standard deviation of uncertainties in thermal transmittance, density, and specific heat is 5%, 1%, and 12.25% respectively as shown in Table 3.2. For solar absorptance and emissivity, the uncertainty range quantified as standard deviation differs depending on the type of materials, which is well summarized in Table 3.3. Using the standard deviation values, we use the 95% confidence interval as the minimum and the maximum values.

Table 3.2 Uncertainty quantification for impermeable materials (MacDonald, 2002)

Category	Conductivity	Density	Specific heat
Impermeable	5%	1%	12.25%

Table 3.3 Solar absorptance and emissivity of material surfaces (MacDonald, 2002)

Category	Absorptance		Emissivity	
	Mean	Std. dev	Mean	Std. dev
Metals polished	0.32	0.07	0.05	0.01
Metals	0.56	0.12	0.24	0.06
Brick (light)	0.49	0.04	0.90	0.02
Brick (dark)	0.76	0.04	0.90	0.02
Stone (natural)	0.63	0.10	0.91	0.02
Plaster	0.40	0.03	0.09	0.02
Concrete	0.68	0.04	0.90	0.02

In addition, normative models use an aggregate parameter called effective heat capacity C, which approximates the dynamic heat storage (thermal mass) of the building envelope as a whole. We follow a calculation procedure in the CEN-ISO standard (EN ISO 13786, 2007) to calculate the cumulative effect of each component heat capacity. The standard calculates the effective heat capacity as an area-weighted function of density ρ (kg/m³) and specific heat capacity c (J/Kg.K) of building elements, starting from the internal surface up to the first insulating layer, the maximum thickness 10cm, or the middle of the wall and roof assembly, either of which comes first. We propagate uncertainty in the density and the specific heat of building elements through this calculation method to derive the uncertainty range for the effective heat capacity. For concrete buildings, the resulting value of C ranges between 160 and 275 (kJ/m².K). As the base value, we use the value recommended by the CEN-ISO standards for each construction class that are summarized in Table 3.4 (ISO 13790, 2008).

Table 3.4 Default values for the effective heat capacity (ISO 13790, 2008)

Class	C (kJ/m².K)
Very light	80
Light	110
Medium	165
Heavy	260
Very heavy	370

3.3.2 Infiltration

Infiltration is the unavoidable introduction of outside air into a building depending on the air-tightness of the building envelope and indoor /outdoor climatic conditions. The infiltration rate is one of the most unknown parameter because attaining accurate values can be possible only through fan pressurized tests on a building under consideration. The fan pressurized tests are conducted by installing a fan, often mounted in a door, to maintain a certain pressure (typically 50Pa or 75Pa) across the building envelope and measuring airflow rates induced through the fan. Equation 3.1 is used to convert measured airflow rates to predict airflow rates through the envelope at any pressure difference. Q denotes the airflow rate induced to maintain the pressure difference Δp, κ is flow coefficient, and n is pressure exponent (typically assumed as 0.65).

$$Q = \kappa(\Delta p)^n \tag{3.1}$$

Infiltration rates vary from one building to another. In energy models, they are quantified by estimating the volumetric flow rate of outside air into a building $\dot{V}$ (m³/h) or air changes per hour ACH (1/h). It is generally understood that the infiltration rate of a building is a function of its age, its construction quality, and weather conditions (pressure difference between the outside and the inside of the building). However, researchers have

not been able to identify the correlation between building age / construction type and infiltration rate after having analyzed measured air-tightness data from 139 commercial and institutional buildings (Persily, 1998; Persily, 1999).

National standards and guidelines establish recommended air permeability values that correspond to the normal and the best practice for different types of buildings. Table 3.5 lists the air permeability values for naturally ventilated office buildings suggested by ATTMA (2010) and CIBSE TM23 (2000). They suggest that the air permeability of the envelope should be in the range of 7.0 - 10.0 $m^3/m^2 \cdot h$ at 50Pa for the normal practice and in the range of 3.0 - 5.0 $m^3/m^2 \cdot h$ at 50Pa for the best practice. These values correspond to pressurization tests at specific conditions, and are translated into an average annual air change rate *ACH* using an empirically derived correction factor. Table 3.6 shows annual air change rate values corresponding to pressurized test results from (CIBSE Guide A, 2006).

Table 3.5 UK recommended infiltration rates for naturally-ventilated office buildings

Standard	Air permeability ($m^3/ m^2.h$ at 50Pa)	
	Normal	Best Practice
ATTMA, 2010	7.0	3.0
CIBSE TM23, 2000	10.0	5.0

Table 3.6 Empirical values of air infiltration rates for naturally ventilated office buildings with partial exposure (CIBSE Guide A, 2006)

Air permeability ($m^3/m^2.h$ at 50Pa)	Infiltration rate (ACH) for a given building size (1/h)			
	4 stories: 2000 m^2		6 stories: 3000 m^2	
	Peak	Average	Peak	Average
20.0 (leaky)	0.75	0.55	0.75	0.55
10.0 (Part L, 2002)	0.40	0.30	0.40	0.30
7.0 (Part L, 2005)	0.25	0.20	0.30	0.20
5.0	0.20	0.15	0.20	0.15
3.0	0.15	0.10	0.15	0.10

In addition, Perera (1997) measured the air-tightness of 10 UK office buildings that are naturally ventilated. The measured values ranged between 8.3 $m^3/m^2.h$ and 32.0 $m^3/m^2.h$ at 50Pa with their mean value of 17.9 $m^3/m^2.h$ at 50Pa. Disparity between the measured data and the standard values indicates that actual infiltration rates of existing buildings are often much higher than the values recommended in the standards. Hence, based on both the standard and the measured data, we quantify the minimum and maximum values as 0.10 and 1.25 (1/h) respectively with 0.50 (1/h) as the base value.

Table 3.7 Pressurized test results for 10 office buildings in UK (Perera, 1997)

Buildings	Air Permeability ($m^3/ m^2.h$ at 50Pa)			
	Mean	Std. Dev	Min	Max
10 BRE office buildings	17.9	9.15	8.3	32.0

3.3.3 Natural Ventilation

Natural ventilation introduces outdoor fresh air into a building through window openings due to the pressure difference across the window. The volumetric flow rate q_v

(m³/s) through single-sided open windows can be estimated from Equation 3.2. C_D refers to the flow discharge coefficient, A_e the effective window opening area, Δp the pressure difference across the opening (Pa), and ρ the air density (kg/ m³). The pressure difference Δp represents the sum of direct wind pressure, thermal buoyancy, and fluctuations. Since wind pressure is generally known as the dominant parameter, we ignore the terms representing thermal buoyancy and fluctuations. Following the relationships derived in Larsen and Heiselberg (2008), Δp across the window opening can be thus approximated by the ratio of local air velocity at the window and a reference mean velocity of outside air, at different wind directions, and for a given configuration of the building. Since Δp is derived mainly as a function of external weather conditions, we do not consider it as an uncertainty parameter in this study. C_D represents the fractional airflow loss due to the geometry of the windows. We use the empirically derived values of C_D found in de Wit (2001), which estimates their values to range between 0.60 and 0.75 for rectangular openings. A_e represents half the window opening area only for the ingoing airflow, and is thus a proportion of the total operable window area A of the building depending on the percentage of windows open at a given time denoted as f_a.

$$q_v = C_D \cdot A_e \cdot \sqrt{2 \cdot |\Delta p| / \rho} \tag{3.2}$$

$$A_e = f_a \cdot A/2 \tag{3.3}$$

Quantifying f_a for a building can be quite difficult since the occupant action of opening or closing windows is controlled by a set of diverse factors. Studies have shown that occupants can have significantly different levels of activeness in relation to opening and closing windows. The act of opening or closing a window, and the duration for which they are left in one state or the other is triggered by a whole range of physical, environmental, and psychological factors. Borgeson and Brager (2008) study a large set of variables potentially influencing occupant control of windows, and summarize existing body of work in this area, many of which are empirical studies for quantifying f_a as a

function of local environmental conditions. From these studies, we consider a logistic regression model derived from a study of fifteen office buildings in UK (Rijal, 2007). The empirical model is able to compute the probability f_a of a window being open as a function of outdoor air temperature by the logit link function;

$$f_a = \frac{e^{b \cdot T + c}}{1 + e^{b \cdot T + c}} \tag{3.4}$$

where T is the outdoor temperature at a given time, b is the regression coefficient for T, and c is the intercept in the regression equation. This study also derived the coefficient and the intercept values fit for each office building and for all buildings as summarized in Table 3.8. The intercept *c* varies from -3.90 to -2.09, depending on occupant behavior in a building with -2.92 as the base value. Coefficient *b* is also uncertain, but the range of values reported seems small enough for us to ignore them.

Table 3.8 Regression coefficient and intercept values from field surveys (Rijal, 2007)

Building	*b*	*c*
Each	0.160	[-3.80, -2.09]
All	0.157	-2.92

3.3.4 Heating System

Normative models parameterize the heating system of the building by its two main components: the seasonal efficiency of the heat generation system and the losses in the distribution system. Thermal efficiency of the heat generating equipment is typically documented in manufacturers' catalogs from experiments under full-load standard testing conditions. However, its seasonal efficiency can differ depending on its actual operation conditions: the frequency of occurring partial loads and the return water temperature to the boiler. Table 3.9 summarizes the range of steady-state efficiency for the three major types of boilers depending on the operation conditions: (a) the temperature regime (i.e.,

80/60°C, 50/30°C) and (b) loads on the boiler (i.e.,100% load, 30% load). The ranges are estimated by the ECOBOILER model that integrates the boiler model with the building model to realistically estimate the boiler performance in the context of actual operations (Kemna, 2007).

Table 3.9 Ranges of steady-state efficiency for three types of boilers (Kemna, 2007)

Boiler type	80/60°C regime		50/30°C regime	
	full-load	part-load	full-load	part-load
Condensing	84 - 88 %	83 - 87 %	94 - 98 %	93 - 97 %
Low temperature	80 - 83 %	79 - 85 %	-	-
Standard	78 - 81 %	76 - 79 %	-	-

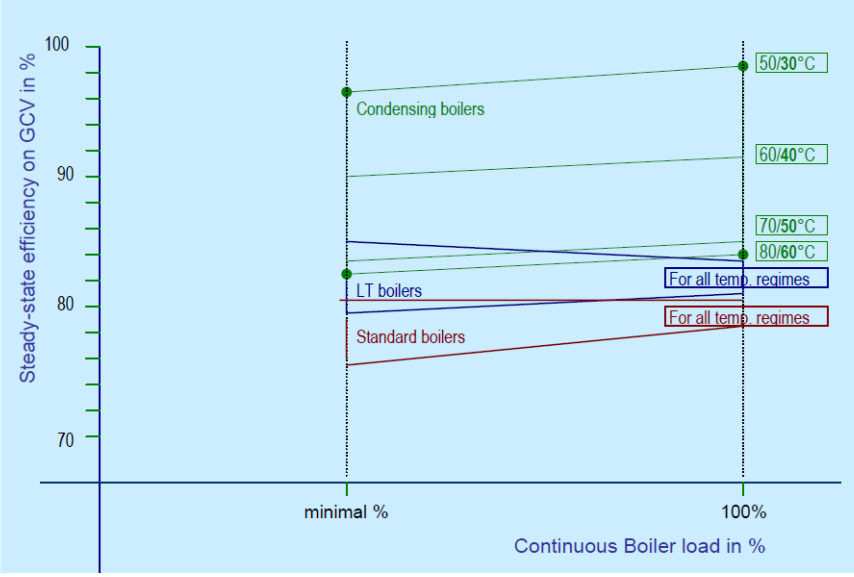


Figure 3.1 Steady-state efficiencies for the three types of boilers (Kemna, 2007)

Figure 3.1 plots the range of steady-state efficiency for each boiler type for all temperature regimes. For standard and low-temperature boilers, their efficiency noticeably degrades when they operate under partial loads, but it does not change much due to the temperature regime. On the contrary, for condensing boilers the effects of partial loads on the system efficiency are negligible while the effects of the temperature

regime are significant. Hence, for condensing boilers, the main source of uncertainty about the boiler efficiency is due to the fact that the return water temperature is not known precisely. Provided the boiler efficiency is steady over partial loads, Lazzarin and Schibuola (1986) developed three empirical relationships that estimate average, minimum, and maximum performance efficiency as a function of the return water temperature within the temperature range between 50℃ and 70℃. We follow these relationships to quantify the bounds of uncertainty for the seasonal efficiency of condensing boilers.

$$\text{For the maximum:} \quad \eta_{boiler} = 97.0 + \frac{1.4(70-T_r)}{20} \tag{3.5}$$

$$\text{For the minimum:} \quad \eta_{boiler} = 94.8 + \frac{0.4(70-T_r)}{20} \tag{3.6}$$

$$\text{For the average:} \quad \eta_{boiler} = 96 + \frac{(70-T_r)}{20} \tag{3.7}$$

Heat generated by the boiler is delivered through a distribution system, and part of this heat energy is lost during delivery. The efficiency of the distribution system depends on its distribution length, its insulation level, and its operation and surrounding conditions. Heat losses in hydraulic distribution systems $Q_{H,dis,ls}$ can be quantified by the following relationship (EN 15316-2-3, 2007). $\Psi_{L,j}$ is the linear thermal transmittance of the pipes in zone j, Θ_m is the supply water temperature, Θ_j is the temperature of the surrounding spaces, L_j is the total length of pipes in zone j, and t_{op} is the number of hours when zone j is heated. While most of these parameters are observable, it is quite difficult to obtain them even from the most thorough building audit. In fact, most transient simulation models tend to ignore heat losses in the distribution system. Therefore, we use our best estimates for parameters described in Equation 3.8 based on operation manuals and construction drawings and specifications. With their uncertainty range obtained from the literature study, we derive the range of $Q_{H,dis,ls}$ through the

equation while using the normative value recommended by the CEN-ISO standards as the base value.

$$Q_{H,dis,ls} = \sum_{j=1}^{n} \Psi_{L,j} \cdot (\Theta_m - \Theta_j) \cdot L_j \cdot t_{op} \quad (3.8)$$

Table 3.10 Parameter uncertainty pertaining to the calculation of distribution heat losses

Parameter	Uncertainty Range	Reference
Ψ_L	± 10%	McDonald, 2002
Θ_m	± 1.5 ℃	ASHRAE, 1999
Θ_j	± 2.0 ℃	ASHRAE, 1999
L_j	± 15%	-

3.3.5 Cooling System

Normative models parameterize the cooling system of the building by its two main components: the seasonal efficiency of the cold generators and the losses in the distribution system. As described in the heating system, the seasonal efficiency of the cold generators also varies from its nominal efficiency measured under the standard testing conditions because the testing conditions cannot capture dynamic operation conditions under which the cooling system actually operates. Equation 3.9 in (EN 15243, 2007) calculates the seasonal energy efficiency ratio ($SEER$) from the steady-state energy efficiency ratio at full load (EER) while accounting for the system efficiency loss under partial loads with the use of part load factor PLF. The system efficiency loss arises from cyclic effects due to the fact that the cooling system operates under partial loads, often resulting in cycling between on and off states.

$$SEER = EER \cdot PLF \quad (3.9)$$

$$PLF = 1 - C_d \cdot (1 - PLR) \quad (3.10)$$

ANSI/AHRI Standard 210/240 (2008) presents a formula (Equation 3.10) that calculates PLF as a function of part load ratio PLR and degradation coefficient C_d. PLR

refers to the ratio of the partial load on the cooling system to its design output, and C_d refers to the system cycling loss coefficient measured by cycling loss tests and documented in manufacturers' catalogs. Southern California Edison (2005) analyzed over 23,000 air-conditioners and heat pumps listed in the California Utilities database, and estimated C_d to range between 0.02 and 0.25. Hu (2009) reviewed existing field measurement data for air-source heat pumps for cooling, and quantified the bounds of C_d to be in the range of 0.066 - 0.26 with 0.25 as the base value. From these studies, we quantify the bounds of uncertainty for the coefficient to be between 0.02 – 0.26.

While cold generated by the cold generators is delivered to spaces, part of this cold energy is lost during delivery. For air distribution systems, part of conditioned air is lost through the duck leakage, and the volume of the loss depends on the duct system design, its construction quality, and its operation conditions (pressure difference between the duct and the surroundings). Equation 3.11 from (EN 15242, 2007) calculates the airflow through the duct leakage $q_{v,ductleak}$ (m³/h) from the duct area A_{duct} (m²), the duct air-tightness K (m³/s· m²), and the pressure difference dP_{duct} (Pa). Table 3.11 summarizes a normative duct leakage value for each duct system class provided by the CEN-ISO standard. Based on the standard, we estimate the cooling distribution loss factor to fall between 0% - 15%.

$$q_{v,ductleak} = \frac{A_{duct} \cdot K \cdot dP_{duct}{}^{0.65}}{3600} \tag{3.11}$$

Table 3.11 Typical values of duct leakages (EN 15242, 2007)

	K	Distribution loss factor lost/airflow (%)
2.5. class A	0.0000675	0.15
class A	0.0000270	0.06
class B	0.0000090	0.02
class C or better	0.0000030	0.00

3.3.6 Domestic Hot Water System

Normative models characterize the energy performance of domestic hot water (DHW) system in terms of the DHW generation efficiency and the distribution efficiency. The water heater efficiency can vary depending on its nominal efficiency, its thermostat settings, its operation schedule, and surrounding conditions. Fanney (1996) investigated the effects of various off-peak schedules on the thermal efficiency of electric water heaters through laboratory tests, and concluded that the efficiency can vary up to 7%. Healy (2001) tested electric water heaters from five manufacturers, the rated efficiency of which ranges between 0.92 and 0.94. Table 3.12 summarizes the rated and the measured efficiency in the laboratory which varied from 0.87 to 0.95. This difference between the measured and rated efficiency suggests that in reality water heaters may often underperform the expected efficiency under actual operation conditions. From these test results, we determine the range of thermal efficiency for electric water heaters to be between0.87 and 0.95. Regarding the distribution efficiency, in common cases in which water heaters are locally distributed to support ancillary areas, the delivery system length is quite short. Thus, heat loss during delivery can be regarded as negligible, so the distribution efficiency is assumed to be ideal.

