Proceedings of BS2013:
13th Conference of International Building Performance Simulation Association, Chambéry, France, August 26-28

EVALUATION OF CALIBRATION EFFICACY UNDER DIFFERENT LEVELS OF UNCERTAINTY

Yeonsook Heo[1, 2], Diane Graziano[2], Leah Guzowski[2], and Ralph T. Muehleisen[2]
[1]Department of Architecture, University of Cambridge, Cambridge, UK
[1]Decision and Information Sciences Division, Argonne National Laboratory, Lemont, USA

ABSTRACT

This paper examines how calibration performs under different levels of uncertainty in model input data. It specifically assesses the efficacy of Bayesian calibration to enhance the reliability of EnergyPlus models. A Bayesian approach can quantify uncertainty in uncertain parameters while updating their values given measurement data. We assess the efficacy of Bayesian calibration under a controlled virtual-reality setup, which enables researchers to rigorously validate the accuracy of calibration results in terms of both calibration parameter values and calibrated model predictions. Case studies demonstrate the performance of Bayesian calibration of base models developed from audit data with differing levels of detail in building design, usage and operation.

INTRODUCTION

The objective of calibrating building energy models is to calculate feasible values for uncertain model parameters that are typically unattainable from a pool of available data. Typically, calibration of energy simulation models has been applied to reliably evaluate energy-savings potentials from energy efficiency measures (EEMs) (Pan, 2007; Zhu, 2006; Pedrini et al., 2002). Also, current standards, including the international performance measurement and verification protocol (IPMVP, 2010) and ASHRAE guideline 14 (ASHRAE, 2002), endorse the whole-building calibrated simulation approach for measuring and verifying energy savings achieved from EEMs implemented for existing buildings.

For energy retrofit projects, analysts use audit data to develop and calibrate building energy models for predicting energy savings from EEMs. Indeed, attaining correct values for model parameters depends on the level of data available for constructing an energy model, which is tightly related to the audit level. The ASHRAE research project 1051-RP summarizes six levels of calibration, depending on the building description and performance data, which can cover the majority of calibration cases for retrofit projects (Reddy et al., 2006). The audit level directly determines the amount of data available for modeling. As a result, the audit level impacts the accuracy of the base model and consequently the reliability of the calibrated model in capturing actual building behavior with high confidence. Hence, the efficacy of calibration should be scrutinized under different levels of uncertainty residing in the base models.

For energy retrofits of individual buildings, ASHRAE Guideline 14 provides a standard analysis procedure for the calibrated simulation approach. From data collected from a detailed audit, one constructs a building simulation model. Then, one estimates uncertain parameter values by comparing model outcomes with measured data until discrepancies between predicted and monitored energy use meet an acceptable tolerance. ASHRAE Guideline 14 stipulates that the coefficient of variation of the root mean square error (CVRMSE) should be within 15% and 30% with use of monthly and hourly energy-use data, respectively, in order for the model to be validated.

In practice, standard calibration techniques have two major drawbacks. First, the calibration process is often subject to experts' judgment, especially for selection of calibration parameters and manual testing of parameter values. High dependency on expertise has been recognized as a major problem that undermines the quality of calibration results (Reddy et al., 2006). Second, the calibration techniques follow a deterministic approach, and accordingly compute a single set of parameter values that minimizes the discrepancy while ignoring uncertainty in the model inputs and the model itself.. In order to quantify uncertainty in calibrated models, Heo et al. (2012a, 2012b) applied Bayesian calibration to building energy models, which resulted in probabilistic distributions of calibration parameters and demonstrated the importance of uncertainty information for retrofit decision-making, especially in the context of performance-based contracts.

Beyond retrofit projects for individual buildings, recent research attempts to apply calibration to derive unknown model parameter values for developing building stock models. A major challenge for building stock modeling is that detailed information about the building portfolio is not available, and accordingly large uncertainties reside in most model parameters. In order to overcome this challenge, Tian and Choudhary (2012) developed a representative

EnergyPlus model for school buildings in London, and applied Bayesian inference to obtain probability distributions of the four main parameters that can cover diversity across different school buildings, using published summaries of energy-use data. Following the same approach, Zhao (2012) derived building design and operational parameter values to replicate an example building stock with use of building energy models and the Commercial Buildings Energy Consumption Survey (CBECS) 2003 database.

