Option D Calibrated Simulation

# Situation

Option D involves using a calibrated building energy simulation (BES) model to predict facility or sub-facility energy use and savings and is usually carried out using either monthly or hourly energy data during a stable period of operation. Option D is applicable where baseline energy data is unavailable, such as in a new construction project or where metering will only be available after the energy conservation measure (ECM) installation ().

A significant challenge associated with Option D is the credibility of the BES model used for the predictions. These models need to be calibrated to predict confidently using simulations. Simple approach of uncertainty is unavoidable in BES because of the complexity and interactions between different building systems and sub-systems, often characterized by a large number of parameters that are related to the building geometry, internal loads, HVAC systems, operational schedules, etc. Therefore, searching for a single “representative” BES model is often unrealistic given that varying combinations of input parameters would match the calibration data equally well, a phenomenon known as equifinality. This has led to a calibration paradigm known as “calibration under uncertainty” that aims to address the topic of model calibration more holistically. In general, sources of uncertainties in BES have been classified into four main categories.

1. *Specification uncertainty*: Uncertainty arising from partial or inaccurate specification of the building systems modelled.
2. *Model uncertainty*: Uncertainty arising from assumptions, abstractions, and approximations of the actual physical system modelled.
3. *Operational uncertainty*: Uncertainty arising from a lack of feedback regarding actual use and operation of the building systems modelled.
4. *Scenario uncertainty*: Uncertainty arising from specifying the conditions imposed on the simulation such as weather and occupancy.

From a computer simulation standpoint, these sources of uncertainties often cumulate in the form of (1) inaccuracies in the values used as inputs to the model, (2) discrepancies between reality and the model selected to represent the actual physical system, and (3) errors in observed values. Therefore, general frameworks for uncertainty quantification in BES simulations usually classify uncertainties as *parameter uncertainty*, *model inadequacy or model discrepancy*, and *observation errors.*

# Approach

Uncertainty quantification in BES can generally be classified as either forward or inverse. Forward approaches quantify uncertainty by propagating uncertainties from model input(s) to the model output(s) and often involve applying sampling techniques such as Monte Carlo simulation. In contrast, inverse approaches start from a set of observations, quantifying various sources of uncertainties given the observed evidence. Among these inverse approaches, most notable in BES is the Bayesian calibration paradigm due to its ability to naturally incorporate uncertainty and combine prior information with measured data.

## Monte Carlo simulation

Using Monte Carlo simulation has become a popular method for uncertainty quantification in BES because of its easy implementation and the increased computing power that made it easier to run a large number of energy simulations.

A Monte Carlo approach to uncertainty quantification can be summarized by the following steps:

1. Identify uncertain input parameters and output of interest.
2. For each uncertain parameter, determine the probability distributions that best represent the uncertainties in the input parameters.
3. Generate random samples from the set of probability distributions.
4. Run energy simulations using the generated samples as inputs to the model.
5. Analyze the simulation results.

Choosing the sample size for the Monte Carlo simulation is problem specific. It depends on factors such as the number of uncertain parameters and their distributions, the simulation engine used, and the output of interest. A rule of thumb is to increase the sample size until the statistic of interest (e.g., standard deviation, 95% confidence interval, etc.) stabilizes.

The resulting output distribution is expected to converge towards a Normal distribution based on the Central Limit Theorem despite having different parameter distributions. In BES, direct sampling from the probability distribution is typically carried out, assuming independence between each uncertain parameter. When using Monte Carlo simulation, it is important to know that it is an approximate approach and that the results can be sensitive to the choice of probability distributions used to describe the uncertainties in the model parameters.

## Bayesian calibration

Bayesian calibration is an approach to estimating uncertainties in the energy model based on Bayes’ theorem. Specifically, the posterior estimates of the uncertain model parameters are derived from a combination of background knowledge expressed in the form of prior probability distributions and observed data expressed in the form of a likelihood function.

