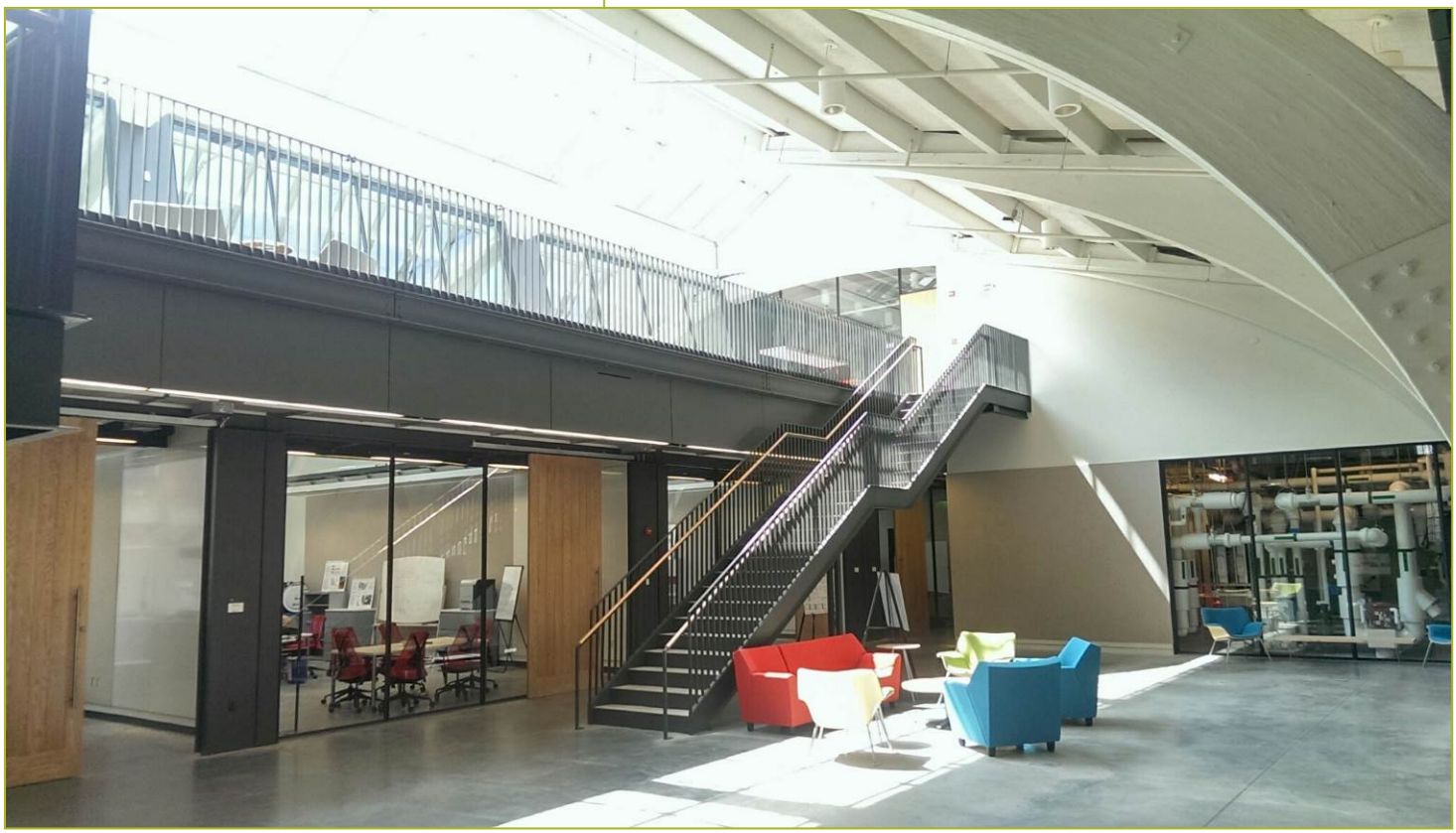
CONSORTIUM for BUILDING ENERGY INNOVATION

REPORT

Title: Uncertainty Quantification in Energy Efficient Building Performance Simulations

Report Date: January 2012

Report Author(s): Slaven Peles, Sunil Ahuja, Satish Narayanan



CBEI was referred to as the Energy Efficiency Buildings HUB at the time this report was developed.

Funded by U.S. DOE

Report Abstract

Whole building energy models do not always provide satisfactory predictions to facilitate decision making during design, due to large number of uncertainties in model input parameters. CBEI presents a computationally efficient process for uncertainty quantification, sensitivity analysis and automated calibration of building models. This is demonstrated using an energy simulation model of a medium sized office building.