This paper scrutinizes the efficacy of calibration under different levels of uncertainty in the model associated with the level of available data. It specifically assesses the efficacy of Bayesian calibration to enhance the reliability of building energy models. A Bayesian approach can quantify uncertainty in uncertain parameters while updating their values given measurement data. Case studies are conducted to evaluate the performance of Bayesian calibration with base models constructed from different levels of detail in available data.

BAYESIAN CALIBRATION

We apply Bayesian inference as a new approach to calibrate uncertain parameters in energy models while accounting for uncertainty in the calibration process. Bayesian calibration is an alternative to traditional, expert-intensive approaches that require "tweaking" of energy model input parameters to match measured data. Typically, traditional approaches deterministically search for a single solution that minimizes the discrepancy between predicted and measured energy use while ignoring many feasible solutions that may have higher likelihoods. Instead, the Bayesian approach derives the most likely distributions, referred to as posterior distributions, of uncertain parameters in the building energy model. The resulting calibrated model is able to compute probabilistic outcomes while accounting for uncertainty in the model inputs. The comparison between the deterministic and Bayesian calibration has been well summarized by Heo et al. (2012a).

The Bayesian paradigm treats a probability as a numerical estimate of the degree of belief in a hypothesis. Under this paradigm, our prior belief in true values of calibration parameters, θ, is quantified as prior density functions $p(\theta)$. The prior distributions are updated, given measured data on building performance, through the Bayes' theorem defined in Equation 1. $p(y|\theta)$ refers to a likelihood function that drives the updating process by comparing how closely model outcomes created with testing parameter values match the measured data, y. The likelihood function is derived from the mathematical formulation developed by Kennedy and O'Hagan (2001).

$$p(\theta|y) \propto p(\theta) \times p(y|\theta) \qquad (1)$$

Since the posterior distributions, $p(\theta|y)$, cannot be analytically derived for nonlinear energy models, they are numerically approximated from one joint multivariate distribution through the Metropolis-Hastings method (one of the Markov Chain Monte Carlo methods). The method iteratively explores the parameter space by sampling a proposed point based on the current point, and accepts the proposed point when it meets an acceptance criterion (Gelman et al., 2004). As a result, the method provides a set of accepted parameter values that approximate theoretical posterior distributions. Detailed information about the Bayesian calibration setup has been provided by Heo et al. (2012b).

Bayesian calibration is deployed in a formal process designed to minimize the role of expert judgment in the calibration process. In the proposed process, expert judgment for selecting calibration parameters and their parameter space is replaced with two pre-steps: (1) prior-uncertainty quantification and (2) parameter screening. First, uncertainties in model input parameters are quantified from evidential knowledge collected from a pool of sources (e.g., site surveys, technical papers, industry reports and standards). Then, a parameter screening method, specifically the Morris method, is applied to identify the most influential parameters with respect to their effects on energy use. The Morris method draws samples in the parameter space by changing one parameter value at a time and computes an elementary effect per parameter that explains the average change in the model output resulting from the change in the parameter value (Morris, 1991). The method efficiently evaluates the sensitivity of many uncertain parameters with a small number of samples, and can still explain the effects of individual parameters on the model output in a global sense. The most dominant parameters identified by the parameter screening are calibrated by the Bayesian calibration module with three types of inputs: (1) prior density functions of calibration parameters, (2) a set of model inputs and outputs exploring the parameter space, and (3) measured energy use (monthly utility bills, in our analysis).

EVALUATION FRAMEWORK

This paper evaluates the efficacy of Bayesian calibration in enhancing the reliability of a baseline energy model under different levels of energy audit. Table 1 summarizes three levels of data available for modeling depending on the audit level, modified from the ASHRAE research project 1051-RP (Reddy et al., 2006). Level 1 provides information about building geometry and thermal properties from as-built drawings, but provides no information about HVAC system characteristics, operational states, building use and operation strategies. In addition to as-built drawings, Level 2 obtains HVAC system inventory/specifications and building use/operation strategies from walk-through site visits. For instance, at this level, lighting and appliance power densities are estimated from equipment inventory. In addition to the information noted above, Level 3 provides measured data on system operational states and end-use energy data. These measured data enable accurate estimation of control state variables and lighting and appliance power densities with much reduced uncertainty arising only from measurement errors.