Table 3.12 Rated and measured efficiencies of electric water heaters (Healy, 2001)

Type	Rated Efficiency	Measured Efficiency 1	Measured Efficiency 2
1	0.93	0.896	-
2	0.92	0.908	-
3	0.93	0.884	0.876
4	0.94	0.888	0.881
5	0.94	0.894	-
6	0.93	0.918	0.949
7a	0.93	0.909	0.936
7b	0.93	0.904	-
8a	0.93	0.896	0.881
8b	0.93	0.895	-

3.3.7 Internal Gains

Internal heat gains refer to the heat produced by occupants, lights, and plug-in appliances in a building. Heat gains from occupants depend on the number of occupants and their metabolic rates in spaces. Even if the number of occupants in the space temporarily fluctuates, building occupancy schedule is considered as the fixed profile that capture the average building occupancy pattern. Occupant metabolic rates depend on individuals' activity level. Table 3.13 shows the range of metabolic rates for the four groups of activities based on Macdonald (2002).

Table 3.13 Uncertainty range of metabolic rates for the four groups of activities

Activity	Metabolic rate (W)	
	Min	Max
Sedentary	70	130
Light work	130	250
Medium work	200	425
Exercising	425	950

Uncertainty in plug-in equipment loads is quantified based on the survey of 30 UK buildings from (Dunn, 2005). The survey suggests that plug-in equipment loads range between 124 W and 229 W per person in a building with 158 W per person as the mean value. Based on these values, we calculate the range of possible plug-in equipment loads from the number of occupants for a particular building under investigation.

3.4 Goals of Uncertainty Analysis

3.4.1 Identification of Dominant Parameters

Building energy models contain a large number of uncertain parameters, and it is infeasible to calibrate all uncertain parameters given the limited measurement data on aggregate-level energy uses. Hence, we apply a parameter screening technique in order to select a smaller number of calibration parameters more objectively than using our own judgment. Particularly, we employ the Morris method (Morris, 1991) to rank uncertain parameters with respect to the effects of their uncertainty on energy consumption. The Morris method has been acknowledged as a suitable screening technique for building energy models (de Wit, 2001; Moon, 2005). First, this method is computationally efficient to test the sensitivity of many uncertain parameters with relatively small samples. Moreover, the method does not assume the relationship between parameters and model outcomes as linear, and evaluates the effects of parameters on the model outcome over the whole parameter space by exploring multiple regions sampled from the parameter space. Hence, the method can capture nonlinear effects of individual parameters and interaction effects among parameters.

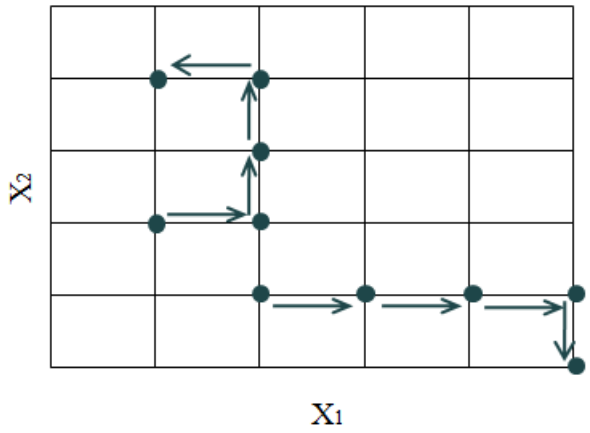


Figure 3.2 Illustration of the Morris method (six-level, two-dimensional space)

Figure 3.2 illustrates the Morris method with two-dimensional parameter space. The Morris method first discretizes the parameter space of the entire design X; it divides each parameter space into a chosen number of levels that correspond to a pre-selected number of quantiles of the corresponding parameter. This forms a grid of values in the parameter space. After starting from an initial fixed point in that grid, the move to the next step is done by changing one parameter value at a time while the other parameter values stay the same; there is no diagonal move, only moves along axes. Eventually, this allows moves in all directions. At the end of each step, we obtain a number: the elementary effect equal to the change in the model outcome as the result of the change Δ in one input value (Equation 3.12). At the end of the entire procedure, we obtain distributions of elementary effects for all parameters. The mean value of each distribution represents the overall importance of an individual parameter. This study uses Simlab version 2.2 to execute the Morris method (SIMLAB, 2009).

$$d_i(X) = \frac{y(x_1,x_2,\ldots,x_{i-1},x_i+\Delta,x_{i+1},\ldots,x_k)-y(X)}{\Delta} \quad (3.12)$$

3.4.2 *Propagation of Uncertainty*

Uncertainties in projected energy-savings predictions arise broadly from uncertainties existing in the calibrated model and additional uncertainties coming from candidate ECMs. Figure 3.3 illustrates how to propagate these uncertainties through the energy model to obtain probability distributions of outcomes. Uncertainties in the calibrated model are conveyed in a form of posterior distributions of individual calibration parameters. Since the posterior distributions are correlated, we cannot treat individual posterior distributions independently. Instead, we randomly select a row of calibration parameter values from posterior realizations to propagate uncertainty existing in the baseline model. On the contrary, additional uncertainties in model parameters pertaining to a testing ECM can be considered as independent since each parameter represents a specific property of the system. This is a valid assumption that has been underlined in most studies of uncertainty analysis (Hu, 2009; Moon, 2005; Macdonald, 2002; de Wit, 2001).

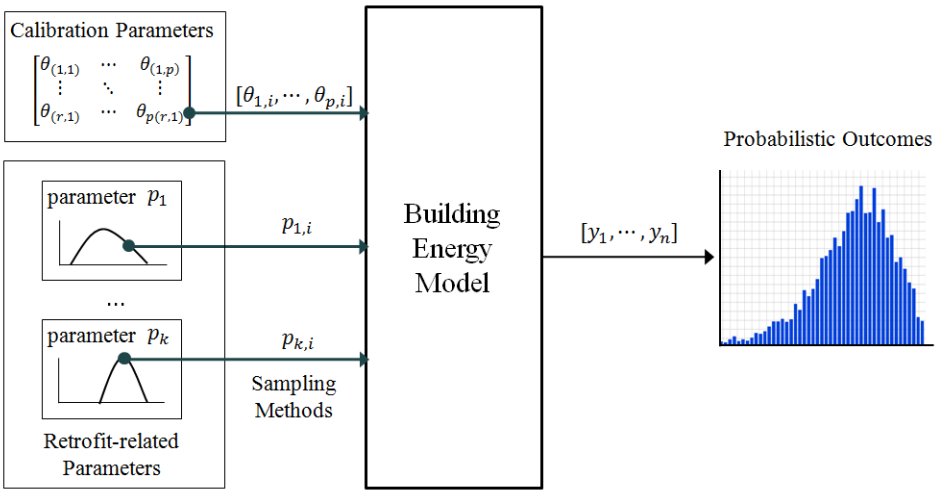


Figure 3.3 Illustration of uncertainty propagation for probabilistic outcomes

For uncertainty propagation, we need to sample values from the probability density functions of uncertain parameters. The most commonly used sampling methods

are Monte Carlo method and Latin Hypercube Sampling (Wyss, 1998). Monte Carlo method randomly draws values from uncertainty distributions, which often requires a large number of samples to ensure convergence to the true probability density. Hence, when a high-fidelity simulation model is deployed, this method is often not efficient due to high computational burden. Latin Hypercube sampling alleviates computational burden by efficiently capturing the real variability of the distributions of uncertain parameters. Latin Hypercube Sampling partitions a probability density function into segments by the same magnitude of probability, and draws a sample once from each of the segments. As a result, this method ensures the reliability of probabilistic outcomes with a much smaller sample size. Hence, we apply the Latin Hypercube Sampling method to propagate uncertainty, and execute the method with the use of Simlab version 2.2 (SIMLAB, 2009).

CHAPTER 4 ANALYSIS ON THE FEASIBILITY OF THE NEW FRAMEWORK ON RETROFIT DECISION-MAKING

4.1 Case Study 1

4.1.1 Building Description

The first case building is the Faculty of English building in the University of Cambridge in the UK. The building consists of offices and seminar rooms as shown in Figure 4.1. The building is equipped with a condensing gas boiler with radiators for space heating. It does not have mechanical cooling systems, but utilizes natural ventilation for both ventilation requirements and space cooling. It has high-frequency fluorescent lighting, electric heaters decentralized for domestic hot water supply, and nominal office plug-in appliances. Since heating is the dominant energy consumer in this building, we evaluate only Energy Conservation Measures (ECMs) that reduce gas consumption in the retrofit analysis process. Accordingly, we built an energy model related to space heating, and calibrated the energy model with monthly gas utility bills.

Figure 4.1 Elevation and typical floor plan of the first case building

We follow the standard process of consulting design specifications, construction documents and operation manuals to build the normative energy model of the building. The main parameters of the energy model can be broadly summarized within following groups: (a) building envelope properties (e.g., thermal transmittance, emissivity, solar absorptance, heat capacity), (b) internal loads (plug-in appliances, lighting, occupants), (c) properties of the HVAC systems, (d) properties describing ventilation and infiltration, and (e) external environment (weather data). Then, following audits and interviews with the building manager, we have adjusted initial values assigned to the model parameters such that the energy model aligns with the actual building. Figure 4.2 shows predicted gas energy uses by the normative model (red color) and the transient simulation model (blue color) against the three-year gas utility data (black color). The comparisons demonstrate that there is considerable difference between the outputs from the energy models and actual gas consumption of the building.

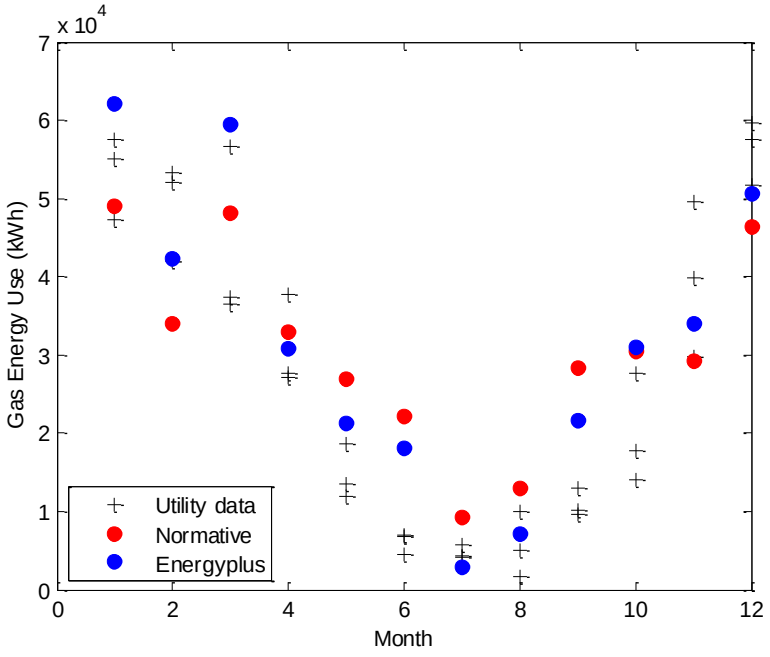


Figure 4.2 Three-year gas utility data against model predictions

4.1.2 Bayesian Calibration of Normative Model

4.1.2.1 Prior Uncertainty Quantification

As the first step, we need to quantify the uncertainties in model parameters by reviewing published literature and industry standards. Table 4.1 summarizes the uncertainties around the initial values assigned to model parameters by listing the minimum and the maximum limit of the values that can be assigned to them. Chapter 3 describes the process of quantifying the uncertainties in the normative model in detail. The uncertainty information is essential for the following two steps in the analysis process: parameter screening and Bayesian calibration. The Morris method used for parameter screening utilizes only the bounds of parameter uncertainty to determine dominant uncertain parameters. On the contrary, Bayesian calibration exploits full information about prior uncertainty distributions. Hence, we translate these values in Table 4.1 into prior uncertainty distributions $p(\theta)$ by assigning a triangular distribution to parameters. The base value is the top of the triangle, and its probability decays linearly to zero at the edges of the assigned interval.

Table 4.1 Uncertain parameters and their ranges in the normative model

Model Parameters	Base	Min	Max	Reference
Thermal Properties				
Roof U value (W/m²·K)	0.19	0.17	0.21	see Chapter 3.3.1
Roof solar absorptance	0.68	0.60	0.76	
Roof Emissivity	0.91	0.87	0.95	
Wall U-value (W/m²·K)	0.32	0.29	0.36	
Wall solar absorptance	0.63	0.43	0.83	
Wall emissivity	0.91	0.87	0.95	
Window U-value (W/m²·K)	2.62	2.36	2.88	
Window solar transmittance	0.77	0.76	0.79	
Window emissivity	0.84	0.75	0.92	
Envelope heat capacity (kJ/m²·K)	260	160	275	
Internal Loads				
Lighting power density (W/m²)	13	11	15	building log book
Appliance power density (W/m²)	15	12	22	see Chapter 3.3.7
Occupant metabolic rate (W)	80	70	130	see Chapter 3.3.7
Control				
Indoor heating temperature (℃)	22	20	24	building log book
Ventilation				
Infiltration rate (1/h)	0.50	0.10	1.25	see Chapter 3.3.2
Discharge coefficient	0.68	0.60	0.75	see Chapter 3.3.3
Intercept c	-2.92	-3.80	-2.09	see Chapter 3.3.3
Heating System				
Heating generation efficiency	0.97	0.95	0.98	see Chapter 3.3.4
Heating distribution loss factor	0.08	0.06	0.16	

4.1.2.2 Parameter Screening

We apply the Morris method described in the Chapter 3.4 to identity dominant uncertain parameters according to their effects on gas energy use. We generated five independent samples to calculate the elementary effects of individual parameters, which resulted in 100 runs for the nineteen uncertainty parameters. Table 4.2 shows the ranking of uncertain parameters by their relative importance on the energy consumption of the building. We selected the top four parameters to calibrate out energy model since it is a reasonable number given that we have three years of monthly gas consumption data (36 observations).

Table 4.2 Ranking of model parameters in the normative model by relative importance

Rank	Model Parameter
1	Intercept c for windows open
2	Indoor temperature
3	Infiltration rate
4	Discharge coefficient
5	Appliance power density
6	Window U-value
7	Heating distribution loss factor
8	Lighting power density
9	Envelope heat capacity
10	Occupant metabolic rate

4.1.2.3 Model Calibration

The Kennedy and O'Hagan formulation of Bayesian calibration (2001) requires three sets of data as input: (1) the prior probability density functions of calibration parameters, (2) computer outputs from exploring the space of calibration parameters, and (3) observation data (monthly utility bills in this study). For the parameter-space

exploration, we apply Latin Hypercube Sampling technique (Wyss, 1998) to generate model outputs given the calibration parameter space since this technique can sufficiently explore the space with much reduced samples. For the observation data, we utilize monthly gas bills over the three years.

Figure 4.3 shows the posterior distributions of the four calibration parameters against the prior distributions assigned to them. For the intercept c, the posterior distribution is quite strongly towards the lower bound, and its mode is at -3.6 closely near the lower bound. The posterior distribution suggests that the proportion of windows open in this case building is far smaller than the average in UK buildings. For the indoor temperature, the posterior distribution shifts to the lower bound by 1℃, and its expected value is around 21℃ deviating from the originally estimated value (22℃). Regarding the infiltration rate, the posterior distribution indicates that the infiltration rate of the case building is likely to be much higher than average UK buildings ($0.5h^{-1}$), and the spread of uncertainty about it is much smaller. For the discharge coefficient, the posterior distribution does not change significantly from the prior distribution. It should be re-emphasized that the posterior distributions of individual parameters are derived from the joint multivariate distribution, and are hence correlated. This means that a specific value of one parameter coincides with a certain value of other parameters. Hence, posterior distributions of all four calibration parameters should be applied in conjunction when exercising the model for retrofit analysis.