Contact Information for Lead Researcher

Name: Satish Narayanan

Institution: United Technologies Corporation

Acknowledgement

This material is based upon work supported by the Consortium for Building Energy Innovation (CBEI) sponsored by the U.S. Department of Energy under Award Number DE-EE0004261.

Uncertainty quantification in energy efficient building performance simulations

Conference Paper · January 2012

CITATIONS
3

READS
68

3 authors:

Slaven Peles
Lawrence Livermore National Laboratory
20 PUBLICATIONS **175** CITATIONS
SEE PROFILE

Sunil Ahuja
United Technologies Research Center
17 PUBLICATIONS **173** CITATIONS
SEE PROFILE

Satish Narayanan
United Technologies Corporation
69 PUBLICATIONS **712** CITATIONS
SEE PROFILE

Uncertainty quantification in energy efficient building performance simulations

Slaven Peleš*, Sunil Ahuja, Satish Narayanan

United Technologies Research Center
411 Silver Lane, East Hartford, CT 06108
Phone: (860) 610-7689, Email: peless@utrc.utc.com

*Corresponding author

ABSTRACT

Whole building energy models do not always provide satisfactory predictions to facilitate decision making during design, due to large number of uncertainties in model input parameters. In this work we present a computationally efficient process for uncertainty quantification, sensitivity analysis and automated calibration of building models. We demonstrate our methodology using an energy simulation model of a medium sized office building.

1. INTRODUCTION

Buildings are complex systems with large number of parameters that drive its energy and comfort performance, such as equipment efficiencies and infiltration rates that are often not accurately measured, estimated or recorded. Furthermore, building performance during operation is influenced by processes that are inherently uncertain, such as occupancy or weather patterns, that are never well characterized or understood during design. Thus, uncertainty quantification and management throughout the building modelling, design and operation process is essential for reliable performance prediction and tracking. This is particularly important for energy efficient building retrofit design where energy simulation models must often be calibrated to as-operated building state prior to evaluation of retrofit options that provide the most favorable cost-benefit trade-off. In order to guarantee a commercially viable retrofit, the designer needs to provide reliable confidence bounds for the retrofitted building performance in the early design stages, and not just annual simulations for nominal conditions. However, characterizing the effects of uncertainties in nominal building simulations is a computationally challenging exercise and is therefore ignored if not simplified.

Uncertainty quantification methods either converge slowly (e.g., Monte Carlo) or are cursed by dimensionality (e.g., probabilistic collocation), and are therefore effective only when the number of uncertain parameters is small (< 10). Building models typically have several hundred parameters, so such uncertainty quantification methods cannot be applied directly. Fortunately, in most cases, each aspect of building energy performance is strongly influenced by only a handful of parameters. Therefore, by conducting efficient sensitivity analysis, one can identify key building parameters and conduct further analysis for those parameters alone.

We present a methodology to automate building model calibration and uncertainty quantification in a computationally tractable manner that would have turnaround times that are acceptable within current design and simulation environment.

Our methodology is based on probabilistic global sensitivity analysis [1–5]. Probabilistic data required for the sensitivity analysis is collected using commonly used building simulation packages, such as TRNSYS or EnergyPlus. This is the most computationally expensive part of the process and requires large scale parallel simulation runs. Based on sensitivity results, we derive a reduced order model, which is then used for parameter calibration with respect to available calibration data (e.g., from sensor measurements). Calibrated parameters are then fed back to the original model to improve predictive capability. The methodology also provides output ranges and confidence bounds for the calibrated model. The reduced order model could also be used for retrofit design optimization and optimal control design.

In Section 2 we outline our building analysis and calibration process. We describe the sampling methodology that was used to generate probabilistic data for sensitivity analysis in Section 3. Reduced order model generation and sensitivity analysis is demonstrated in Section 4, and model calibration is presented

in Section 5. Finally, we discuss mathematical methods and technical challenges involved with uncertainty management during energy efficient building retrofit design in Section 6. We also discuss limitations of our methodology and describe technology gaps.

2. ANALYSIS PROCESS PROTOTYPE

Since uncertainty quantification methods are either slow to converge or cursed by dimensionality, the computational cost for uncertainty analysis typically grows exponentially as the size of the problem increases. This means that simply employing more computational power may not be sufficient. For example, to perform uncertainty quantification for 200 uncertain building parameters, taking into consideration parameter interactions up to $O(10)$ and using full collocation grid, it would take roughly 351 billion years for computation to complete on a single processor machine. There is no physical solution in sight that can compress such simulation to run within a reasonable amount of time. Therefore, it is necessary to develop mathematical methods that scale better with the size of the problem.