Most notable in BES is the approach proposed Kennedy and O’Hagan (2001) that explicitly models various sources of uncertainty. In BES, typical uncertain sources include parameter uncertainties, model discrepancy or inadequacy, and observation error.

Where is the observed data, is the BES output, is the discrepancy between the model and reality, and is the observation errors. and represents the observed inputs (e.g., inputs such as outdoor dry-bulb temperature and humidity levels that form the boundary conditions of the simulation) and calibration parameters (e.g., infiltration rate, lighting power density, etc.), respectively. Here is used to represent the true but unknown values of the calibration parameters to illustrate that the BES model is a biased representation of the actual physical building system even in the ideal case where .

Markov Chain Monte Carlo (MCMC) is used to sample from the posterior distribution since an analytical solution is not possible given the complexities of the BES model. Consequently, using metamodels as a substitute for the BES model to reduce simulation runtime is an essential part of Bayesian calibration. A metamodel is a simplified model of the energy simulation model representing the input and output relationships. Gaussian processes (Williams and Rasmussen, 2006) are a popular metamodel choice because of their flexibility, generality, and proven accuracy from past BES calibration studies.

# Examples

Both examples presented in this section uses free and open-source software. Energy simulation is carried out using EnergyPlus. R programming language is used for statistical computing and graphics. Specifically, we use the eplusr R package (<https://cran.r-project.org/web/packages/eplusr>) to interface with EnergyPlus using R and ease running parametric simulations and the programmatic manipulation of simulation results.

## Monte Carlo simulation

This example involves the new construction of a medium office building in San Diego, California. Total electricity energy savings are to be analyzed for energy efficiency measures that reduce the following parameters in the baseline building by a certain percentage:

* Lighting Power Density (LPD)
* Electric Equipment Power Density (EPD)

### Energy Model and Variable Characterization

The energy model for this example is created by using the following DOE Commerical Reference Building Model (https://www.energy.gov/eere/buildings/commercial-reference-buildings):

* Template: DOE Commercial Prototype Building Models
* Building Type: Medium Office ANSI/ASHRAE/IES Standard 90.1-2022
* Representative City: San Diego, CA
* Climate Zone: ASHRAE Thermal Climate Zone 3C (Warm Marine)
* Weather Location: San Diego/Brown Field Municipal Airport, California
* Weather Data: TMY3

To understand the effect that these parameters have on total building electricity and natural gas usage as well as understanding the sensitivity and uncertainty in the resulting energy savings for these measures, we first need to define valid ranges of the input uncertainty (a minimum and maximum) and a possible distribution type (triangular, uniform, normal, etc.) for each input parameter (LPD and EPD). These input uncertainties can come from many sources such as installation and implementation variations as well as other factors beyond our control. This means that even though we intend to reduce both LPD and EPD by exactly 10% in the building, from a practical standpoint, we will not see a reduction of exactly 10% but more of a distribution of reductions around that 10% target reduction.

To simulate this effect, we need to try and quantify the variation in the input that we will expect to have. For this example, we will consider a 5% variation in the 10% reduction that we wish to apply to the building. This gives us the minimum and maximum values for our distribution and bounds the input parameters at 5% and 15% of the baseline model values of LPD and EPD. We also need to choose a distribution shape to finish quantifying our input uncertainty. We could choose a flat or uniform change from 5% to 15%, a normal or bell curve distribution or one of several others that we think identifies the uncertainty we have about our ability to implement the energy efficiency measures. For this example, we will make the choice of using a triangle distribution for the uncertainty that peaks to its maximum at a 10% reduction and tails off at the bounds as listed in Table 1. This implies there is a higher probability of achieving the 10% reduction and a decreasing probability to achieve reduction values towards the bounds of 5% and 15%.