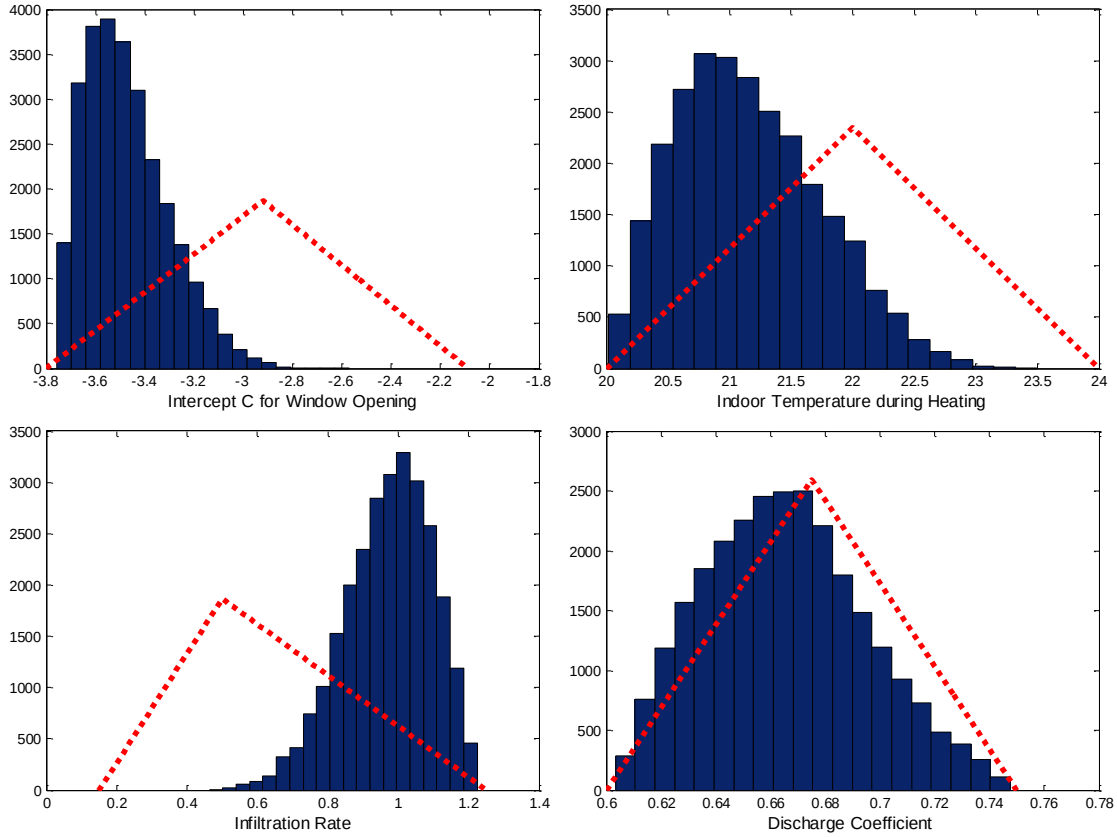


Figure 4.3 Posterior distributions of calibration parameters with the normative model (posterior - blue, prior - red)

4.1.3 Bayesian Calibration of Transient Simulation Model

4.1.3.1 Prior Uncertainty Quantification

Quantification of uncertainties in model parameters depends on the choice of a building energy model. For instance, the normative model describes the characteristics of major components whereas the transient simulation model requires a detailed level of model parameters that describe the physical behavior of individual parts. Owing to the different granularity level, parameter uncertainty in the transient simulation model can differ from that in the normative model. Chapter 3 describes the process of quantifying parameter uncertainty in the normative model: investigating uncertainties in detailed

model parameters and aggregating them to capture uncertainty in macro-level parameters. The first step yields quantified parameter uncertainty in the detailed transient simulation model. Table 4.3 summarizes the list of uncertain parameters in the energyplus model and their uncertainty ranges.

Table 4.3 Uncertain parameters and their ranges in the energyplus model

Model Parameters	Base	Min	Max	Reference
Thermal Properties				
Stone conductivity (W/m²·K)	3.17	2.85	3.49	see Chapter 3.3.1
Stone density (kg/m³)	2560	2509	2611	
Stone specific heat (kJ/kg·K)	0.79	0.60	0.98	
Stone solar absorptance	0.63	0.43	0.83	
Stone emissivity	0.91	0.87	0.95	
Concrete conductivity (W/m²·K)	2.15	1.94	2.37	
Concrete density (kg/m³)	2400	2352	2448	
Concrete specific heat (kJ/kg·K)	0.90	0.68	1.12	
Concrete solar absorptance	0.68	0.60	0.76	
Concrete emissivity	0.91	0.87	0.95	
Insulation conductivity (W/m²·K)	0.03	0.0	0.03	
Insulation density (kg/m³)	43	42.14	43.86	
Insulation specific heat (kJ/kg·K)	1.21	0.91	1.51	
Plaster conductivity (W/m²·K)	0.16	784	0.18	
Plaster density (kg/m³)	800	823	816	
Plaster specific heat (kJ/kg·K)	1.09	0.82	1.36	
Plaster solar absorptance	0.40	0.34	0.46	
Plaster emissivity	0.90	0.86	0.94	
Glass conductivity (W/m²·K)	0.90	0.81	0.99	
Glass emissivity	0.84	0.76	0.92	
Glass solar transmittance	0.77	0.76	0.79	

Table 4.3 continued

Internal Loads				
Lighting power density (W/m²)	13	11	15	building log book
Appliance power density (W/m²)	15	12	22	see Chapter 3.3.7
Occupant metabolic rate (W)	80	70	130	see Chapter 3.3.7
Control				
Indoor temperature (°C)	22	20	24	building log book
Ventilation				
Infiltration rate (1/h)	0.50	0.10	1.25	see Chapter 3.3.2
Discharge coefficient	0.68	0.60	0.75	see Chapter 3.3.3
Intercept c	-2.92	-3.80	-2.09	see Chapter 3.3.3
Heating System				
Boiler nominal efficiency	0.97	0.95	0.98	Lazzarin (1986)

4.1.3.2 Parameter Screening

With the Morris method, we generated five independent samples to obtain the elementary effects of individual parameters. As a result, we ran 150 simulation runs to determine calibration parameters among 29 uncertain parameters. Table 4.4 shows the ranking of uncertain parameters in the energyplus model ordered by their relative importance regarding the gas energy consumption. The five most dominant parameters are the same as those ranked as the most dominant parameters in the normative model shown in Table 4.2. Further below, the ranking of parameters in the two model differs which is expected since they compute heat and mass flows in a building at very different resolutions.

Table 4.4 Ranking of model parameters in the energyplus model by relative importance

Rank	Model Parameter
1	Intercept c for windows open
2	Indoor temperature during heating
3	Infiltration rate
4	Discharge coefficient
5	Appliance power density
6	Lighting power density
7	Glass emissivity
8	Boiler nominal efficiency
9	Insulation conductivity
10	Glass solar transmittance

4.1.3.3 Model Calibration

Figure 4.4 shows the posterior distributions of the four calibration parameters with the energyplus model against their prior distributions. They are, in general, very similar to the posterior distributions derived from the calibrated normative model. However, some difference between the calibration results is unavoidable since the two models have a quite different model resolution level. For the intercept c, the normative model produced the posterior distribution that is slightly more towards the lower bound than the energyplus model. For the infiltration rate, the normative model resulted in the posterior distribution that is more strongly towards the upper bound than the energyplus model. This discrepancy between the two calibration results can be regarded as trivial given the magnitude of uncertainty reduced by Bayesian calibration and coincidence between the two results. The significance of this discrepancy will be further investigated in the following section with respect to its effect on predictions and decisions in the decision-making context.

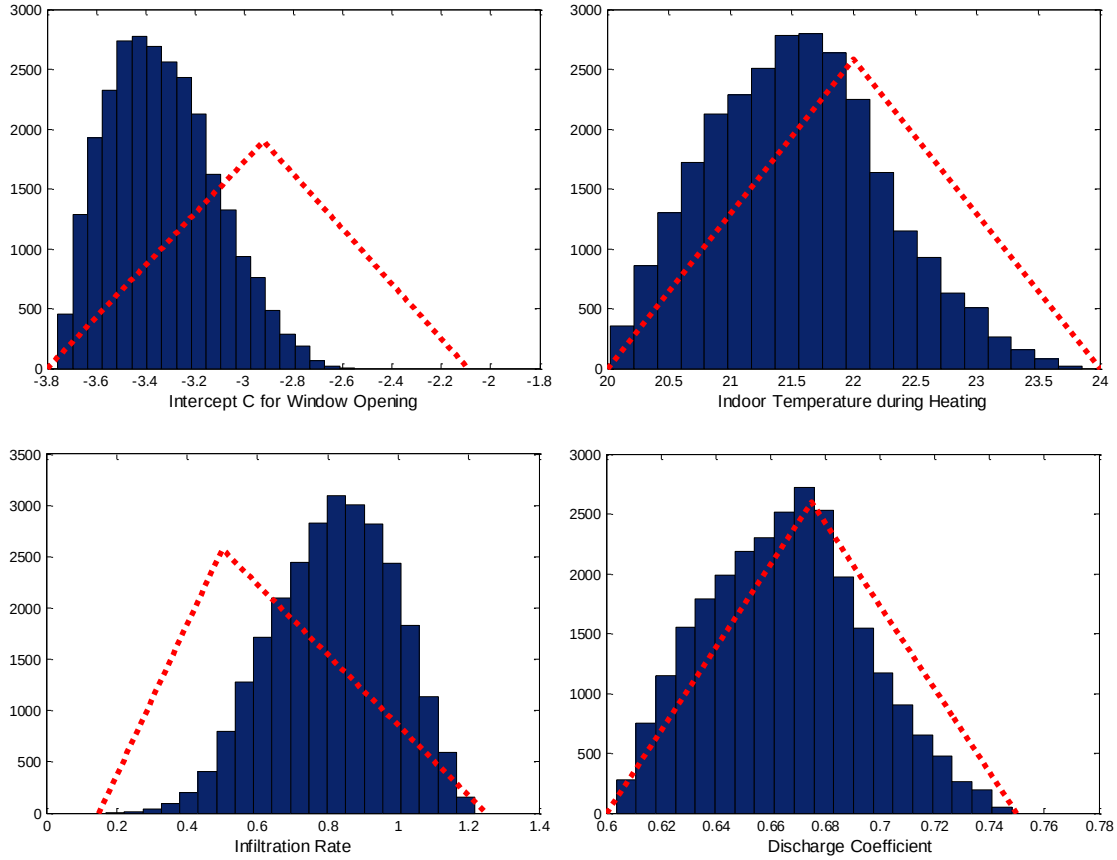


Figure 4.4 Posterior distributions of calibration parameters with the energyplus model (posterior-blue, prior-red)

4.1.4 Evaluation: Comparison between Normative and Transient Energy Model

We compare the calibrated normative model with the calibrated energyplus model under the three evaluation criterion summarized in Chapter 2.3. The first criterion evaluates the accuracy of the calibrated models by using standard validation metrics such as index of agreement *d* and coefficient of variation of the root mean square error *CVRMSE*. These metrics compare the outputs from the calibrated model with observed values of energy consumption. Table 4.5 shows the values of *d* and *CVRMSE* of the model predictions before calibration and after calibration. These validation metric values indicate that the calibrated normative model predicts as accurately as the calibrated

energyplus model and better than the uncalibrated energy model. Furthermore, the validation measures demonstrate that Bayesian calibration enhances the accuracy of the baseline model as it reduced the *CVRMSE* value by half. Figure 4.5 also depicts that the calibration improves the predictive power of the normative model for the particular building at the same confidence level as the energyplus model.

Table 4.5 Validation measures for uncalibrated and calibrated models

Type of Model	*d*	*CVRMSE*
before calibration		
Normative	0.76	0.34
Energyplus	0.88	0.30
after calibration		
Normative	0.97	0.17
Energyplus	0.97	0.17

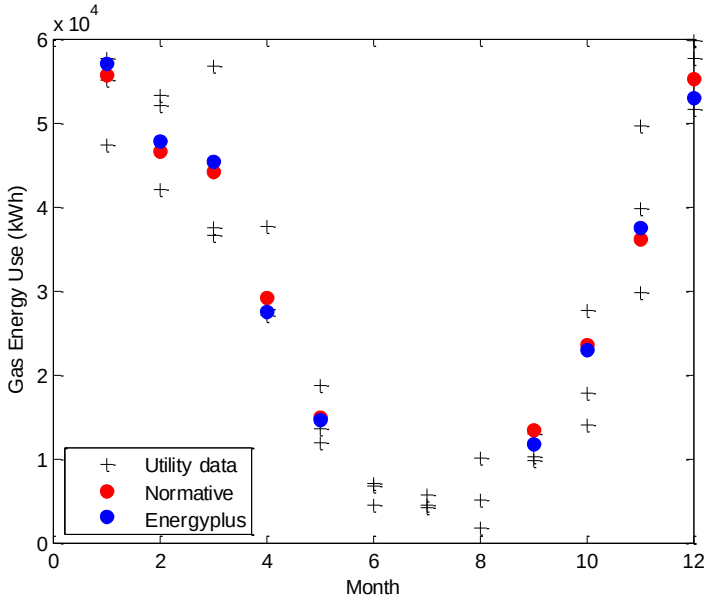


Figure 4.5 Predicted gas energy uses from the calibrated models against the monitored gas uses

To evaluate the feasibility of the Bayesian-calibrated normative model as a method to assist retrofit decision-making, we compare the normative model with the transient simulation model in the context of retrofit decision-makings. We consider three

ECMs for this building: (1) insulation addition, (2) window replacement, and (3) air-tightness improvement. Table 4.6 lists the parameter values specified for these three improvements and their uncertainty ranges, and Table 4.7 lists the capital costs roughly estimated for the improvements based on the BCIS price book for building refurbishments (BCIS, 2010). For the calculation of annual saving costs, we keep the gas price at 2.4p/kWh (DECC, 2010b).

Table 4.6 Uncertain parameters and their ranges for the three ECMs

Parameters	Base	Min	Max	Reference
ECM1: Insulation Addition				
Insulation U-value (W/m²·K)	0.30	0.27	0.33	see Chapter 3.3.1
ECM2: Window Replacement				
Window U-value (W/m²·K)	1.41	1.27	1.55	see Chapter 3.3.1
Window solar transmittance	0.65	0.63	0.67	
Window emissivity	0.05	0.04	0.06	
ECM3: Air-tightening				
Infiltration rate reduction (%)	11	1	31	Jacobson (1986)

Table 4.7 Cost estimates of the three ECMs (in 1000£)

Retrofit Options	Base	Min	Max
ECM1: Insulation Addition	16	15	17
ECM2: Window Replacement	204	194	214
ECM3: Air-tightening	10.5	10	11

To verify the prediction accuracy of the normative model (evaluation criterion 2), we assess the disparity of predictions between the two calibrated models. First, the two-sample K-S test is used to examine whether the two SPT distributions generated by the two models come from the same cumulative distribution. Table 4.8 shows that for the three ECMs the K-S tests reject the hypothesis that the two samples of outcomes are from

the same distribution. Although the two models do not yield identical outcomes for ECM predictions, graphical comparison in Figure 4.6 suggests that the SPT distributions predicted by the normative model coincide well with those predicted by the energyplus model. These comparisons conclude that the calibrated normative model can adequately estimate the cost-effectiveness of ECMs.

Table 4.8 Two-sample K-S tests for predictions from the two calibrated models

Retrofit Options	Two-sample K-S test	
	H	p-value
ECM1: Insulation Addition	1	0.00
ECM2: Window Replacement	1	0.00
ECM3: Air-tightening	1	0.00

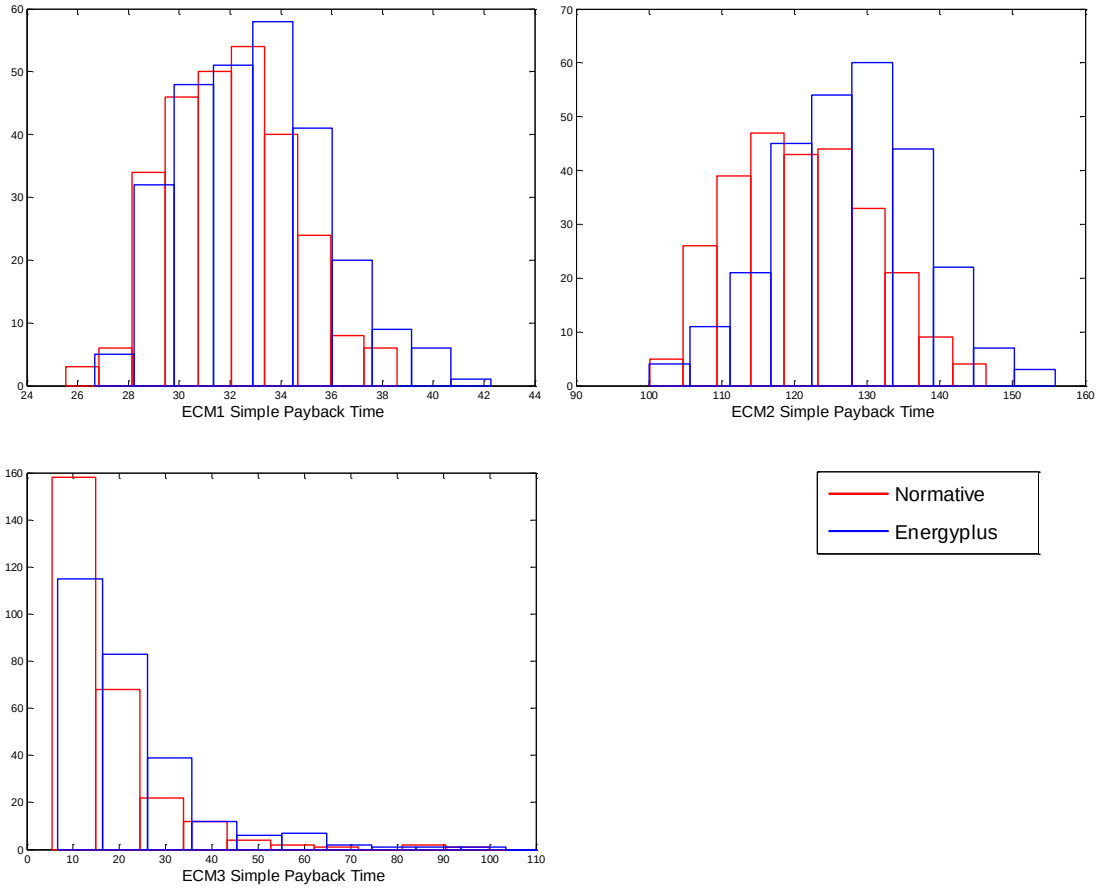


Figure 4.6 Simple Payback Time distributions of the three ECMs from the two models

Finally, we evaluate the two calibrated models to see if they derive the consistent rankings of candidate ECMs for supporting decisions despite the difference in the predictions. The ECMs are evaluated under the three decision-making scenarios which corresponds to different levels of decision-makers' risk awareness, which is summarized in Chapter 2.3.4. Table 4.9 shows the ranking of the three ECMs in the three decision-making scenarios. This result shows that the ranking of the ECMs differs based on the decision-makers' willingness to accept a certain level of risks. In addition, the result shows that the normative and the energyplus model result in the same ranking of the three ECMs. This result demonstrates that lower resolution of the normative model does not bias decisions in the retrofit analysis process, and the normative model can adequately support retrofit analysis.