Current state of the art methods for uncertainty quantification do not remove the curse of dimensionality altogether. Practical improvements have been accomplished recently by increasing number of uncertain parameters that can be studied simultaneously to several hundreds from less than ten. Particularly effective are methods for detecting uncertain parameters to which system is not sensitive. These methods scale linearly and can help reduce the size of the problem significantly. There are other techniques, such as quasi-random number generation, which improve convergence of Monte Carlo methods [6], as well as sparse grid methods [7], which can significantly reduce amount of probabilistic samples needed to cover given volume in parameter space. To be effective, uncertainty analysis must be carefully tailored to the system being studied. Therefore, there is a need to develop not only new uncertainty analysis methods, but also to develop processes that utilize those methods in the most efficient way.

Commonly used buildings simulation packages rely heavily on legacy code and embedding numerical solvers within such packages is very difficult, often impossible. Fortunately, efficient uncertainty analysis can be carried out entirely with black-box methods. These methods use repeated simulations of a deterministic model to produce statistical data needed for uncertainty quantification. Each deterministic simulation is performed with input parameters perturbed to reflect uncertainty associated with them. General black-box uncertainty analysis process has three key steps: (i) generating probabilistic parameter samples, (ii) running simulations for each sample and (iii) performing analysis on simulation data.

Parameter perturbations are performed according to probabilistic rules. For example, parameters can be selected randomly from input probability density function. Each deterministic simulation is run for different sample of input parameters. Running deterministic simulations is the most time consuming part of the process, which typically takes days, whereas the rest of the process takes only minutes to execute. Output data from simulations is then analysed in order to assess and quantify effects of input uncertainty. Analysis can be as simple as calculating average value of all simulation trials, or it can be an elaborate multi-stage process [8].

The guiding principle for the proposed uncertainty analysis process is to reuse much of existing tools and resources to make it easy to adopt by building design practitioners. We designed our uncertainty analysis code as a wrapper around a standard buildings simulator so that same models can be used for both, deterministic simulations as well as uncertainty analysis. The analysis itself does not depend on the specific simulator. The prototype code described here has been designed to work with TRNSYS, but with a suitable plug-in, the code can be interfaced to any building simulator.

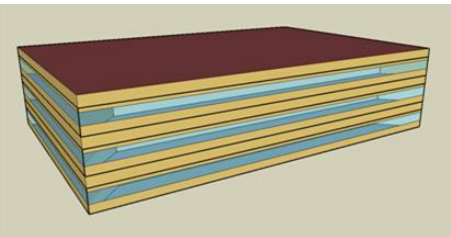


FIG. 1: Model of DOE medium-size office benchmark building.

We demonstrate analysis using a TRNSYS whole building energy simulation model of DOE medium-sized office benchmark building [9], shown in Figure 1. This is a three storey building with total inside area of 15,000 ft^2, divided into 15 heating/cooling zones.

3. GENERATING PROBABILISTIC SAMPLES FOR UNCERTAINTY ANALYSIS

3.1 Sampling Techniques

In order to perform analysis, sufficient amount of probabilistic data must be produced by running repeated deterministic simulations. Each deterministic simulation is run with uncertain parameters perturbed according to a mathematical prescription. In the simplest case, we choose parameters by sampling them randomly from probabilistic density functions for each uncertain input (Monte Carlo method). The sampling range and the type of the probabilistic density functions must be set by the modeller/designer. The uncertain input properties are not always known. Equipment or material manufacturers may have to conduct extensive material testing and compile deviations from nominal values into a probability density function. More often, however, no reliable information is available, so the modeller must conduct testing on his/her own or estimate uncertain input properties based on prior experience.