Table 1 Input parameter ranges for energy efficiency measures

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | Minimum | Mode | Maximum |
| LPD (% reduction from baseline) | 5% | 10% | 15% |
| EPD (% reduction from baseline) | 5% | 10% | 15% |

### Varying each input independently

We will sample the two variables, one at a time while holding the other constant (at their default values) resulting in an “OAT” (One-At-a-Time) sampling method. This method is also described as a local method since we are holding all the non-sampled variables to their default values. Thus, while we are getting some sense of the effect of the perturbed variable, this perturbation is only with respect to the same locally fixed values for the other variables. For illustration purposes, the number of samples in this example for each variable was chosen to be 50. This resulted in a total of 400 simulations to run (200 samples x 2 variables = 400 total runs).

A comparison of a graph

Description automatically generated with medium confidence

Figure 1 Histogram of 200 samples for lighting power density (left) and equipment power density (right).

Histograms of Total Electric Use and Savings variation due to the variation in the variables are depicted in Figure 2. The pink histogram characterizes the Total Electric Use after the change in LPD is applied to the baseline model while the yellow represents the change due to the EPD measure. The histograms are plotted as density functions so that they are normalized, with the area under each curve is one. Thus they can be treated as probability distributions. The uncertainty of each variable on total electric use can be determined by comparing the width of each colored distribution to the other. The larger the width of the resulting distribution means a larger uncertainty in that output to the input variable.

In our example, reduction in LPD results in total electricity end use with mean 296733 kWh and standard deviation 1036 kWh as highlighted by the pink histogram in Figure 2. In contrast we see that reduction in EPD results in greater savings in total electricity energy use (mean = 280902 kWh) but with significantly larger uncertainty (standard deviation = 4628 kWh). These results are in comparison to the base model, illustrated in Figure 2 with a dotted red line, which does not incorporate any reductions in LPD and EPD, and has a total electricity end use of 302425 kWh.

A graph of a graph of electricity

Description automatically generated

Figure 2 Variation in total electric end use.

### Varying both variables simultaneously

In this subsection, we attempt to compute the precision of the estimated Total Electric Use and Savings due to both LPD and EPD combined. To create the data necessary, we sample **all** the Input Variables at the **same** time. For illustrative purposes, we will use 400 as the sample size.

The result of sampling 400 values of LPD and EPD changes is 400 output values of total electricity energy end use. Figure 3 shows the distribution of total electricity energy end use (mean = 275204 kWh, standard deviation = 4620 kWh) when both LPD and EPD were varied simultaneously.

One can imagine repeating this analysis for a much larger set of building characteristics. In most buildings, there will be uncertainty in operating schedules, HVAC equipment efficiencies, baseline lighting density, plug load density, etc. Some of these parameters are not changed by the energy efficiency measures but have a significant impact on the savings estimates. The analysis will show the expected range of savings for the expected uncertainty in these inputs.

A graph of electricity

Description automatically generated

Figure 3 Variation in total electricity end use when both inputs are varied simultaneously.

## Bayesian calibration

This example is adapted from the paper by Chong and Menberg (2018). Figure 4 below shows the overall process flow of Bayesian calibration, which includes:

1. Generate samples using Latin hypercube samples.
2. Parse the generated samples to the energy model to quantify the impact of varying the calibration parameters on the calibration target or the output of interest.
3. Combine the measured and simulated data in a Gaussian process metamodel and explore the joint posterior distribution using Markov Chain Monte Carlo (MCMC).
4. Parse posterior samples generated using MCMC back to the energy model to check the predictive performance of the calibrated energy model.

A diagram of a scientific experiment

Description automatically generated with medium confidence

Figure 4 Overview process flow of Bayesian calibration.

In this example, Bayesian calibration at a monthly time resolution (over a one-year period from January 2013 to December 2013) is demonstrated using the following scenario.