Table 4.9 Ranking of the three ECMs for the three decision-making scenarios

ECMs	Scenario 1	Scenario 2	Scenario 3
Normative			
ECM1	2	1	1
ECM2	3	3	3
ECM3	1	2	2
Energyplus			
ECM1	2	1	1
ECM2	3	3	3
ECM3	1	2	2

4.2 Case Study 2

4.2.1 Building Description

The second case building is a four-story office building located in London, UK. The three floors above the ground consist of open offices and meeting rooms while the basement floor includes an open office and a copy room with heavily loaded printers. The

building has two gas boilers that provide hot water to radiators for space heating. All floors except the basement are naturally ventilated without any auxiliary cooling. Due to high equipment density, the basement floor is air-conditioned. Electric lighting is provided by T-8 high frequency fluorescents, and domestic hot water is supplied by decentralized electric heaters. For this case study, we constructed a whole-building model, and calibrated the model with monthly gas and electricity utility bills. Figure 4.7 shows energy uses predicted by the two energy models after operational adjustments (adjusting initial values of observable parameters according to site visits and surveys) against five-year utility bills. The figure demonstrates that the energy models without parameter estimation still yield substantial discrepancy between model outcomes and actual energy consumptions.

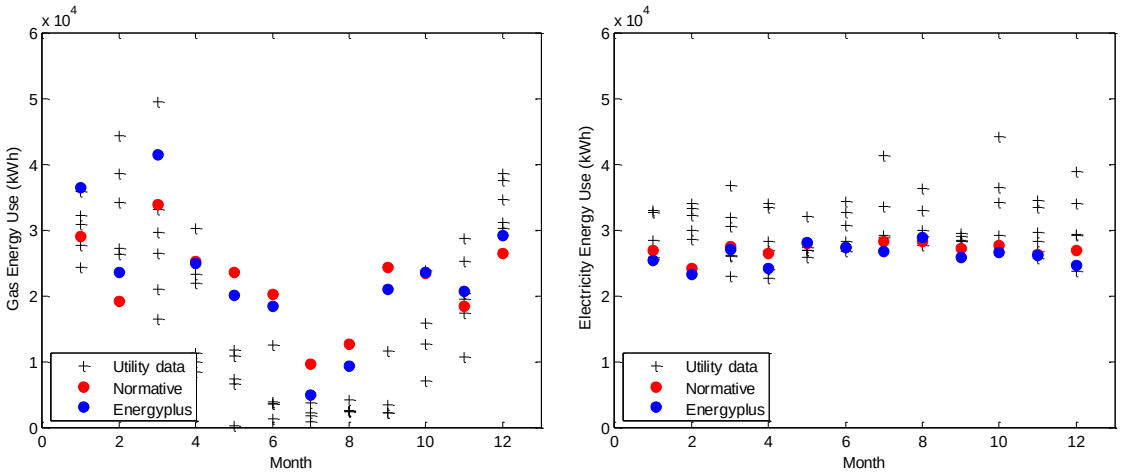


Figure 4.7 Monitored gas (left) and electricity (right) energy uses against model predictions

4.2.2 Bayesian Calibration of Normative Model

4.2.2.1 Prior Uncertainty Quantification

We need to estimate uncertainties in model parameters specific for this building case. First, we assigned base values to model parameters based on design documents, operation manuals, and industry standards to make the model as close to existing building conditions as we can. Then, based on collective expert knowledge from the literature study, we quantified uncertainties in model parameters, which refer to possible deviations from base values. Table 4.10 shows the list of uncertain parameters in the normative model and their uncertainty ranges with the base values.

Table 4.10 Uncertain parameters and their ranges in the normative model

Model Parameters	Base	Min	Max	Reference
Thermal Properties				
Roof U value (W/m²·K)	0.51	0.46	0.56	see Chapter 3.3.1
Roof solar absorptance	0.40	0.34	0.46	
Roof Emissivity	0.90	0.86	0.94	
Wall U-value (W/m²·K)	0.52	0.47	0.57	
Wall solar absorptance	0.40	0.34	0.46	
Wall emissivity	0.90	0.86	0.94	
Window U-value (W/m²·K)	3.16	2.84	3.47	
Window solar transmittance	0.84	0.76	0.92	
Window emissivity	0.84	0.82	0.85	
Envelope heat capacity (kJ/m²·K)	260	160	275	
Internal Loads				
Lighting power density (W/m²)	15	11	19	measurements
Appliance power density multiplier	1	0.78	1.45	see Chapter 3.3.7
Occupant Metabolic Rate (W)	80	70	130	see Chapter 3.3.7

Table 4.10 Continued

Control				
Indoor heating temperature (℃)	22	20	24	measurements/
Indoor cooling temperature (℃)	24	22	26	operation manuals
Ventilation				
Infiltration rate (1/h)	0.50	0.10	1.25	see Chapter 3.3.2
Discharge coefficient	0.68	0.60	0.75	see Chapter 3.3.3
Intercept c	-2.92	-3.80	-2.09	
Heating System				
Heating generation efficiency	0.86	0.84	0.88	see Chapter 3.3.4
Heating distribution loss factor	0.08	0.06	0.15	
Cooling System				
Mean Partial Load Factor	0.84	0.83	0.96	see Chapter 3.3.5
Cooling distribution loss factor	0.06	0.00	0.15	
Domestic Hot Water System				
DHW generation efficiency	0.91	0.87	0.95	see Chapter 3.3.6

4.2.2.2 Parameter Screening

Table 4.11 lists the ranking of uncertain parameters with respect to the effect of their uncertainty on the total energy consumption. We selected the top five uncertain parameters for calibration: (1) intercept c for windows open, (2) indoor temperature during heating, (3) infiltration rate, (4) appliance power density multiplier, and (5) discharge coefficient. The top three parameters have an exceptionally higher impact on the energy consumption than the other parameters in this case building as well as the first case building. This similarity of the results between the two case studies implies that the sensitivity of uncertain parameters can be possibly generalized with proper classification of buildings when much more case buildings are investigated.

Table 4.11 Ranking of model parameters in the normative model by relative importance

Rank	Model Parameter
1	Intercept c for windows open
2	Indoor temperature during heating
3	Infiltration rate
4	Appliance power density multiplier
5	Discharge coefficient
6	Envelope heat capacity
7	Heating distribution loss factor
8	Lighting power
9	Heating generation efficiency
10	Window U-value

4.2.2.3 Model Calibration

Figure 4.8 shows the posterior distributions of the five calibration parameters compared with their prior distributions. For the intercept c, the posterior distribution shifts toward the lower bound. This change suggests that the proportion of open windows in this case is smaller than the average in UK buildings. For the indoor temperature during heating, the posterior distribution shifts to the lower bound by around 1℃. This update indicates that spatially-averaged indoor temperatures during heating in reality is most likely to be lower than the set-point temperature (22℃) due to vertical and horizontal stratifications in spaces. For the infiltration rate, the posterior distribution tells that the building is leakier than average UK buildings. For the appliance power density multiplier the posterior distribution is refined the most from the prior distribution. The expected appliance power density in reality is most likely 20% higher than our prior estimates, and the spread of uncertainty is significantly reduced. On the contrary, the posterior distribution of the discharge coefficient does not change much from the prior distribution.

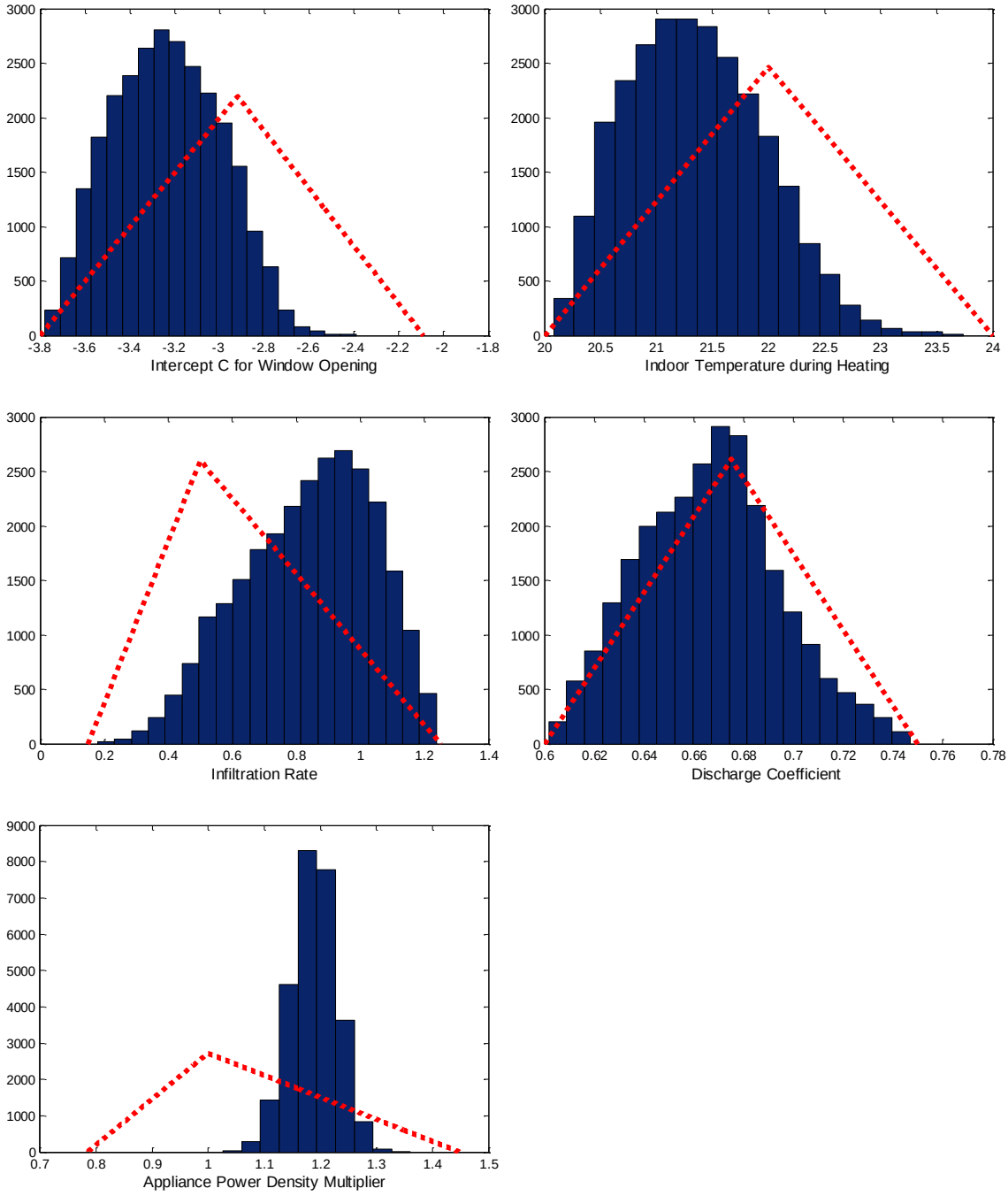


Figure 4.8 Posterior distributions of the five calibration parameters with the normative model (posterior - blue, prior - red)

4.2.3 Bayesian Calibration of Transient Simulation Model

4.2.3.1 Prior Uncertainty Quantification

Table 4.12 summarizes the list of uncertain parameters in the energyplus model and their uncertainty ranges. In addition to uncertainty quantification described in Chapter 3, we further investigated technical papers and industrial reports that analyze the performance of commercial products in the market (e.g., Energy-using Products reports) in order to quantify uncertain parameters in the fan and the pump system model.

Table 4.12 Uncertain parameters and their ranges in the energyplus model

Model Parameters	Base	Min	Max	Reference
Thermal Properties				
Concrete conductivity (W/m²·K)	2.15	1.94	2.37	see Chapter 3.3.1
Concrete density (kg/m³)	2400	2352	2448	
Concrete specific heat (kJ/kg·K)	0.90	0.68	1.12	
Concrete solar absorptance	0.68	0.60	0.76	
Concrete emissivity	0.91	0.87	0.95	
Insulation conductivity (W/m²·K)	0.03	0.0	0.03	
Insulation density (kg/m³)	43	42.14	43.86	
Insulation specific heat (kJ/kg·K)	1.21	0.91	1.51	
Plaster conductivity (W/m²·K)	0.16	784	0.18	
Plaster density (kg/m³)	800	823	816	
Plaster specific heat (kJ/kg·K)	1.09	0.82	1.36	
Plaster solar absorptance	0.40	0.34	0.46	
Plaster emissivity	0.90	0.86	0.94	
Glass conductivity (W/m²·K)	0.90	0.81	0.99	
Glass emissivity	0.84	0.76	0.92	
Glass solar transmittance	0.84	0.82	0.86	

Table 4.12 Continued

Internal Loads				
Lighting power density (W/m²)	15	11	19	measurements
Appliance power density multiplier	1	0.78	1.45	see Chapter 3.3.7
Occupant metabolic rate (W)	80	70	130	see Chapter 3.3.7
Control				
Indoor heating temperature (°C)	22	20	24	measurements/
Indoor cooling temperature (°C)	24	22	26	operation manuals
Ventilation				
Infiltration rate (1/h)	0.50	0.10	1.25	see Chapter 3.3.2
Discharge coefficient	0.68	0.60	0.75	see Chapter 3.3.3
Intercept c	-2.92	-3.80	-2.09	
Heating System				
Boiler nominal efficiency	0.86	0.84	0.88	Lazzarin (1986)
Cooling System				
Degradation coefficient	0.25	0.06	0.26	see Chapter 3.3.5
Fans				
Supply fan total efficiency	0.65	0.60	0.70	Radgen (2006)
Supply fan motor efficiency	0.82	0.75	0.87	de Almeida (2008)
Pumps				
Motor efficiency	0.77	0.70	0.83	de Almeida (2008)
Domestic Hot Water System				
Heater thermal efficiency	0.91	0.87	0.95	see Chapter 3.3.6

4.2.3.2 Parameter Screening

Table 4.13 shows the uncertain parameters, ranked in the order of dominance using the Morris method. The top five dominant parameters are the same as those ranked

by the normative energy model (shown in Table 4.11), and we thus calibrated the same five parameters with the energyplus model.

Table 4.13 Ranking of model parameters in the energyplus model by relative importance

Rank	Model Parameter
1	Intercept c for windows open
2	Indoor temperature during heating
3	Appliance power density
4	Infiltration rate
5	Discharge coefficient
6	Lighting power density
7	Boiler nominal efficiency
8	Indoor temperature during cooling
9	Insulation conductivity
10	Glass emissivity

4.2.3.3 Model Calibration

Figure 4.9 shows the posterior distributions of the five calibration parameters with the energyplus model. The posterior distributions from the energyplus model are quite the same as those from the normative model. This high coincidence proves the feasibility of the normative model to adequately capture a building as operated through the calibration process. However, this similarity does not necessarily guarantee that predictions between the two models are consistent because the normative model approximates the heat transfer processes in a building and may bias model outcomes. Hence, we will further investigate the feasibility of the normative model for retrofit analysis by comparing predictions and decisions derived by the normative model with those by the energyplus model.