Table 1 Levels of data available for three audit levels

DATA	LEVEL 1	LEVEL 2	LEVEL 3
Utility bills (1 yr)	X	X	X
As-built drawings	X	X	X
Walk-through site visits		X	X
Detailed audit			X
Monitored end-use data			X

Applying the three audit levels, we evaluated the efficacy of Bayesian calibration of EnergyPlus models for three types of office buildings constructed before 1980, located in Chicago: (1) small office building, (2) medium office building, and (3) large office building. We used the U.S. Department of Energy (DOE) commercial reference buildings and associated EnergyPlus models developed by the National Renewable Energy Laboratory for this study (DOE, 2012). The efficacy of Bayesian calibration is assessed under a controlled virtual-reality setup. For each type of office building, we generated a "real" building by quantifying uncertainty in model parameters from Audit Level 3, randomly selecting model input values from the ranges of quantified uncertainties, and adding random measurement errors ranging between -2% and 2%. Then, the energy consumptions predicted by the "real" building model became the "utility bills" against which base models were calibrated. Calibrated models were evaluated against "real" buildings under two evaluation criteria: (a) the accuracy of calibration parameter values and (b) the accuracy of calibrated model predictions. We used CVRMSE, specified in Equation 2, as a statistical measure:

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

where P_i denotes a predicted variable value for period i, O_i an observed value for period i, and $\bar{O}$ the mean of all observed variable values. CVRMSE quantifies the discrepancy between testing values and targeting values in a normalized manner; where a value of 1.0 indicates that the discrepancy is equivalent to the average targeting value. The first criterion compares posterior distributions against true values from the "real" buildings, while the second criterion compares model predictions against utility bills.

CASE STUDIES

Calibration of accurate-level model (Level 3)

For Level 3, the base models are constructed from audit data that provide accurate estimates for most model parameters except infiltration and HVAC system efficiency. As a result, infiltration rate is the most dominant parameter, followed by heating system efficiency, infiltration rate reduction (while the mechanical system is on), and fan system efficiency. Although heating setpoint temperature during occupied hours is one of the top four parameters for the large office building, the three office buildings yield similar rankings of uncertain parameters; infiltration rate is by far the most dominant parameter, while the system-related parameters have much smaller effects on energy use. For all the cases at Level 3, since the top four parameters have a much greater effect on energy use than the other parameters, we naturally selected them for calibration. We also selected the top four parameters for calibration at the other levels to equivalently compare the effect of Bayesian calibration across different model levels.

CVRMSE values in Table 2 compare the posterior distributions of the top four parameters against true values from the "real" buildings. All the parameters selected for calibration except infiltration rate already have quite low CVRMSE values, ranging between 0.01 and 0.20 before calibration. For these parameters, calibration does not noticeably improve the accuracy of parameter values compared to the prior distributions. However, for infiltration rate, calibration greatly reduces the discrepancy between the parameter values and the true value. As shown in Figure 1, Bayesian calibration greatly narrows the range of feasible values while updating the posterior distribution to move toward the true value regardless of where it is located.

Table 2 Evaluation of calibration parameter values against true parameter values at Level 3

	CVRMSE IN PARAMS		DIFF.
	Uncalibrated	Calibrated	
Small Building			
Infiltration	1.08	0.34	0.74
Heating sys. eff.	0.19	0.19	0.01
Infiltration reduc.	0.20	0.21	-0.00
Fan sys. eff.	0.16	0.15	0.00
Medium Building			
Infiltration	6.35	0.80	5.56
Fan sys. eff.	0.09	0.06	0.04
Heating sys. eff.	0.08	0.07	0.01
Infiltration reduc.	0.17	0.17	0.00
Large Building			
Infiltration	0.65	0.19	0.46
Heating sys. eff.	0.08	0.06	0.01
Heating setpoint T.	0.01	0.01	0.00
Infiltration reduc.	0.20	0.21	-0.01