1. Observed output or the simulation output that the model is calibrated against:
   * Whole facility electricity usage (kWh)
2. Observed inputs or inputs to the simulation model that are observable or measurable:
   * Outdoor air drybulb temperature (°C)
   * Outdoor air relative humidity (%)
3. Calibration parameters or the parameters that are to be estimated:
   * Zone infiltration flow rate per exterior surface area (m3/s-m2)
   * Electric equipment design power (W)
   * Lighting levels (W)
   * AHU total fan efficiency (-)

### Energy model and measured data

This example is based on an open model and data that was made publicly available on the National Renewable Energy Laboratory’s (NREL) Parametric Analysis Tool (PAT) GitHub Repository (https://github.com/NREL/OpenStudio-PAT/). The one-story building is the site entrance building to NREL and is located in Colorado, USA (Figure 5). Data used to calibrate the model include:

1. Twelve months of actual electricity energy usage (January 2013 to December 2013).
2. Actual Meteorological Year (AMY) hourly weather data for the year 2013.

A picture containing cube, LEGO

Description automatically generated with medium confidence

Figure 5. Visualization of EnergyPlus model used in example.

### Gaussian process metamodel

As mentioned in Section 2.2, using metamodels to emulate the energy model is an essential part of Bayesian calibration so that it is computationally tractable. The first step to constructing a metamodel or Gaussian process model is to generate an inputs-output simulation dataset for training the metamodel. This is achieved by varying the calibration parameter across a reasonable range and simulating the given input parameter values to obtain the corresponding output value. Latin hypercube sampling (Stein, 1987) instead of random sampling is often used to sample the search space more efficiently so that a smaller sample size is required to train the metamodel to similar levels of predictive performance. Figure 6 below shows the interactions between the generated samples to check for uniformity and spread of the samples across all four parameter, which is important for ensuring a robust metamodeling fitting process.

A graph of a function

Description automatically generated with medium confidence

Figure 6 Input parameter values generated using Latin hypercube sampling in R.

The samples were generated using a seed value of 1 to ensure reproducibility. Note that the input values are generated in the normalized range [0, 1] and can be considered the fraction of the expected maximum value or upper bound. The generated input parameter values were parsed into the EnergyPlus model to simulate the corresponding output value, i.e., the facility’s monthly electricity usage. Figure 6 compares the simulated output values with the actual observations over the same period. It is essential that the actual values fall within the range of simulated output values, as illustrated in Figure 7 since metamodels are defined over the same inputs-output range as the simulation model, and caution is needed to avoid extrapolating input parameter values beyond the range of values simulated.

A graph with orange and green dots

Description automatically generated

Figure 7 Visual comparison to ensure observed values are within the range of the simulation dataset.

### Posterior sampling using MCMC

After generating the simulation dataset, the next step is to combine the simulation dataset with the observed data using a Gaussian process model following the formulation by Kennedy and O’Hagan (Eq. 2). MCMC is then used to sample from the joint posterior distribution. Detailed formulation of the mean and covariance function used to define the Gaussian process model can be found in Higdon et al. (2004). In this example, we use RStan, the R interface to Stan, a probabilistic programming language for the MCMC sampling. The Stan code is from Chong and Menberg (2018) and is freely available on GitHub (https://github.com/adChong/bc-stan).

Before running the MCMC simulation, the prior distribution of the calibration parameters needs to be specified. Prior distributions are an important element in any Bayesian inference, especially when data is sparse. This is the case in this example since we are calibrating the model with only twelve data points (twelve months of observations given a one-year period). Before observing any data, the prior distribution represents the energy modeler’s knowledge of the calibration parameters. For instance, an energy modeler with access to audit reports that convey detailed information about the calibration parameters would use more informative priors to represent confidence in their prior knowledge or belief. In contrast, an energy modeler who relies on past case studies or research without access to specific building reports would more likely specify weakly informative priors to represent their limited knowledge about the calibration parameters. When the modelers know nothing about the calibration parameters, a noninformative or flat prior is often specified.

The use of weakly informative priors is encouraged, while energy modelers should, in general, move away from the use of noninformative priors. Weakly informative priors have been shown to be a good compromise between highly informative and noninformative priors. Consequently, they are preferred because they influence the resulting posterior distribution but not so strongly as to rule out values that might make sense after incorporating the observed data. Noninformative priors, however, represent total ignorance. In this example, noninformative priors were assigned since we do not have any prior information about the calibration parameters.