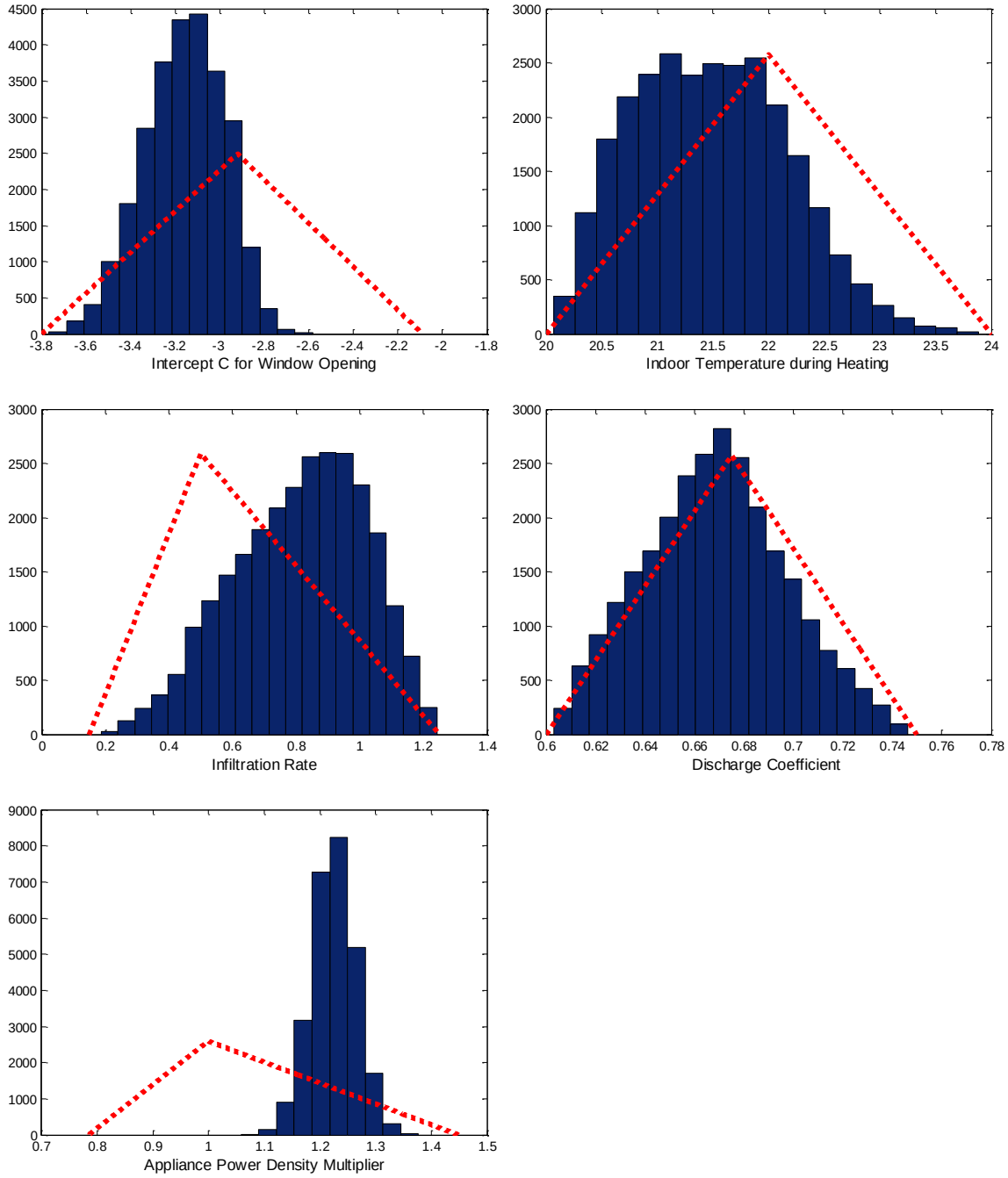


Figure 4.9 Posterior distributions of calibration parameters with the energyplus model (posterior-blue, prior-red)

4.2.4 Evaluation: Comparison between Normative and Transient Energy Model

Under the first criterion (accuracy of calibrated models), we evaluate the validity of the calibrated models in terms of agreements between predicted and monitored energy uses. Table 4.14 shows *CVRMSE* values of the uncalibrated models and the Bayesian calibrated models. The statistical measures indicate that the normative model, supported by Bayesian calibration, can predict the energy consumption as accurately as the calibrated energyplus model. Also, the validation measures tell that Bayesian calibration improves the accuracy of the baseline model by reducing the *CVRMSE* values by about 65 percent for both gas and electricity consumption. The *CVRMSE* values of the calibrated models are still higher than the 15% stipulated in ASHRAE Guideline 14 (for a model to be deemed valid). However, it should be noted that even with lower agreements with measured data Bayesian calibration models still outweigh deterministically calibrated models because they can quantify uncertainties remaining in the model and propagate them in model predictions.

Table 4.14 *CVRMSE* measures for uncalibrated and calibrated models

Type of Model	Gas	Electricity
before calibration		
Normative	0.95	0.38
Energyplus	0.94	0.23
after calibration		
Normative	0.34	0.14
Energyplus	0.39	0.15

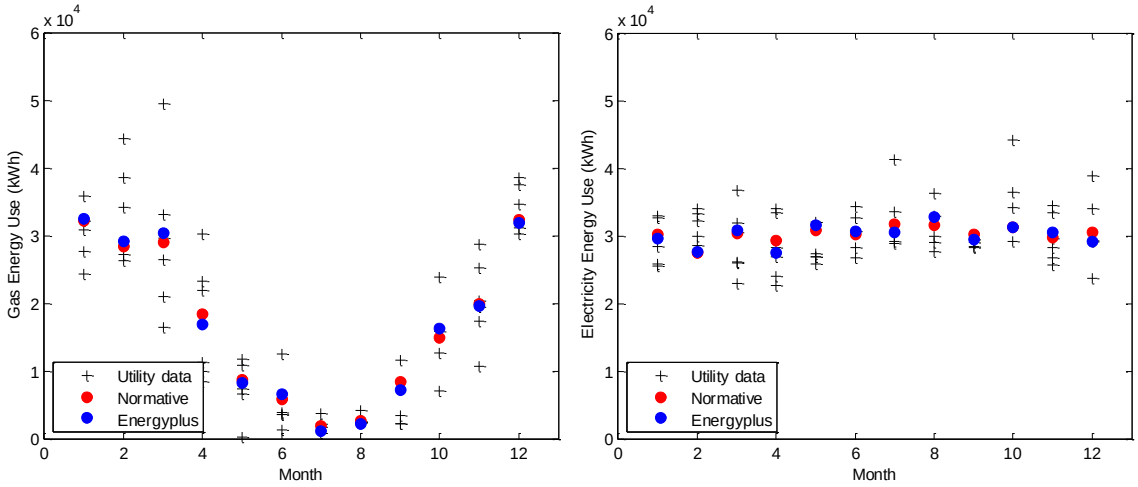


Figure 4.10 Predicted energy uses from the calibrated models against the monitored uses (left - gas energy use, right - electricity energy use)

We further compare the two calibrated models in the context of retrofit decision-makings. We exercise the calibrated models to evaluate six ECMs: (1) insulation upgrade, (2) window replacement, (3) infiltration air-tightening, (4) boiler upgrade (seasonal efficiency = 0.97), (5) air-conditioning upgrade (COP = 5), and (6) lighting upgrade (T-5 lamps). Table 4.15 summarizes uncertainty ranges for uncertain parameters pertaining to the six ECMs.

Table 4.16 lists the capital investment costs, including equipment and labor costs, estimated for the six ECMs based on the BCIS price book (BCIS, 2010). For energy costs, the gas price was fixed at 2.4 pence/kWh, and the electricity price at 8.6 pence/kWh (DECC, 2010b).

Table 4.15 Uncertain parameters and their ranges for the three ECMs

Parameters	Base	Min	Max	Reference
ECM1: Insulation Addition				
Insulation U-value (W/m²·K)	0.30	0.27	0.33	see Chapter 3.3.1
ECM2: Window Replacement				
Window U-value (W/m²·K)	1.41	1.27	1.55	see Chapter 3.3.1
Window solar transmittance	0.65	0.63	0.67	
Window emissivity	0.05	0.04	0.06	
ECM3: Air-tightening				
Infiltration rate reduction (%)	11	1	31	Jacobson (1986)
ECM4: Boiler Upgrade				
Boiler seasonal efficiency	0.97	0.95	0.98	see Chapter 3.3.4
ECM5: Air-conditioning Upgrade				
Degradation coefficient	0.25	0.06	0.26	see Chapter 3.3.5
ECM6: Lighting Upgrade				
Lighting power density (W/m²)	13	11	15	see Chapter 3.3.7

Table 4.16 Cost estimates of the six ECMs (in 1000£)

Retrofit Options	Base	Min	Max
ECM1: Insulation Addition	11	10.5	12
ECM2: Window Replacement	55	52	58
ECM3: Air-tightening	7.2	6.8	7.5
ECM4: Boiler Upgrade	3.4	3.2	3.6
ECM5: Air-conditioning Upgrade	3.2	3.0	3.3
ECM6: Lighting Upgrade	5.7	5.4	6.0

Under the second criterion (accuracy of model predictions), we compare SPT predictions from the normative model with those from the energyplus model. Table 4.17 shows the statistical results from the two-sample K-S tests that evaluate the hypothesis:

whether the two samples come from the same cumulative distribution. This hypothesis is rejected at the 5% significance level for all the ECMs except ECM 3. This result indicates that mostly the normative model may not yield probabilistic outcomes equivalent to those by the energyplus model. However, Figure 4.11 shows that overall the normative model predictions is similar to the energyplus model predictions. Nonetheless, the normative model tends to slightly overestimate potential energy-savings of ECMs 1 and 2, resulting in lower SPTs than the energyplus model. This prediction bias in the normative model will be further looked into with respect to its effect on decision-making.

Table 4.17 Two-sample K-S tests for model predictions from the normative and energyplus model

Retrofit Options	Two-sample K-S test	
	H	p-value
ECM1: Insulation Addition	1	0.00
ECM2: Window Replacement	1	0.00
ECM3: Air-tightening	0	0.16
ECM4: Boiler Upgrade	1	0.00
ECM5: Air-conditioning Upgrade	1	0.00
ECM6: Lighting Upgrade	1	0.01

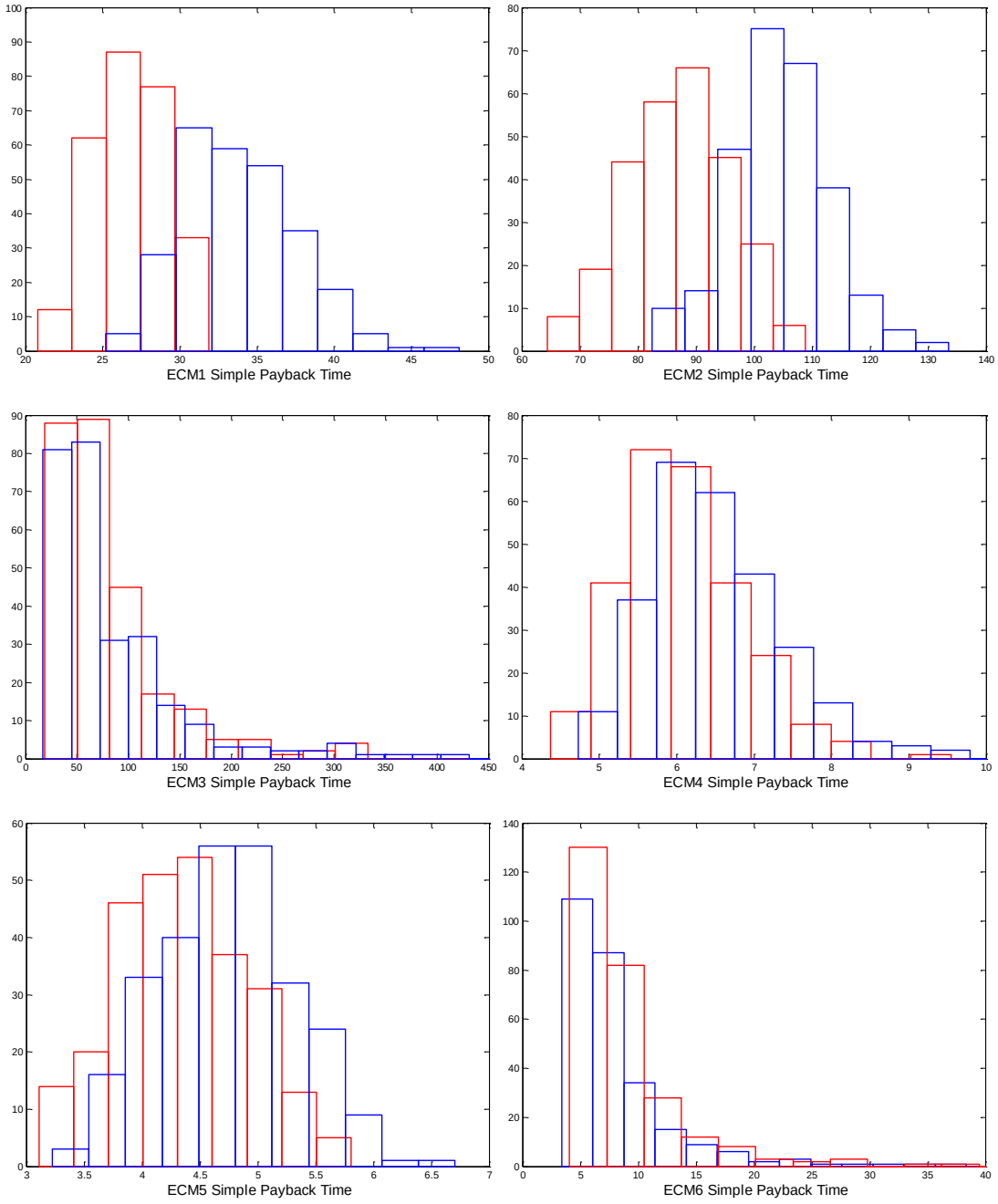


Figure 4.11 SPT predictions of the six ECMs (red - normative, blue- energyplus)

Finally, we evaluate the feasibility of the calibrated normative model by comparing decisions supported by the normative model with those by the energyplus model in a set of plausible decision-making scenarios. Table 4.18 lists the ranking of the six ECMs under the three scenarios that capture different decision-makers' willingness to

accept a certain level of risks. Unlike the first case study in which the preferred ECM differs depending on the scenario, the preferred ECMs are the same regardless of the scenario because their overall performance is far superior to the others. At this point it should be re-emphasized that information about underperforming risks of ECMs does not necessarily lead to better decisions but guarantee that final decisions sufficiently reflect decision-makers' intentions. In addition, the comparison demonstrates that the normative and the energyplus model results in the consistent ranking of the ECMs for all the scenarios. The consistent ranking in the second case study confirms that the normative model can adequately support retrofit decision-makings while accounting for major sources of uncertainty.

Table 4.18 Ranking of the three ECMs for the three decision-making scenarios

ECMs	Scenario 1	Scenario 2	Scenario 3
Normative			
ECM1	4	4	4
ECM2	6	5	5
ECM3	5	6	6
ECM4	2	2	2
ECM5	1	1	1
ECM6	3	3	3
Energyplus			
ECM1	4	4	4
ECM2	6	5	5
ECM3	5	6	6
ECM4	2	2	2
ECM5	1	1	1
ECM6	3	3	3

4.3 Discussion

4.3.1 Weather Data for Calibration

Generally it is recommended to use actual weather data for calibrating building energy models (ASHRAE, 2002). No doubt, actual weather data covering the same period as the metered energy consumption data provides the most reliable scenario for calibration. However, actual weather data is not always accessible. Hence, we investigate if the TMY data is good enough for the calibration. Figure 4.12 plots observed monthly outdoor temperatures over a three-year period against TMY temperatures. The plot demonstrates that the TMY temperatures well coincide with the average of the three-year observations. This implies that TMY data is good enough for the calibration when the calibration is based on monthly utility data over a multiple-year period.

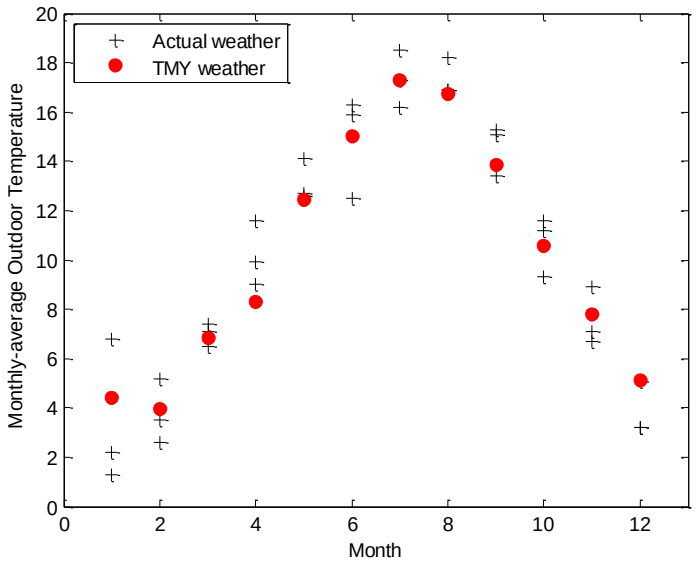


Figure 4.12 Three-year actual temperatures against TMY temperatures

4.3.2 Effects of Prior Estimates on Calibration Results

Bayesian calibration updates prior estimates of calibration parameters given observed data on building performance. Ideally, a large number of observations at various

levels (e.g., utility data, sub-metered data) can result in similar posterior distributions close to true values even without good prior estimates. However, in reality, metered energy consumption data (observations) is often only available for a limited period. Hence, it is expected that calibration results can be considerably influenced by the prior estimates. In order to investigate the effects of prior estimates on calibration results, we calibrate the model with two different prior distributions:

- Scenario 1: increases the upper and lower limits of the original prior distributions by 50% while maintaining the distribution shape.
- Scenario 2: uses uniform distributions within the limits specified for the original prior distributions.

Figure 4.13 overlays posterior distributions from Scenario 1 (red color) and Scenario 2 (blue color) against those from the original prior estimates (black color). Increasing the ranges of prior estimates results in wider ranges of the posterior distributions because the observations are insufficient to curtail wider uncertainty assigned in the prior distributions. However, except the spread, the two posterior distributions have similar distribution characteristics: both the distribution shpaes and the expected values are similar. On the contrary, change in the distribution shape significantly impacts the posterior distributions. With the uniformly distributed priors, the resulting posterior distributions are strongly weighted toward one bound. But, in both Scenario 2 and Original Scenario, the posteriors shift toward the same bound due to the same likelihood function given the monitored data. Particularly for the appliance power density, the three posteriors (Original, Scenario 1, and Scenario 2) are quite similar despite the different priors since the monitored data contains enough information to derive the postier estimate.