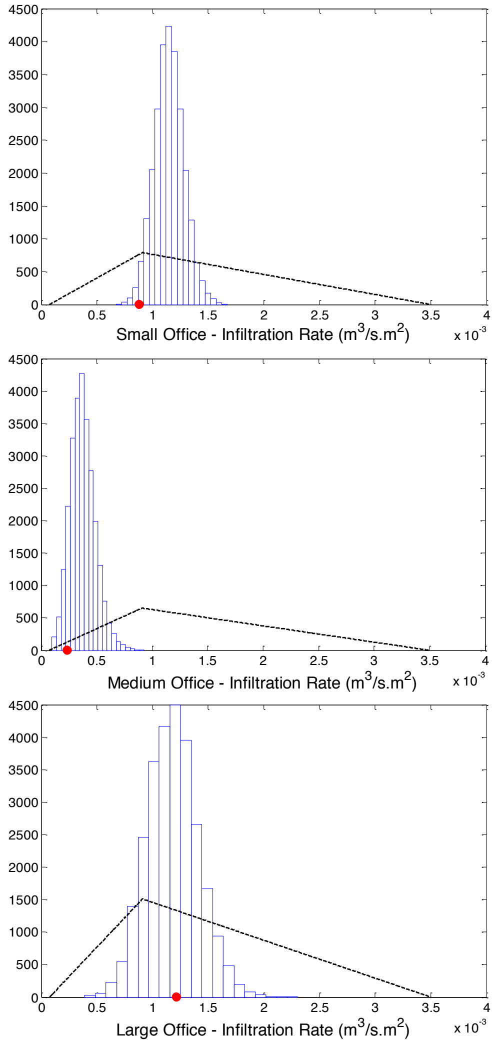


Figure 1 Posterior distribution of infiltration rate for the three cases at Level 3: priors (dashed line), posteriors (blue bars), and true value (red point)

Table 3 shows CVRMSE values that quantify the accuracy of energy consumptions predicted by the calibrated model and the uncalibrated model at Level 3. Overall, calibration enhances the accuracy of model predictions for the three office buildings. Calibration significantly reduces the discrepancy between model predictions and utility bills for gas consumptions, and has a lesser effect on reducing the already small discrepancy in electricity consumptions. This trend is expected because the calibration significantly updates infiltration rate, which heavily influences gas use for heating and has a smaller effect on electricity use for cooling. This trend may be valid only for studies in cold climate zones, including the present studies.

Table 3 Evaluation of model predictions (Level 3) against utility bills

	CVRMSE (ELECTRICITY)		**CVRMSE (GAS)**	
	Uncal.	Cal.	Uncal.	Cal.
Small building	0.19	0.04	0.76	0.13
Medium building	0.07	0.03	2.56	0.30
Large building	0.02	0.02	0.52	0.13

Calibration of intermediate-level model (Level 2)

In comparison to Level 3, the base models for Level 2 are constructed without measured data about system operation states and end-use energy use (i.e., lighting, appliances, and DHW). As shown in Table 4, model parameters related to that missing information exhibit a much higher magnitude of uncertainty compared to Level 3, in which their uncertainty arises only from measurement errors. As a result, the most dominant parameters identified by the Morris method include outside air flow, appliance power density, and fan pressure rise, in addition to infiltration rate and heating system efficiency, identified as the dominant parameters at Level 3.

Table 4 Evaluation of calibration parameter values against true parameter values at Level 2

	CVRMSE IN PARAMS		**DIFF.**
	Uncalibrated	Calibrated	
Small Building			
Infiltration	1.05	0.32	0.73
Outside air flow	4.28	1.29	3.00
Heating sys. eff.	0.19	0.21	-0.01
Fan pressure rise	0.43	0.27	0.16
Medium Building			
Outside air flow	4.22	0.38	3.84
Appliance power	0.73	0.24	0.49
Infiltration	6.35	1.93	4.42
Heating sys. eff.	0.08	0.07	0.01
Large Building			
Outside air flow	4.20	0.43	3.77
Appliance power	0.71	0.10	0.61
Infiltration	0.65	0.37	0.28
Heating sys. eff.	0.08	0.06	0.01