### Model evaluation

Figure 8 shows histograms of the posterior distributions for the two calibration parameters. From Figure 4, it can be observed that the calibration with measured data significantly reduced uncertainties in electric equipment design power . In contrast, the measured data did not significantly affect the infiltration rate and the posterior distribution is mainly driven by the prior .

A graph of value and values

Description automatically generated

Figure 8 Posterior distribution of the four calibration parameters.

Following the utilization of MCMC to derive the posterior distributions of the calibration parameters, we proceed to sample these distributions. In this example, we sampled (with replacement) fifty sets of calibration parameters. These samples represent a range of plausible parameter sets informed by the Bayesian calibration process. Using these sampled parameter sets, we ran a new series of energy simulations. This allows us to generate a calibrated energy model capable of supporting various “what if” scenarios.

Figure 9 visually compares the agreement between the predictions from the calibrated energy model (green box plots) and the measured data (orange points). This visual comparison reveals a reasonable trend agreement, particularly noticeable in the slightly higher electricity usage in January which initially decreases, then peaks in July and August, before decreasing again. Both the measured data and calibrated simulation predictions follow this trend. The calibrated model yields Coefficient of Variation of the Root Mean Squared Error (CV(RMSE)) and Normalized Mean Bias Error (NMBE) of 11% and -6%, respectively. To achieve lower CV(RMSE) and NMBE values, the discrepancy term (referenced in Equation 2) could be omitted during the Bayesian calibration process, contingent upon the specific calibration goals and desired accuracy level.

A graph with green and blue squares

Description automatically generated

Figure 9 Comparison of posterior predictions (green box-plots) with measured data (orange points).

### Simulating energy conservation measures (ECMs)

Having quantified the uncertainties in the calibration parameters through Bayesian calibration, we can now assess the impact of various Energy Conservation Measures (ECMs) on potential energy savings. In this study, we demonstrate the effects of two specific ECMs: reducing equipment load by 30% and upgrading the Air Handling Unit (AHU) fan to achieve a total efficiency of 0.7. Prior to implementing the ECMs, the energy usage averaged 17,308 kWh, with a standard deviation of 4,499 kWh (Figure 10 and Table 2). After the measures were applied, the mean energy consumption dropped to 12,755 kWh, accompanied by a decreased standard deviation of 3,032 kWh, indicating variability in the savings across different simulation scenarios.

A graph of electricity and electricity

Description automatically generated with medium confidence

Figure 10 Distribution of total electricity end use before and after implementing the energy conservation measures (ECMs).

|  |  |  |
| --- | --- | --- |
| Statistic | Base (kWh) | With ECMs (kWh) |
| Mean | 17308 | 12755 |
| Median | 17335 | 12683 |
| SD | 4888 | 3032 |
| IQR | 6474 | 4350 |
| Min | 7803 | 5878 |
| Max | 36397 | 20811 |
| 25th Percentile | 13594 | 10500 |
| 75th Percentile | 20069 | 14850 |

Figure 11 Summary statistics of total electricity end use before and after implementing the energy conservation measures (ECMs).

# Data and Code Availability

The data and code for the two examples can be found at XXX, hosted on GitHub.

# References

Chong, A., & Menberg, K. (2018). Guidelines for the Bayesian calibration of building energy models. *Energy and Buildings*, *174*, 527-547.

Higdon, D., Kennedy, M., Cavendish, J. C., Cafeo, J. A., & Ryne, R. D. (2004). Combining field data and computer simulations for calibration and prediction. *SIAM Journal on Scientific Computing*, *26*(2), 448-466.

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

Stein, M. (1987). Large sample properties of simulations using Latin hypercube sampling. *Technometrics*, *29*(2), 143-151.

Williams, C. K., & Rasmussen, C. E. (2006). *Gaussian processes for machine learning* (Vol. 2, No. 3, p. 4). Cambridge, MA: MIT press.