Random sample $x^i = (x^i_1, \ldots, x^i_d)$ is computed under assumption that all uncertain inputs are uncorrelated, so uncertain parameter values x^i_j are calculated independently as

$$x^i_j = \bar{x}_j + \varepsilon_j \rho_j^{-1}(\xi^i_j) \tag{1}$$

where $\bar{x}_j$ is nominal value of the uncertain parameter, ε_j is relative tolerance, ρ_j is cumulative probability density function describing type of the uncertainty, and $0 < \xi_j < 1$ is a random number. In our analysis we use Monte Carlo and quasi Monte Carlo methods. The two differ in the way random numbers ξ^i_j are generated. The former uses plain random number generator, while the latter uses quasi random number sequences, which ensure more uniform sampling space coverage. We refer to $x^i = (x^i_1, \ldots, x^i_d)$ as a physical sample, and to $\xi^i = (\xi^i_1, \ldots, \xi^i_d)$ as mathematical sample. Convergence of Monte Carlo methods is slow and numerical error scales as $1/\sqrt{N}$, where N is the number of random samples. The numerical error of quasi Monte Carlo method is theoretically evaluated to be of order $(\log N)^d/N$, where N is the number of samples [6]. In some cases it is observed that the error diminishes even faster, at $1/N$ rate. Quasi Monte Carlo implementations we used are Joe-Kuo variant of Sobol sequence [10] and Halton sequence with Kocis-Whiten scrambling [11].

3.2 Input Data

Model developer needs to supply following data for each uncertain parameter in the building model: (i) nominal value, (ii) tolerance, (iii) type of uncertainty and (iv) polynomial expansion order. The nominal values for all parameters are already included in the model. In our approach we enter additional information required for uncertainty analysis through TRNSYS graphical user interface, as well.

TRNSYS mark-up uses the exclamation mark symbol to denote a comment in the input file. Everything entered after this symbol is ignored by the simulation engine. The graphical user interface allows user to enter comments next to variables as shown in Figure 2. To take advantage of this feature, we defined tag "!UNC" to denote uncertain input data entry. For each uncertain input user adds a comment that begins with this tag and is followed by value for the tolerance, uncertainty type tag and polynomial expansion order separated by blank spaces.

Tolerance is entered as a relative tolerance, a number in interval $(0, 1)$. The absolute tolerance is calculated by multiplying this number to the nominal value of the uncertain parameter. If the nominal value is zero, then the tolerance is interpreted as the absolute tolerance expressed in the same units as the nominal value. Type of uncertainty is entered as a string, e.g. "UNIFORM" for uniformly distributed input, "NORMAL" for Gauss distributed input, etc. Polynomial expansion order is entered as a positive integer and is used to specify polynomial order for response surface fitting.

Our code can extract all of this information from TRNSYS input file and use it for the analysis. Once samples are computed, the code generates TRNSYS input files for each sample, where nominal parameter values are replaced with sampled values. TRNSYS simulations are then scheduled for each input sample, and information necessary for the analysis is extracted from TRNSYS output files once simulations are completed.

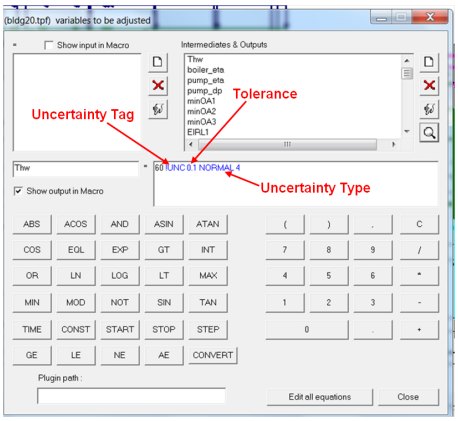


FIG. 2: TRNSYS graphic user interface is used to enter uncertain inputs data. The data is stored as a comment in the TRNSYS input file.

4. SENSITIVITY ANALYSIS

Typical building model has hundreds of uncertain parameters and because of that sensitivity analysis methods that can handle high dimensional problems are at the core of the building performance analysis. Variance-based methods are particularly important since even when they fail to capture all sensitivities in the system, they can provide a quantitative measure of the amount of sensitivity in the system that has not been accounted for.

For sensitivity analysis we first account for sensitivities due to single parameter perturbations. Such computation scales *linearly* with the number of uncertain parameters, and large problems can be handled with more computational power. From there we find first order sensitivity indices and amount of sensitivity not captured by first order analysis. If unaccounted sensitivity is a small fraction of the overall sensitivity, then all necessary information for further analysis is contained within sampling data. Otherwise, higher order sensitivity analysis needs to be done, preferably after some model reduction based on first order analysis results.