In the uncalibrated models at this level, the CVRMSE values of parameter values for infiltration rate, outside air flow, and appliance power density are quite large, whereas CVRMSE values for heating system efficiency are as small as those at Level 3. For those parameters with large discrepancies, Bayesian calibration significantly reduces CVRMSE values by moving the posterior distributions toward the true values while greatly reducing uncertainty in the distributions (i.e., distribution width), as shown in Figure 2. However, for the parameters with small discrepancies, the posterior distributions are little changed from the prior distributions, an effect which is also observed at Level 3. This trend implies that calibrating these four dominant parameters may be sufficient to update energy models at this level, but the efficacy of Bayesian calibration with a larger set of parameters is also investigated to confirm that implication.

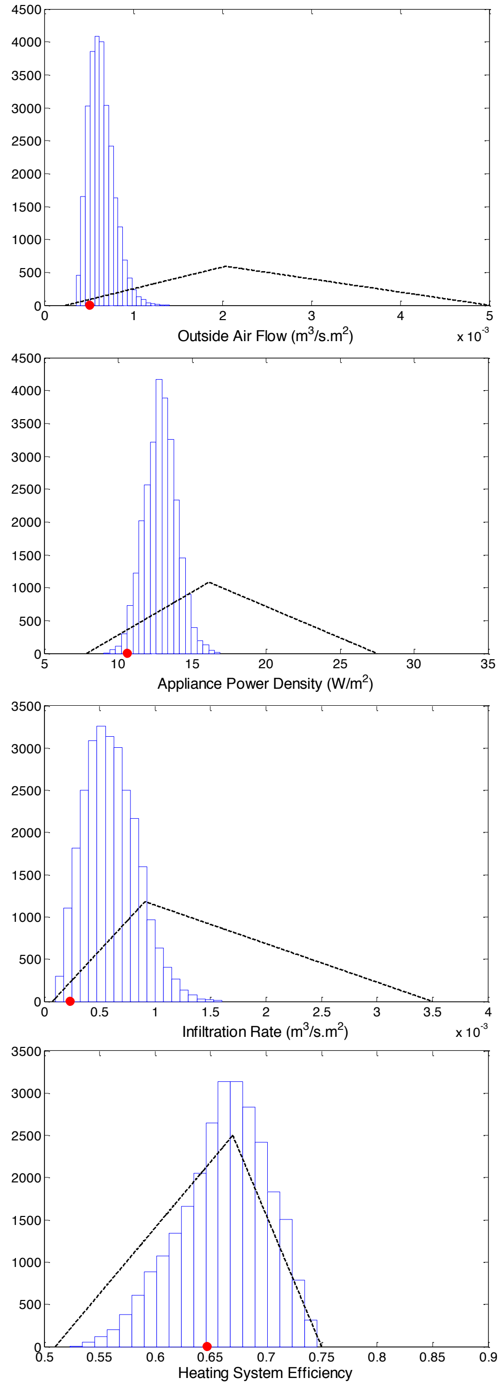


Figure 2 Calibration results for the medium office building at Level 2: priors (dashed line), posteriors (blue bars), and true value (red point)

Table 5 shows CVRMSE values that compare energy use predicted by the calibrated model and the uncalibrated model against utility bills for the three office building models at Level 2. Overall, calibration significantly improves the accuracy of model predictions for both electricity consumptions and gas consumptions. The calibrated models result in low CVRMSE values, less than 0.25 for all the cases except for gas consumptions for the medium-office building. For this particular case, the large discrepancy between predicted and actual gas use arises mainly from the large magnitude of uncertainty still remaining in infiltration rate in comparison to the other office building cases. However, for the same case with the building energy model at Level 3 (shown in Figure 1), the uncertainty range of infiltration rate is slightly less than half of that at Level 2 (shown in Figure 2). This comparison suggests that calibration may have limitations with respect to updating calibration parameter values to accurately correspond to true values when the uncertainty of model inputs is large. In particular, this uncertainty widens the parameter space to be explored during calibration, thereby magnifying interactive effects of calibration parameters on model outcome and confounding the effects of model parameters not included in the calibration. The effect of parameter interactive effects on calibration results is a topic for future study. Nonetheless, calibration is still shown to enhance the reliability of model predictions by an order of magnitude, and the resulting calibrated models at Level 2 are competitive with the uncalibrated models at Level 3.