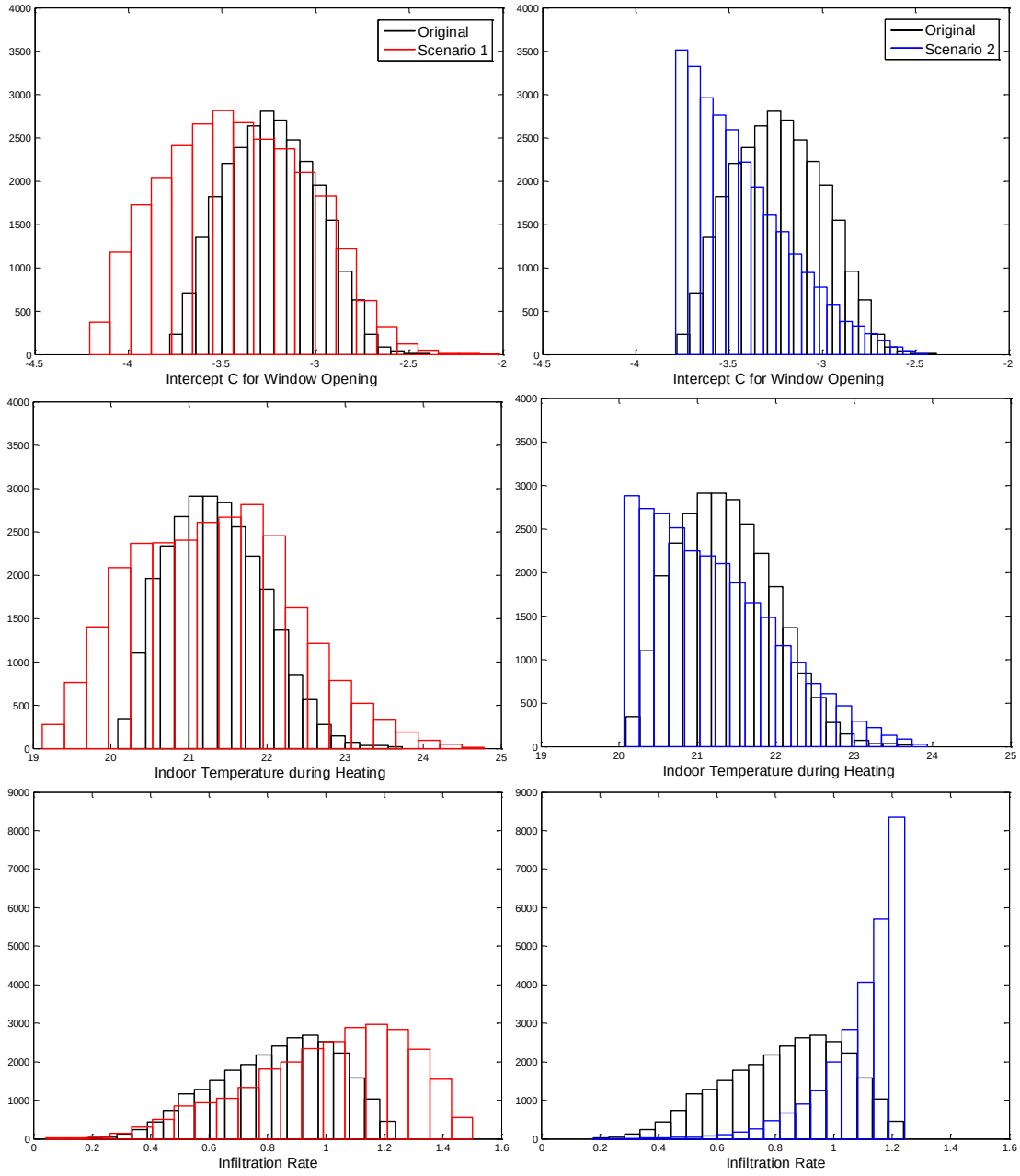
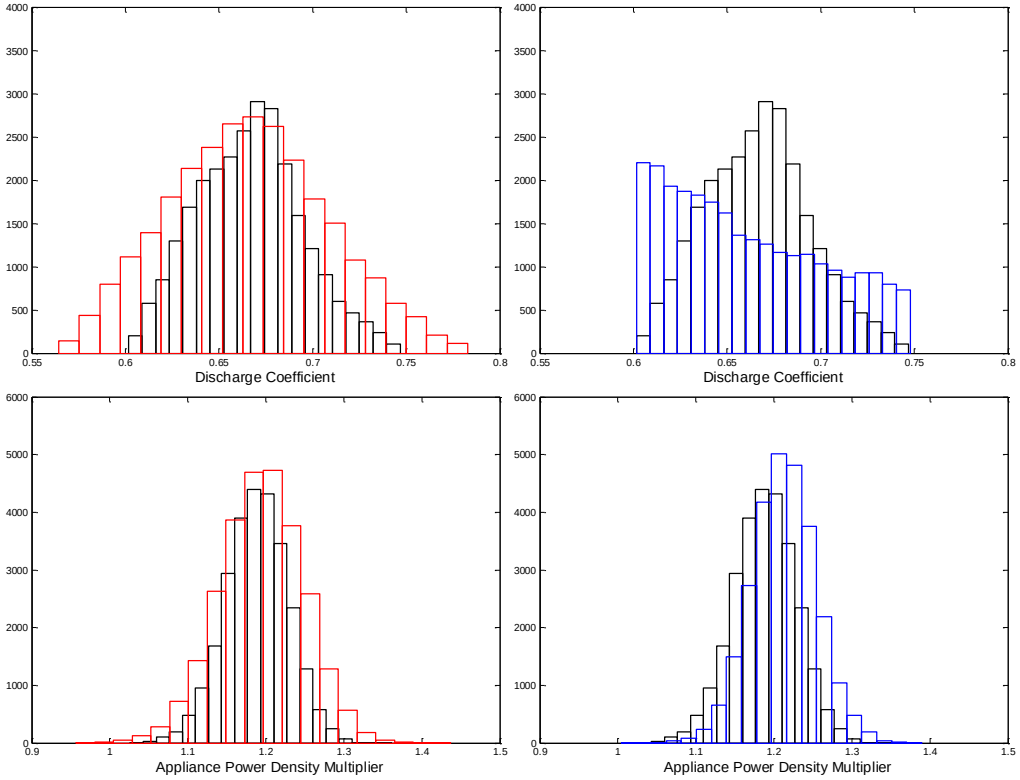


Figure 4.13 Posterior distributions of the five calibration parameters (black - from original priors, red - scenario 1 (from wider priors), blue - scenario 2 (from uniformly distributed priors))

Figure 4.13 Continued



In summary, Bayesian calibration can correct our prior beliefs about true calibration parameter values, but its results still significantly depend on the prior estimates. This relationship implies that prior estimation is important. One point to be emphasized is that prior estimates are set up based on collective expert knowledge and change only when there is additional knowledge in the process of prior uncertainty quantification. Then, given prior estimates are further refined through Bayesian calibration.

4.3.3 Comparison of Posterior Distributions from Case Studies

The calibration exercise enables us to learn from observations and derive better estimates of parameter values by changing prior uncertainty distributions. Resulting

posterior distributions from calibration exercises enhance our understandings of parameter values, and can be further utilized to set up better prior estimates for other case buildings. Beforehand we should inspect how distinctively parameter values differ depending on the case in order to investigate the potential of generalizing from case buildings to a large population of buildings.

Figure 4.14 depicts the posterior distributions of the four calibration parameters derived from the first case study (purple color) and the second case study (green color). The two case buildings have similar posterior distributions for the two parameters: indoor temperature during heating and discharge coefficient. The indoor temperature refers to spatially-averaged indoor temperature that typically deviates from a set-point temperature due to air stratification in spaces. Similar posteriors for the indoor temperature imply that the posteriors can be possibly applied to other typical office buildings. Another parameter, discharge coefficient, depends on the shape and the size of openings, and the consistent calibration results for this parameter suggest that these results can be also used for other office buildings that have typical types of windows although they do not change much from the original prior estimates.

On the contrary, the other two parameters, intercept c for windows open and infiltration rate, have noticeably different posterior distributions for each case building. Intercept c is part of the empirical formula that determine the proportion of windows open as the result of occupant control actions. Since window-opening behavior is triggered by various environmental and social factors, it is likely to differ depending on the building environment (e.g., cultural, organizational, and operational settings). Infiltration rate also differs building by building. Indeed, infiltration rate of a building is determined by case-specific factors such as the construction quality of a building and pressure difference between the outside and the inside of a building. Hence, the calibration results for these two parameters may not be generalized for all other cases.

Before the generalization, proper categorization of buildings should be studied through extensive case studies that enable deep understanding of main factors that impact building energy performance.

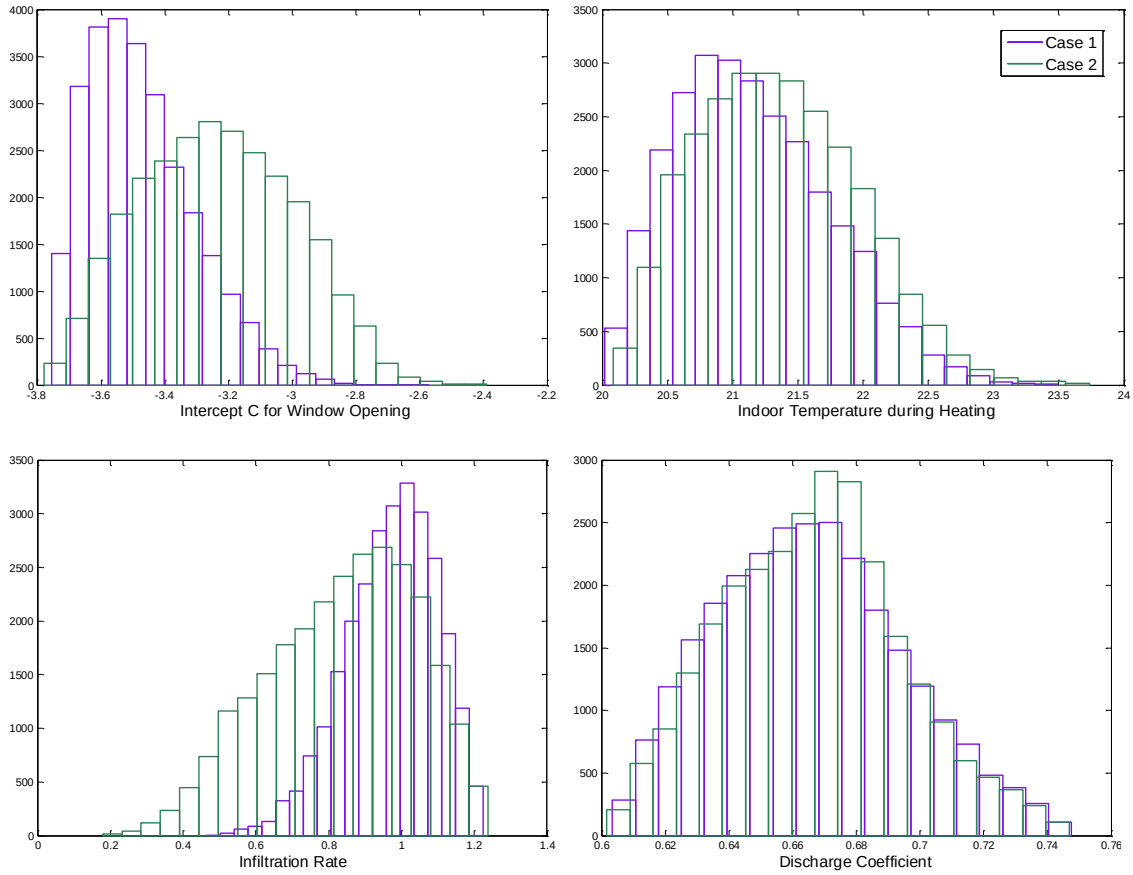


Figure 4.14 Posterior distributions of the four calibration parameters from the two case studies (purple - the first case, green - the second case)

4.4 Concluding Remarks

The objective of this chapter is to verify the first hypothesis: whether normative models have a reasonable level of model resolution to adequately evaluate ECMs and lead to decisions without compromising the degree of confidence in decisions. This chapter demonstrated the feasibility of the new methodology (normative models enhanced by Bayesian calibration) on the retrofit analysis through the case studies that

compared model outcomes and decisions from the calibrated normative models with those from the calibrated energyplus models. The case studies verified that normative models can adequately serve as a tool to support decision-makings under uncertainty as the replacement of transient simulation models when they are supported by Bayesian calibration.

CHAPTER 5 CONTRIBUTIONS OF THE NEW FRAMEWORK IN PRACTICE OF ENERGY-EFFICIENCY PROJECTS

5.1 Introduction

This chapter demonstrates the capability of the new framework to rigorously support uncertainty analysis for risk-conscious decision-making. Chapter 1.2 presents the importance of risk analysis in retrofit decision-making, particularly in the context of Energy Service Companies (ESCOs). Chapter 1.2 also summarizes key limitations of the current methods to support risk analysis: they follow the deterministic approach that ignores uncertainty in the retrofit analysis, and rely on experts' subjective decisions to quantify uncertainty in energy-saving estimates. In order to adequately evaluate all available ECMs with the confidence level, an analysis method should quantify risks in energy savings from ECMs while accounting for uncertainty in the analysis process.

The new methodology proposed in this thesis can serve as a formal method to support rigorous risk analysis in the decision-making stage. Bayesian calibration models quantify uncertainties in the baseline model, and further incorporate additional uncertainties coming from ECMs in order to compute probabilistic predictions of retrofit performance. The resulting probabilistic outcomes systematically capture the effects of all major sources of uncertainty on the outcomes, which can be naturally translated to quantify risks of underperformance associated with ECMs. Indeed, energy-efficiency risks come from a broad range of uncertain factors that often confound ECM performances and increase the volatility in investment decisions (Mills, 2006). Hence, without taking all sources of uncertainty into account, one cannot realize the inherent risks one may undertake for his/her decisions.

This chapter investigates the importance of formal uncertainty analysis on retrofit decision-making. First, the chapter looks into current standard methods with respect to how they handle uncertainty in the process of determining energy savings. Then, the chapter revisits the first case study to illustrate the role of Bayesian calibration models in risk-conscious decision-making. The case study compares outcomes and decisions derived by the Bayesian calibration model with those derived by methods used in current practice in the context of plausible decision-making scenarios.

5.2 Current Practice from the Perspective of Uncertainty

This section describes primary concepts and methods employed in the current practice concerning uncertainty in energy-savings estimates. Energy-efficiency projects in the ESCO industry typically follow the International Performance Measurement and Verification Protocol to estimate ECM energy-savings (Hansen, 2004; IPMVP, 2010). The IPMVP offers several methods for determining energy savings from ECMs for a building, but all of them follow a deterministic approach; they set methods for computing an absolute value of energy-savings from a set of ECMs without quantifying any risks expressing potential underperformance of ECMs not resulting in energy savings as projected in the energy performance contract.

In addition to the IPMVP, ASHRAE Guideline 14 (ASHRAE, 2002) provides guidelines and calculation methods for retrofit analysis in the US. These guidelines recommend deriving a deterministic model best fit to monitored data and estimating a single energy-saving value by subtracting projected (calculated) energy use during the post-retrofit period from baseline energy use during the pre-retrofit period. The deterministic results can be regarded as optimistic because they do not account for sources of uncertainty that can potentially cause ECMs to underperform the expected performance improvements.