Table 5 Evaluation of model predictions (Level 2) against utility bills

	CVRMSE (ELECTRICITY)		**CVRMSE (GAS)**	
	Uncal.	Cal.	Uncal.	Cal.
Small building	0.18	0.12	1.79	0.25
Medium building	0.29	0.08	10.14	1.39
Large building	0.61	0.16	1.42	0.18

Calibration of crude-level model (Level 1)

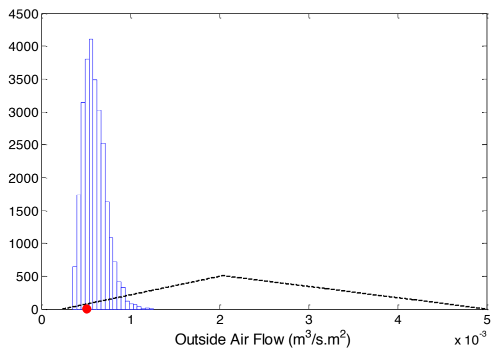
For Level 1, the base models are constructed only from as-built drawings, which provide information only about building geometry and construction specifications. Consequently, the three office buildings at this level yield similar rankings; the most dominant parameters include outside air flow, appliance power density, lighting power density, infiltration, and heating system efficiency. Appliance and lighting power densities become more dominant at this level because their uncertainty range covers internal power consumptions across various office buildings from field surveys (Knight and Dunn, 2003), in comparison to the Level 2 situation, in which they are estimated from the equipment inventory specific to the building being considered. For Level 1, these influential parameters have a major effect on energy use predictions, similar to that of the most dominant parameter at other levels.

The CVRMSE values in Table 6 demonstrate that the posterior distributions coincide with the true values much better than the prior distributions for all the parameters except heating system efficiency. As shown in Figure 3, the posterior distributions for the medium office building substantially reduce uncertainty, while their expected values more closely match the true values. However, for heating system efficiency, little change is observed between the prior and the posterior distributions. In addition, the calibration of crude-level models results in wider ranges of feasible values in the posterior distributions than those at Level 2 (shown in Figure 2). The larger uncertainty in the calibration results can be attributed to the large magnitude of uncertainty in model parameters, given the limited measurement data. Nevertheless, Bayesian calibration still leads to reasonable results that improve the accuracy of the baseline model.

Table 6 Evaluation of calibration parameter values against true parameter values at Level 1

	CVRMSE IN PARAMS		**DIFF.**
	Uncalibrated	Calibrated	
Small Building			
Infiltration	1.08	0.39	0.69
Outside air flow	4.18	1.18	3.00
Heating sys. eff.	0.29	0.31	-0.02
Appliance power	0.96	0.86	0.10
Medium Building			
Outside air flow	4.22	0.30	3.92
Appliance power	0.96	0.49	0.47
Lighting power	0.34	0.27	0.07
Heating sys. eff.	0.13	0.16	-0.03
Large Building			
Appliance power	0.95	0.17	0.78
Outside air flow	4.18	0.48	3.70
Lighting power	0.84	0.18	0.65
Infiltration rate	0.65	0.33	0.32

Table 7 shows CVRMSE values that compare model predictions against utility bills for the three office buildings. Overall, calibration greatly reduces CVRMSE values in model predictions as the result of correcting calibration parameter values. However, for the medium office building case, the calibrated model still results in a high CVRMSE value of 8.63 for gas energy-use prediction, which suggests that calibrating four parameters at this level may not be sufficient to obtain reliable predictions. As expected, the number of influential parameters increases as the model level goes down, with less data collected for modeling. Hence, the next section explores the effect of a larger set of calibration parameters on enhancing the model reliability.