Despite the deterministic approach, ASHRAE Guideline 14 implicitly acknowledges the importance of uncertainty (or rather variability) particularly for validating analysis results. The ASHRAE guideline attempts to quantify three types of uncertainty: (a) modeling uncertainty, (2) measurement errors and (3) sampling uncertainty. In the guideline, modeling uncertainty refers to how well a baseline model captures variability in measured data. The guideline defines modeling uncertainty U_M in terms of the coefficient of variation of the root mean square error (*CVRMSE*) shown in Equation 5.1; P_i is predicted value, O_i is observed value in period i, and the average of all observations is denoted by $\bar{O}$. The guideline requires that *CVRMSE* should be less than 15% for monthly calibration data and 30% for hourly calibration data in order that a calibrated model is deemed valid.

$$U_M = \frac{\sqrt{\sum_{n=1}^{n}(O_i - P_i)^2/n}}{\bar{O}} \tag{5.1}$$

Measurement errors include errors in monitoring both energy use data and independent variables. Measurement equipment error U_E depends on the measurement accuracy of an instrument, and Equation 5.2 calculates the overall instrument error over a range of measured data. r_{rating} denotes reading of an instrument at a point at which its manufacturer verifies its relative error ($RE_{equipment}$) through full-scale tests, and $\bar{r}$ denotes the mean value of a series of instrument readings. Independent variable error U_{iv} refers to error in monitoring independent variables (e.g., outdoor weather conditions) during the post-retrofit period. The ASHRAE guideline recommends quantifying the spread of the additional uncertainty by calculating the difference between calculated savings with the maximum variable values and those with the minimum variable values. The guideline also stipulates the conditions in which measurement equipment error and independent variable error can be ignored: measurement equipment error is assigned zero when metered energy use data is from utility bills: independent variable error is assigned

zero when the weather data used for the post-retrofit analysis comes from a government-operated weather reporting service.

$$U_E = \frac{\sqrt{\sum_{i=1}^{C}\left(RE_{equipment} \times r_{rating,i}\right)^2}}{\sum_{i=1}^{C} \bar{r_i}} \tag{5.2}$$

Sampling uncertainty refers to errors in estimating energy use for the total population from a sample of units. For the pragmatic reason to reduce monitoring costs, projects often measure energy uses in a sample of units, and derive the average energy uses in the entire set of units from the sample. For instance, for a building with ten identical floors, projects measure the lighting electricity use in a random floor, and use the measured sample to estimate the average lighting electricity use in the whole building. Equation 5.3 computes sampling uncertainty U_S; Q is a number of the total units, q is a number of the units selected for measurement; y_i is a monitored value in period i from a randomly sampled unit, and $\bar{y}$ is the estimated mean of the total population Q.

$$U_S = \frac{100}{\bar{y}} \times \sqrt{\left(1 - \frac{q}{Q}\right)\left(\sum_{i=1}^{n} \frac{(y_i - \bar{y})^2}{(q-1)}\right)\left(\frac{1}{q}\right)} \tag{5.3}$$

ASHRAE Guideline 14 provides an empirically driven equation (Equation 5.4) that calculates overall savings uncertainty U from the three types of uncertainty. Savings uncertainty U refers to relative uncertainty in the estimated energy savings, which is defined as the standard deviation ΔE_{saving} divided by the energy-savings estimate E_{saving}. F is the ratio of energy-saving estimate to baseline energy use, and t is a t-statistics for the expected confidence level. m is the number of predictions in the post-retrofit period, and n is the number of observations in the pre-retrofit period. The guideline stipulates that calculated standard deviation should be lower than 50% of the annual savings estimate for the projected energy-savings to be valid.

$$U = \frac{t}{F}\sqrt{\frac{{U_M}^2}{m} \times \left(1.6 + \frac{3.2}{n}\right) + {U_S}^2 + {U_E}^2 + {U_{iv}}^2} \quad (5.4)$$

Equation 5.4 simplifies to Equation 5.5 for typical projects in which (a) no sampling is done ($U_S = 0$), (b) utility bills are used for metered energy uses ($U_E = 0$) and (c) government-published weather data is used for independent variables ($U_{iv} = 0$). The equations are initially developed to validate the baseline model and estimated energy-savings. Furthermore, the ASHRAE guideline implies that the equations can be used as means for parties in the ESCO projects to reach mutual agreements on risks associated with retrofit implementations.

$$\frac{\Delta E_{saving}}{E_{saving}} = t \times \frac{1.26 \times CVRMSE}{F} \times \sqrt{\left(\left(1 + \frac{2}{n}\right) \times \frac{1}{m}\right)} \quad (5.5)$$

ASHRAE Guideline 14 thus provides empirical tests that quantify the confidence in an energy-savings estimate for a building retrofit. However, this method is not suitable for quantifying risk for several reasons. First, ΔE_{saving} should be derived in principle from a probabilistic distribution. Second, the method cannot be used to quantify uncertainties associated with the proposed ECMs. Hence, it leads to the same magnitude of risk for any retrofit option although each retrofit option most likely contains different level of underperforming risks. Furthermore, the empirical relationship shown in the equations is designed for a very specific definition of energy-savings. Hence, it does not allow translation of computed energy-saving uncertainty into other risk measures. In fact, energy retrofit is an investment, and retrofit projects often employ cost-effectiveness measures such as cost/benefit ratio and simple payback time for decision-making. However, the current method cannot provide information about risks in these forms.

5.3 Case Study

This chapter revisits the first case study to compare outcomes derived by the proposed methodology with those by the current methods in practice. Chapter 4.1 summarizes analysis outcomes calibration results and predictions for the ECMs from the analysis process supported by the proposed methodology. As the counterpart, we employ the calibrated model approach endorsed by ASHRAE Guideline 14, which reflects the standard practice of retrofit projects (refer to Chapter 2.1).

5.3.1 Calibration Results

This section compares calibration results between the Bayesian approach and the deterministic approach used in current practice. As the standard practice, we calibrate the energy model of the same building following the standard calibration procedure in ASHRAE Guideline 14. The calibration process is more ad-hoc; the modeler selects calibration parameters according to his/her knowledge and experience, and manually tests different values until there is good match between model predictions and utility data. The high involvement of experts has been recognized as a major problem that makes the calibration process ad hoc and non-scientific (Reddy, 2006). In this study, we leave out the influence of experts on calibration outcomes, but focus on the effects of following a deterministic calibration process on risk analysis.

We use the same four parameters used for the Bayesian calibration to calibrate the energy model as per the deterministic method. The upper and lower limits of parameter values are also the same as those specified in the prior distribution functions of the Bayesian calibration. Although in practice experts heuristically tune calibration parameters, we employ an optimization algorithm because current calibration practices are in principle a heuristic version of deterministic optimization (in terms of minimizing

the discrepancy between model outcomes and observations). Figure 5.1 shows the values of the calibration parameters derived by the deterministic calibration in comparison to the posterior distributions by the Bayesian calibration. All values are at the extreme upper or lower limits and quite far from the expected value derived from the Bayesian calibration process. This large difference between the two calibration results implies that the deterministic calibration process cannot guarantee that resulting parameter values accurately correspond to actual building conditions.

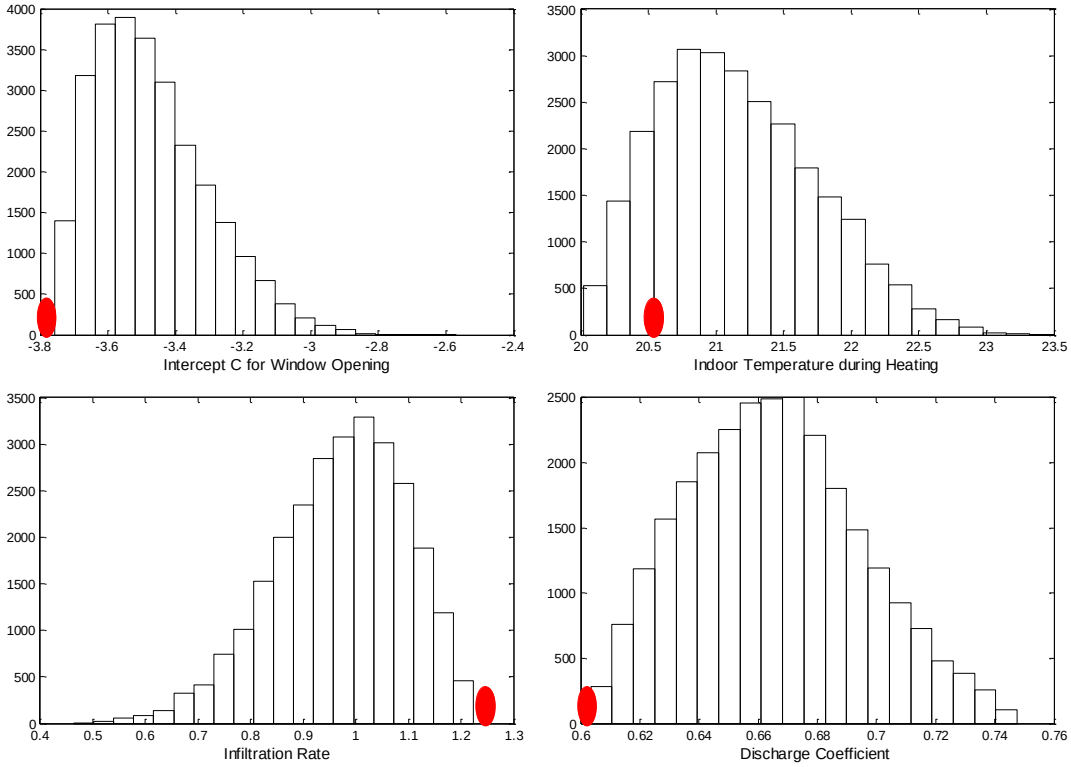


Figure 5.1 Deterministic calibration results for the four parameters (red dot) against the posterior distributions from Bayesian calibration

We evaluate the validity of the calibrated models by *CVRMSE* used in the ASHRAE guideline. It should be mentioned that most other validation metrics in the literature summarized well in (Krause, 2005; Legates, 1999) are similar goodness-of-fit measures between model predictions and observations. Table 5.1 shows the *CVRMSE*

values of the (a) uncalibrated model, (b) model calibrated based on the Bayesian approach, and (c) deterministically calibrated model. It shows that the calibration process enhances the accuracy of the baseline model as it reduced the *CVRMSE* value by half. Still, both calibration processes fail to satisfy the validation criteria in the ASHRAE guideline; *CVRMSE* should be within 15%. However, it should be noted that the guideline evaluates the accuracy of deterministically calibrated models. Even with lower agreements with measured data, calibrated models based on a Bayesian approach still outweigh deterministic models because they can quantify uncertainties in model predictions.

Table 5.1 Validation measures for uncalibrated model, Bayesian calibration model, and deterministically calibrated model

Calibration Method	*CVRMSE*
Un-calibrated model	0.34
Bayesian Calibration	0.17
Deterministic Calibration	0.18

The *CVRMSE* also indicates that the deterministically calibrated model predicts the baseline energy use as closely as the Bayesian calibration model does. However, the standard validation measures (such as *CVRMSE*) cannot truly evaluate whether a model is accurate enough. Calibration techniques identify parameter values that compute predicted energy consumption closely matching monitored energy consumption. Hence, when calibrated models are evaluated against the monitored data, they seem quite valid. However, chose parameter values that attain good agreements between predicted and monitored energy use do not guarantee that they correspond to actual building conditions. Especially, the deterministic calibration technique results in one single solution that maximizes the agreement while ignoring many feasible parameter values that may have higher likelihoods. A deterministically calibrated model can thus potentially bias the

effects of certain ECMs on energy savings, and ultimately lead to wrong decisions. On the other hand, the Bayesian calibration process results in a set of plausible parameter values under which ECMs are assessed. Hence, the Bayesian process can enhance the reliability of model predictions by not only providing uncertainty in calibration parameter values but also assuring the reliability of baseline models.

5.3.2 Retrofit Decision-making

This section demonstrates the value of the Bayesian calibration approach to support retrofit decision-making. In order to investigate the positive effects of risk information on decisions, we deploy both the Bayesian calibrated and deterministically calibrated model in plausible decision-making scenarios, and compare their predictions and resulting decisions.

We evaluate three ECMs: (1) upgrading insulation, (2) replacing windows, and (3) improving air-tightness. Table 4.6 summarizes uncertainties in model parameters associated with these three ECMs. Each ECM brings in varying amounts of uncertainty; for example, infiltration reduction has a higher degree of uncertainty than other ECMs, as the performance of air-tightening techniques depends on diverse factors including workmanship, weather conditions, and indoor conditions. A point to be noted is that a high level of uncertainty in an ECM does not necessarily result in large uncertainty in energy-saving estimates as different ECM parameters have different influence on the energy outcomes. The magnitude of uncertainty in energy-saving is derived through the energy model.

In the decision-making stage, one needs to select a performance indicator to express and quantify the decision-makers' rational objectives. In this analysis, we employ the following performance indicators commonly used in current retrofit projects:

- Performance Indicator 1: annual energy-savings leading to annual utility cost reduction

- Performance Indicator 2: simple payback time (SPT) concerning energy-saving returns that recover initial investment costs

In addition, one needs to a measure to capture the willingness to accept certain levels of risk. For each performance indicator, we compare three scenarios that express different levels of risk-consciousness.

- Scenario 1: represents conventional practice that is not concerned with risks but only with overall performance. We use expected values for this measure.
- Scenario 2: represents guaranteed savings in performance contracts. Guarantees are translated into 5-quantile for annual energy-savings and 95-quantile for simple payback time.
- Scenario 3: represents one of the existing risk measures proposed for actuarial pricing of energy-efficiency projects (Mathew, 2005). We use a saving curve score defined as the mean savings divided by the standard deviation of savings estimates.

5.3.2.1 Decisions by energy-saving measure

Table 5.2 shows the mean and the standard deviation of annual energy-savings predicted by the proposed probabilistic approach and the standard deterministic approach for each ECM. The probabilistic approach used 271 simulation runs to propagate uncertainties in ECM parameters using Latin Hypercube Sampling (Wyss, 1998). As a result, we obtain the distribution of annual energy-savings. With the deterministic approach (following ASHRAE Guideline 14), we obtain a single annual energy-saving estimate and its standard deviation. The probabilistic approach results in different magnitudes of energy-saving uncertainty (ΔE_{saving}) for the three ECMs; ECM 3 has a higher risk of underperformance than ECM 1 and 2. However, the ASHRAE 14 method

computes the same degrees of validity for all ECMs because the calculated uncertainty (or standard deviation in this case) depends only on the *CVRMSE* of the calibrated model. Hence, it is concluded that the ASHRAE guideline cannot adequately quantify risks associated with ECMs.

Table 5.2 Predictions by the proposed and the standard approaches (in *kWh*)

	Probabilistic		Standard	
	E_{saving}	ΔE_{saving}	E_{saving}	ΔE_{saving}
ECM1	21,354	1,486	20,417	24,810
ECM2	76,745	5,423	66,623	24,810
ECM3	33,469	15,687	30,373	24,810

Table 4.9 shows the ranking of the ECMs for the three risk-consciousness scenarios. In the proposed method, the ranking of ECMs differs depending on the scenario. In contrast, the ASHRAE 14 method results in the same ranking regardless of the scenario: window replacement being the most beneficial, followed by air-tightening and insulation improvement. It should be noted that the two methods lead to the same ranking for scenario 1, which is expected since scenario 1 is only concerned with overall performance rather than risk. However, the ECMs are ranked differently when scenario 2 and 3 are considered. When guaranteed energy savings are required (the most risk-averse attitude), insulation improvement ranks highest, and for the moderate risk-averse scenario (scenario 3), window replacement is most suitable. The preliminary conclusion through this comparison is that the ASHRAE guideline provides poor support for risk analysis although it may be adequate for analysing overall performance of ECMs.

Table 5.3 Ranking of the three ECMs by energy-savings

ECMs	Scenario 1	Scenario 2	Scenario 3
Probabilistic			
ECM1	3	1	2
ECM2	1	2	1
ECM3	2	3	3
Standard			
ECM1	3	3	3
ECM2	1	1	1
ECM3	2	2	2

5.3.2.2 Decisions by cost-effectiveness measure

This section compares decisions supported by the two methods with respect to simple payback time (SPT). In addition to technical uncertainty in ECM performances, the cost-effectiveness measure should also quantify uncertainty in economic factors such as fuel costs, labor costs, and equipment costs. Table 4.7 lists estimated investment costs associated with labor and equipment and the range of uncertainty in the estimated costs. For fuel costs, we fix gas price at 2.4 pence/kWh (DECC, 2010b).

Table 5.4 shows the ranking of ECMs for the three risk scenarios when SPT is the performance indicator. For scenario 1, the ranking by cost-effectiveness (SPT) differs from that by energy-savings. Although ECM 2 (window replacement) is likely to result in highest energy-savings, its investment cost is also far higher than the other options; the mean SPT is 32 years for improving insulation, 121 years for window replacement, and 18 years for air-tightening. Using the proposed method, risk-averse decision-makers select ECM 1 (referring to scenarios 2 and 3) while those who are willing to accept higher levels of risks in obtaining potentially higher energy-savings select ECM 3. The ASHRAE 14 method cannot compute uncertainty in the cost-effective measures since it

provides an empirically derived formula to calculate uncertainty in energy-saving estimates. In other words, current standard practice does not support analysis of technical and financial risks associated with ECMs.

Table 5.4 Ranking of ECMs by simple payback time

ECMs	Scenario 1	Scenario 2	Scenario 3
Probabilistic			
ECM1	2	1	1
ECM2	3	3	3
ECM3	1	2	2
Standard			
ECM1	2	-	-
ECM2	3	-	-
ECM3	1	-	-

5.4 Concluding Remarks

Although quantitative risk analysis is essential in energy-efficiency projects, current practice does not offer an adequate approach to quantify uncertainty. Instead it is solely based on a deterministic calibration of building energy models and their use to evaluate and compare ECMs. In order to tackle the limitation of current practice, we have presented a probabilistic risk analysis methodology based on Bayesian calibration. Bayesian calibration models can take all sources of uncertainties into account. The resulting energy model supports probabilistic risk analysis according to decision-makers' objectives and appetite for risk. The case study demonstrates how Bayesian calibration models enhance current practice by offering a formal method that supports decision-making under uncertainty.

CHAPTER 6 CONCLUSIONS AND FUTURE WORK

6.1 Summary and Conclusions

Despite the increasing need to improve the energy efficiency of the existing building stock, the current methods are not capable to support retrofit decision-makings at large scale with adequate risk management due to the two major drawbacks: (1) they cannot scale to support large-scale analysis due to low modeling efficiencies and high reliance on expertise: (2) they cannot adequately support risk-conscious decision-making because they deterministically calibrate an energy model and derive a single prediction of ECM savings. Hence, current methods cannot scale up to large portfolio of buildings and cannot support risk-conscious decision-making in retrofit projects.

In order to overcome these limitations, this thesis proposed a scalable, adaptable methodology that is suitable for large-scale retrofit analysis by enhancing the cost-effectiveness and objectivity of the modeling process. The proposed methodology is based on normative models and Bayesian calibration. In the context of large-scale retrofit projects, the normative model can provide the following strengths:

- The normative model enables modeling a large portfolio of buildings while greatly reducing modeling burdens (i.e., data collection, modeling, and computation).
- The normative model can extensively assess feasible ECMs to select the optimal mix of retrofit technologies.
- The normative model does not require modeling expertise, and thus makes the modeling process transparent.