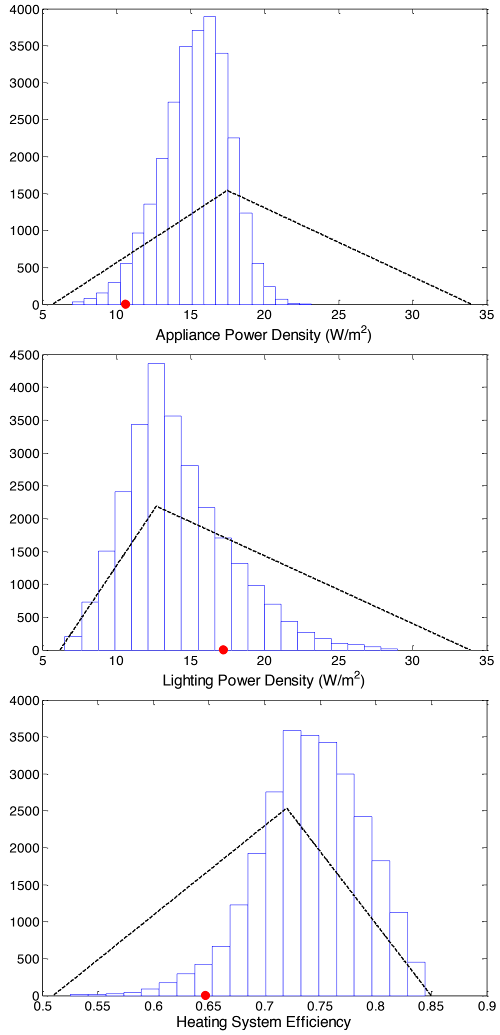


Figure 3 Calibration results for the medium office building at Level 1: priors (dashed line), posteriors (blue bars), and true value (red point)

Table 7 Evaluation of model predictions (Level 1) against utility bills

	CVRMSE (ELECTRICITY)		**CVRMSE (GAS)**	
	Uncal.	Cal.	Uncal.	Cal.
Small building	1.17	0.75	5.64	0.86
Medium building	1.56	0.57	31.47	8.63
Large building	3.65	0.74	4.94	1.05

Calibration of a larger set of parameters

This section examines whether calibrating a larger number of uncertain parameters at lower model levels can help enhance the reliability of the calibrated models. Table 8 summarizes CVRMSE values of parameter values from calibrating six and eight parameters at Levels 1 and 2 for the medium office building case. At Level 2, calibrating the larger number of parameters does not improve the accuracy of posterior distributions noticeably, either for the four parameters in the initial strategy (as compared to Table 8) or for the other parameters introduced in the new calibration exercises. Consequently, calibrating a larger set of parameters does not substantially improve the accuracy of model predictions, as observed in Table 9.

On the contrary, at Level 1, calibrating six parameters significantly reduces the CVRMSE value for the infiltration rate (ranked as the fifth), and as a result significantly reduces CVRMSE values in model predictions, especially for the gas energy prediction. However, calibrating two more parameters does not further improve the accuracy of parameter values and model predictions. These results suggest that Model Level 1 needs to calibrate a larger set of parameters, including outside air flow, appliance power density, lighting power density, heating system efficiency, and infiltration rate, to ensure the reliability of the calibrated model. Nevertheless, the CVRMSE value for gas predictions is still high because one-year monthly utility bills are not able to reduce the high magnitude of uncertainty in uncertain parameters with high interactions. In order to further enhance the model reliability at Level 1, further research is needed on developing advanced Bayesian calibration algorithms that utilize an extensive set of measurement data at different levels, including hourly data and submetered data.

Table 8 Evaluation of calibration parameter values against true parameter values at Levels 2 and 1

	CVRMSE IN PARAMS			
	Uncal.	4Params	6Params	8Params
Model Level 2				
Outside air flow	4.22	0.38	0.32	0.32
Appliance power	6.35	0.24	1.83	1.94
Infiltration	0.73	1.93	0.26	0.26
Heating sys. eff.	0.08	0.07	0.07	0.07
Fan pressure rise	0.44	-	0.46	0.46
Fan sys. eff.	0.09	-	0.09	0.09
Heating setpoint T	0.01	-	-	0.01
Lighting power	0.07	-	-	0.06
Model Level 1				
Outside air flow	4.22	0.30	0.44	0.45
Appliance power	0.96	0.49	0.47	0.42
Lighting power	0.34	0.27	0.27	0.27
Heating sys. eff.	0.13	0.16	0.15	0.15
Infiltration	6.41	-	2.10	2.04
Fan pressure rise	0.38	-	0.34	0.32
Heating setpoint T	0.03	-	-	0.03
Fan sys. eff.	0.31	-	-	0.30

Table 9 Evaluation of model predictions (Levels 2 and 1) against utility bills

	CVRMSE (ELECTRICITY)	**CVRMSE (GAS)**
Model Level 2		
Four params	0.08	1.39
Six params	0.07	1.18
Eight params	0.07	1.25
Model Level 1		
Four params	0.57	8.63
Six params	0.35	1.56
Eight params	0.28	1.56

CONCLUSION

This paper evaluated the efficacy of Bayesian calibration using EnergyPlus building energy models, given uncertainties consistent with different audit levels. Case studies with three types of office buildings demonstrated that for all audit levels, Bayesian calibration yields posterior distributions that correspond well to true values while significantly reducing uncertainty in parameter values. Consequently, calibrated models show enhanced reliability in their predictions, more closely matching utility bills with much-reduced uncertainty compared to uncalibrated models.

This research project is ongoing to extend the Bayesian calibration methods to OpenStudio. In order to enhance the practicality of Bayesian calibration, we will further investigate the following:

1. Alternative algorithms and techniques for speeding up Bayesian calibration (e.g.; parallel processing algorithms for posterior simulations);
2. In-depth guidance for applying Bayesian methods to calibrate EnergyPlus models for different levels of available data, including identification of uncertain parameters, selection of calibration parameters, and analysis of results;
3. A database for uncertainty in building energy models, with the objective of significantly reducing the upfront effort required for uncertainty analysis, collaborating with Georgia Tech researchers who have developed a workbench with a database of uncertainty arising from both parameter uncertainty and model inadequacy (Sun et al., 2011).

ACKNOWLEDGEMENT

This work was supported by the U.S. Department of Energy under Contract No. DE-AC02-06CH11357.

REFERENCES

ASHRAE 2002. ASHRAE guideline 14 – measurement of energy and demand savings, American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Atlanta, GA.

DOE 2012. Commercial reference buildings, http://www1.eere.energy.gov/buildings/commercial/ref_buildings.html.

Gelman, A., Carlin, J.B., Stem, H.S., Rubin, D.B. 2004. Bayesian Data Analysis, Chapman & Hall/CRC, Boca Raton, FL.

Heo, Y., Augenbroe, G., Choudhary, R. 2012a. Quantitative risk management for energy retrofit project, Journal of Building Performance Simulation, in press.

Heo, Y., Choudhary, R., Augenbroe, G. 2012b. Calibration of building energy models for retrofit analysis under uncertainty, Energy and Buildings, 47:550–560.

IPMVP 2010. International performance measurement and verification protocol: concepts and options for determining energy and water savings volume 1, Efficiency Valuation Organization. Washington, D.C.

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

Knight, I.P., Dunn, G.N. 2003. Evaluation of heat gains in UK office environments, in Worldwide CIBSE/ASHRAE Gathering of the Building Services Industry, ISBN 1-903287-43-x, Edinburgh, Scotland.

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

Pan, Y., Huang, Z., 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., Lamberts, R. 2002. A methodology for building energy modeling and calibration in warm climates, Building and Environment, 38:903–921.

Reddy, T.A., Maor, I., Jian, S. 2006. Procedures for reconciling computer-calculated results with measured energy data, American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Atlanta, GA.

Sun, Y., Heo, Y., Xie, H., Tan, M., Wu, J., Augenbroe, G. 2011. Uncertainty quantification of microclimate variables in building energy simulation. Proceedings of the 12th IBPSA Conference, Sydney Australia.

Tian, W., Choudhary, R. 2012. A probabilistic energy model for non-domestic building sectors applied to analysis of school buildings in greater London, Energy and Buildings, 54:1–11.

Zhao, F. 2012. Agent-based modeling of commercial building stocks for energy policy and demand response analysis, Ph.D. Thesis, Georgia Institute of Technology.

Zhu, Y. 2006. Applying computer-based simulation to energy auditing: a case study, Energy and Buildings, 38:421–428.