In addition, Bayesian calibration can enhance retrofit decision-making under uncertainty by providing the following strengths:

- Bayesian calibration enhances the reliability of the normative model by tuning important uncertain parameters in the model to represent the actual building operations.
- Bayesian calibration results in calibrated models that are suitable to uncertainty analysis. Calibrated models provide information about underperforming risks of ECMs while taking into account a full spectrum of uncertainty sources.
- The proposed calibration procedure is designed to objectively quantify uncertainty and select a set of calibration parameters with respect to their importance on model outcomes.

This thesis verified the two major hypotheses: (1) feasibility of the proposed methodology on retrofit applications and (2) the importance of uncertainty analysis on retrofit decision-makings. Chapter 4 verified through the case studies that lower resolution of normative models can correctly evaluate ECMs and derive consistent results as energyplus models when they are supported by Bayesian calibration. Chapter 5 demonstrated the capability of the proposed methodology to support risk management particularly in the context of the ESCOs industry.

6.2 Future Work

6.2.1 From Methodological Perspective

This thesis focused on developing the scalable methodology suitable for large-scale retrofit analysis and verifying its feasibility through the case studies. However, in

order to improve the applicability of the proposed methodology into standard practice, we need to resolve the following issues:

- Feasibility of normative models: this thesis verified the feasibility of the proposed methodology through the two case studies that are limited to office buildings. Therefore, more case studies are necessary to confirm the feasibility of normative models across various buildings (e.g., building function, building design, system design).
- Versatility of Bayesian Calibration: The current Bayesian calibration module is based on Kennedy and O'Hagan's framework, and its applicability is currently limited to the cases in which the source of measured data is at one building level. Hence, the Bayesian calibration module should be extended to calibrate a model with various sources of sparse monitored data (e.g., containing a mix of building specific as well as portfolio-aggregated consumption data).
- Extension of energy-efficiency risks: this thesis provides the analysis framework that is ready to incorporate all sources of uncertainty for ECM predictions. Nonetheless, the sources of uncertainty in the case studies have been limited to physical properties, equipment performance, and investment costs. The case studies ignore other uncertainties such as system degradation over the lifetime and detailed economic factors. In order to correctly evaluate the volatility of ECMs, we need to further quantify the full spectrum of uncertainties related to performance and financial risks of energy retrofit projects and ECM selection.
- Performance indicators for decision-making: this thesis evaluates candidate ECMs solely with respect to their saving performance (e.g., energy-savings, simple payback time). However, this measure typically understates the value

of energy retrofit projects since it does not capture indirect benefits stemming from non-energy impacts: increasing real estate value, improving indoor air quality, and extending equipment life. Mills (2009) stated based on the survey that non-energy benefits are part of desires initiating the projects and they are difficult to be quantified. Nevertheless, in order to arrive at rational decisions, retrofit projects should use adequate performance indicators that reflect decision-makers' preferences accurately.

6.2.2 From Pragmatic Perspective

In the context of large-scale retrofit projects, the proposed methodology can support the two steps: (1) it efficiently evaluates a large set of buildings to identify buildings that need energy-efficiency improvements: (2) for identified buildings it extensively assesses feasible ECMs to select the optimal mix of ECMs according to decision-makers' objectives. In order to strengthen the practicality of the methodology, we need to accomplish the following tasks:

- Intelligent interface for modeling: the current normative calculation module has full functionalities to evaluate available retrofit interventions for any kind of buildings. However, it still requires manual efforts to link information about model parameters to the calculation module for each building or each what-if scenario. Hence, it requires to develop dashboards to link available data to the calculation module and evaluate performances of buildings simultaneously as testing city-scale improvement measures.
- Automatic calibration process: this thesis proposed the calibration and retrofit analysis framework that can potentially be used without deep modeling and calibration expertise. In the current stage, however, the proposed method still relies heavily on experts' judgments especially in the choice of calibration

parameters, quantification of their prior distributions, and quantification of uncertainties in other parameters. The analysis process will remain dependent on experts until a repository for standard estimates of parameter uncertainty for variety of building cases becomes available. The development of such a comprehensive database requires an extensive effort and will rely on collaboration within the research community. Only through such effort, model calibration for retrofit analysis will become consistent and transparent, and ultimately automated.

- Model-based benchmarking: normative models are initially designed to benchmark buildings-as-designed. In principle, we should evaluate individual buildings solely in terms of their energy performance as excluding the effects of extraneous factors ((e.g., building function, occupancy schedule, operation schedule) on energy consumptions. CEN-ISO standards (CEN prEN 15217, 2005) summarize procedures to define references and benchmark buildings for certification. The analysis framework needs to incorporate the benchmarking mechanism based on the normative model to correctly rank buildings and identify those that need energy performance improvements.

REFERENCES

Ahmad, M., and Culp, C. H. (2006). Uncalibrated building energy simulation modeling results. *HVAC&R Research*, 12(4):1141-1155.

ANSI/AHRI Standard 210/240 (2008). 2008 Standard for performance rating of unitary air-conditioning & air-source heat pump equipment. *Air-Conditioning, Heating, and Refrigeration Institute*, Arlington, VA.

ASHRAE (2002). ASHRAE Guideline 14 - measurement of energy and demand savings. *American Society of Heating, Refrigerating, and Air-Conditioning Engineers*, Atlanta, GA.

ASHRAE (1999). ASHRAE Handbook: HVAC Applications. *American Society of Heating, Refrigerating, and Air-Conditioning Enginners*, Atlanta, GA.

ATTMA (2010). Technical standard L2: measuring air permeability of building envelopes (non-dwellings). *Technical Report*, The Air Tightness Testing & Measurement Association.

BCIS (2010). *BCIS Wessex Alterations & Refurbishment Price Book*. Building Cost Information Service of Royal Institution of Chartered Surveyors.

Borgeson, S., and Brager, G. (2008). Occupant control of windows: accounting for human behavior in building simulation. *Technical Report*, Center for the Built Environment, University of California, Berkeley.

Brown, J., and Exstrum, B. (2004). Rebuild America program scope of work. *Final Technical Report*. Aspen Systems Corporation.

City of Chicago Climate Action (2010). Chicago Climate Action Plan. Available from http://www.chicagoclimateaction.org/e

CEN, prEN 15203/15315 (2006). Energy performance of buildings – overall energy use, CO2 emissions and definition of energy ratings.

CEN, prEN 15217 (2005). Energy performance of buildings - methods for expressing energy performance and for energy certification of buildings.

CIBSE, Guide A (2006). *Guide A: Environmental Design*. Chartered Institution of Buildings Services Engineers, London.

CIBSE, TM23 (2000). *CIBSE technical memorandum TM23: testing buildings for air leakage*. Chartered Institution of Buildings Services Engineers, London.

de Almeida, A. T., Ferreira, F. J. T. E., Fong, J., and Fonseca, P. (2008). EUP Lot 11 Motors. *Final Report*, University of Coimbra.

DECC (2010a). Digest of United Kingdom Energy Statistics. *Technical Report*, Department of Energy and Climate Change.

DECC (2010b). Electricity Market Assessment. *Technical Report*, Department of Energy and Climate Change.

Department of Energy (2009). DOE to fund up to $454 million for Retrofit Ramp-Ups in energy efficiency. Available from http://energy.gov/articles/doe-fund-454-million-retrofit-ramp-ups-energy-efficiency

de Wit, S. (2001). Uncertainty in predictions of thermal comfort in buildings. *Ph. D. Thes*is, Delft University of Technology.

de Wit, S., and Augenbroe, G. (2002). Analysis of uncertainty in building design evaluations and its implications. *Energy and Buildings*, 34:951-958.

Dunn, G., and Knight, I. (2005). Small power equipment loads in UK office environments. *Energy and Buildings*, 37:87-91.

EERE (2003). Rebuild America program brief. Available from http://apps1.eere.energy.gov/buildings/publications/pdfs/rebuild_america/essbrief1003.pdf

EIA (2011). Energy explained: your guide to understanding energy. Available from http://www.eia.gov/energyexplained/index.cfm

EIA (2003). 2003 Commercial Buildings Energy Consumption Survey. U.S. Energy Information Administration, Washington, DC. Available from http://www.eia.gov/emeu/cbecs/contents.html

EN 15242 (2007). Ventilation for buildings - calculation methods for the determination of air flow rates in buildings including infiltration.

EN 15243 (2007). Ventilation for buildings - calculation of room temperatures and of load and energy for buildings with room conditioning systems.

EN 15316-2-3 (2007). Heating systems in buildings - method for calculation of system energy requirements and system efficiencies - part 2-3: space heating distribution systems.

EPA (2008). EPA's Report on the Environment. *Final Report EPA/600/R-07/045F*. United States Environmental Protection Agency.

EN 1SO 13786 (2007). Thermal performance of building components - dynamic thermal characteristics - calculation methods.

Fanney, A. H., and Dougherty, B. P. (1996). The thermal performance of residential electric water heaters subjected to various off-peak schedules. *Journal of Solar Energy Engineering*, 118:72-80.

FEMP. (2011). Federal Energy Management Program. Available from http://www1.eere.energy.gov/femp/.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004). *Bayesian Data Analysis*. Chapman & Hall/CRC, USA.

Goldman, C. A., Osborn, J. G., Hopper, N. C., and Singer, T. E. (2002). Market trends in the U.S. ESCO industry: results from the NAESCO database project. *Technical Report*. Lawrence Berkeley National Laboratory.

Guillas, S., Rougier, J., Maute, A., Richmond, A. D., and Linkletter, C. D. (2009). Bayesian calibration of the Thermosphere-Ionosphere Electrodynamics General Circulation Model (TIE-GCM). *Geoscientific Model Development*, 2:137-144.

Hansen, S. J., and Brown, J . W. (2004). *Investment Grade Energy Audit: making smart energy choices*. Fairmont Press, Inc, Lilburn, GA.

Healy, W. M., Lutz, J. D., and Lekov, A. B. (2003). Variability in energy factor test results for residential electric water heaters. *HVAC&R Research*, 9(4):435-449.

Hu, H. (2009). Risk-conscious design of off-grid solar energy houses. *Ph. D. Thesis*. Georgia Institute of Technology.

IPMVP (2010). International Performance Measurement and Verification Protocol: concepts and options for determining energy and water savings volume 1. Efficiency Valuation Organization.

ISO 13790 (2008). Thermal performance of buildings - calculation of energy use for space heating and cooling.

Jacobson, D. I., Dutt, G. S., and Socolow, R. H. (1986). Pressurization testing, infiltration reduction, and energy savings, in: *Trechsel, H. R., and Lagus, P. L. (Eds.), Measured Air Leakage of Buildings*. American Society for Testing and Materials.

Kemna, R., van Elburg, M., Li, W., and Holsteijn, R. (2007). Eco-design of CH boilers: technical analysis. *Task 4 Final Report*, VHK for European Commission.

Kennedy, M. C., and O'Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society, Series B*, 63(3):425-464.

Krause, P., Boyle, D. P., and Base, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, 5:89-97.

Larsen, T. S., and Heiselberg, P. (2008). Single-sided natural ventilation driven by wind pressure and temperature difference. *Energy and Buildings*, 40:1031-1040.

Lazzarin, R. M., and Schibuola, L. (1986). Performance analysis of heating plants equipped with condensing boilers. *Journal of Heat Recovery Systems*, 6:269-276.

Legates, D. R., and McCabe, Jr., G. J. (1999). Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, 35:233-241.

Lee, S. H., Zhao, F., and Augenbroe, G. (2011). The use of normative energy calculation beyond building performance rating systems. *Proceedings of the 12th International Building Performance Association Performance Simulation Association Conference*, Sydney Australia.

Liu, Y., Yang, P., Hu, C., and Guo, H. (2008). Water quality modeling for load reduction under uncertainty: A Bayesian approach. *Water Research*, 42:3305-3314.

Macdonald, I. A. (2002). Quantifying the effects of uncertainty in building simulation. *Ph. D. Thesis*, University of Strathclyde.

Mathew, P., Kromer, J. S., Sezgen, O., and Meyers, S. (2005). Actuarial pricing of energy efficiency projects: lessons foul and fair. *Energy Policy*, 33:1319-1328.

Mills, E. (2009). Building commissioning: a golden opportunity for reducing energy costs and greenhouse gas emissions. *LBNL Report*. Lawrence Berkeley National Laboratory.

Mills, E., Friedman, H., Powell, T., Claridge, D., Haasl, T., and Piette, M. A. (2004). The cost-effectiveness of commercial-buildings commissioning: a meta-analysis of energy and non-energy impacts in existing buildings and new construction in the United States. *LBNL Report-56637*. Lawrence Berkeley National Laboratory.

Mills, E., Kromer, S., Weiss, G., and Mathew, P. A. (2006). From volatility to value: analysing and managing financial and performance risk in energy savings projects. *Energy Policy*, 34:188-199.

Mills, E. (2003). Risk transfer via energy-savings insurance. *Energy Policy*, 31:273-281.

Moon, H. J., and Augenbroe, G. (2007). Application of probabilistic simulation and Bayesian decision theory in the selection of mold remediation actions. *Proceedings of the 10th International Building Performance Association Performance Simulation Association Conference*, Beijing China.

Moon, H. J. (2005). Assessing mold risks in buildings under uncertainty. *Ph. D. Thesis*, Georgia Institute of Technology.

Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2):161-174.

Pan, Y., Huang, Z., and Wu, G. (2007). Calibrated building energy simulation and its application in a high-rise commercial building in Shanghai. *Energy and Buildings*, 39:651-657.

Pedrini, A., Westphal, F. S., and Lamberts, R. (2002). A methodology for building energy modeling and calibration in warm climates. *Building and Environment*, 37:903-912.

Perera, M. D. A. E. S., Henderson, J., and Webb, B. C. (1997). Simple air leakage predictor for office buildings: assessing envelope airtightness during design or before refurbishment. *Proceedings of CIBSE National Conference*, 21-26.

Persily, A. K. (1999). Myths about building envelops. *ASHRAE Journal*. 39-45.

Persily, A. K. (1998). Airtighness of commercial and institutional buildings: blowing holes in the myth of tight buildings. *Proceedings of Thermal Performance of the Exterior Envelopes of Buildings*, 829-837.

Qian, S. S., Reckhow, K. H., Zhai, J., and McMahon, G. (2005). Nonlinear regression modeling of nutrient loads in streams: a Bayesian approach. Water Resources Research, 41(7).

Radgen, P., Oberschmidt, J., and Cory, W. T. W. (2006). EuP Lot 11: fans for ventilation in non residential buildings. *Report*. Institute Systems and Innovation Research.

Rasmussen, C. E., and Williams, C. (2006). *Gaussian Processes for Machine Learning*. The MIT Press, USA.

Reddy, T. A., Maor, I., Jian, S., and Panjapornporn, C. (2006). Procedures for reconciling computer-calculated results with measured energy data. *ASHRAE Research Project 1051-RP*.

Reddy, T. A. (2005). Literature review on calibration of building energy simulation programs: uses, problems, procedures, uncertainty, and tools. *ASHRAE Transactions*, 112:226-240.

Rijal, H., Tuohy, P., Humhreys, M., Nicol, J., Samuel, A., and Clarke, J. (2007). Using results from field surveys to predict the effect of open windows on thermal comfort and energy use in buildings. *Energy and Buildings*, 39:823-836.

Roth, K. W., Westphalen, D., Feng, M. Y., Llana, P., and Quartararo, L. (2005). Energy impact of commercial building controls and performance diagnostics: market characterization, energy impact of building faults and energy savings potential. *Report prepared for U.S. Department of Energy*. TIAX LCC, Cambridge, MA.

Ryan, J. D., and Nicholls, A. (2004). Commercial building R&D program multi-year planning: opportunities and challenges. *Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings*. American Council for an Energy-Efficiency Economy, Washington, DC.

Satchwell, A., Goldman, C., Larsen, P., Gilligan, D., and Singer, T. (2010). A survey of the U. S. ESCO industry: market growth and development from 2008 to 2011. *Technical Report*, Lawrence Berkeley National Laboratory.

Sargent, R. G. (2005). Verification and validation of simulation models. *Proceedings of the 37th Conference on Winter Simulation*, 130-143.

SECO (2007). LoanSTAR Revolving Loan Program. Available from http://www.seco.cpa.state.tx.us/ls/

SIMLAB (2009). Software package for uncertainty and sensitivity analysis. Joint Research Centre of the European Commission.

Southern California Edison (2005). DEER Residential SEER-rated units performance maps - phase 2 report: performance maps and methodology development. California.

van Oijen, M., Rougier, J., and Smith, R. (2005). Bayesian calibration of process-based forest models: bridging the gap between models and data. *Tree Physiology*, 25:915-927.

White House (2011). President Obama's plan to win the future by making American businesses more energy efficient through the "Better Buildings Initiative". Available from http://www.whitehouse.gov/the-press-office/2011/02/03/president-obama-s-plan-win-future-making-american-businesses-more-energy

Wyss, G., and Jorgensen, K. (1998). *A user's guide to LHS: Sandia's Latin Hypercube Sampling Software*. Sandia National Laboratories, Albuquerque, NM.

Yoon, J., Lee, E. J., and Claridge, D. E. (2003). Calibration procedure for energy performance simulation of commercial building. Journal of Solar Energy Engineering, 125:251-257.

Zhu, Y. (2006). Applying computer-based simulation to energy auditing: A case study. *Energy and Buildings*, 38:421-428.