A building model $f(x)$, with uncertain parameters $x = (x_1, \ldots, x_d)$, $\quad 0 \leqslant x_i \leqslant 1$, can be expanded in terms of analysis of variance (ANOVA) decomposition as

$$f(x) = f_0 + \sum_{i=0}^{d} f_i(x_i) + \sum_{i=0}^{d} \sum_{j<i} f_{ij}(x_i, x_j) + \ldots + f_{1,2,\ldots,d}(x_1, x_2, \ldots, x_d) \tag{2}$$

where

$$\int_0^1 f_{i_1 i_2 \ldots i_s}(x_{i1}, x_{i2}, \ldots, x_{is}) dx_{i_k} = 0, \ \forall k : \ 1 \leqslant k \leqslant s \tag{3}$$

Here we assume without loss of generality that all uncertain parameters $x_i \in (0, 1)$. From the definition (2) it follows that f_0 is the mean value of $f(x)$. Variance expansion is then given as

$$D = \sum_i D_i + \sum_i \sum_{j<i} D_{ij} + \ldots + D_{1,2,\ldots,d} \tag{4}$$

where

$$D_{i_1 i_2 \ldots i_s} = \int_0^1 f^2_{i_1 i_2 \ldots i_s}(x_{i1}, x_{i2}, \ldots, x_{is}) dx_{i_1} \ldots dx_{i_s} \tag{5}$$

is a partial variance and

$$D = \int_0^1 (f(x) - f_0)^2 dx \tag{6}$$

is the total variance. Sobol sensitivity index is defined as a fraction of partial and total variance

$$S_{i_1 i_2 \ldots i_s} = \frac{D_{i_1 i_2 \ldots i_s}}{D}. \tag{7}$$

By definition all Sobol indices sum up to one.

$$\sum_i S_i + \sum_i \sum_{j<i} S_{ij} + \ldots + S_{1,2,\ldots,d} = 1 \tag{8}$$

First order Sobol index is defined as

$$S_i = \frac{D_i}{D} \tag{9}$$

When $S_i \approx 1$ the system is sensitive only to x_i. Sum of all first order indices gives the fraction of uncertainty due to single parameter perturbations. The value of this sum helps determine if higher order sensitivity analysis is necessary to assess overall system sensitivity. Another useful quantity is first order *total* Sobol index, which is defined as

$$T_i = \frac{D_i + \sum_j D_{ij} + \sum_j \sum_{k<j} D_{ijk} + \ldots + D_{1,2,\ldots,d}}{D} \tag{10}$$

When $T_i \approx 0$ the system is not sensitive to x_i. Total Sobol index is typically used to eliminate parameters that the system is not sensitive to from further analysis. Finally, derivative based first order sensitivity index is defined as

$$v_i = \int_0^1 \left(\frac{\partial f}{\partial x_i} \right)^2 dx \tag{11}$$

Relationship between the three types of first order sensitivity indices is:

$$0 \leqslant S_i \leqslant T_i \leqslant v_i \leqslant 1 \tag{12}$$

Derivative based sensitivity index is the upper bound of the total Sobol index, but in practice it is much cheaper to evaluate, and is used instead of the total Sobol index to eliminate low sensitivity parameters.

Mean value f_0 and total variance D are computed directly from the simulation data. Partial variances could be calculated from separate sets of simulation data where only a subset of uncertain parameters is perturbed. However, computationally more efficient way to compute partial variances is response surface method where the data for the mean and total variance calculation can be reused. Let us assume that model $f(x)$ can be expanded in terms of orthonormal polynomials ϕ_k and consider response surface in the form

$$f_j(x_j) = \sum_{k=0}^{\infty} a_{jk} \phi_k(x_j), \tag{13}$$

where

$$\int_0^1 \phi_i(x) \phi_j(x) \rho(x) dx = \delta_{ij}. \tag{14}$$

For simplicity and without loss of generality we will assume that $\rho(x) = 1$. Partial variance then can be expressed in terms of polynomial expansion coefficients as

$$D_j = \sum_{k=1}^{\infty} (a_{jk})^2, \tag{15}$$

and therefore Sobol first order index can be computed as

$$S_j = \frac{1}{D} \sum_{k=1}^{\infty} (a_{jk})^2. \tag{16}$$

If polynomial expansion converges, we can make a cut-off at suitable order P and approximate Sobol index as

$$\hat{S}_i = \frac{1}{D} \sum_{k=1}^{P} a_k^2. \tag{17}$$

Obviously, only a finite approximation at relatively small order is useful for practical applications. Response surface for the entire system is

$$f(x) \approx f_0 + \sum_{j=1}^{d} \sum_{k=1}^{P} a_{jk} \phi_k(x_j) \tag{18}$$

Polynomial coefficients can be calculated from simulation data using for example least square fit. The number of polynomial coefficients to find, and therefore computational cost, grows linearly with the dimension of the system d.

4.1 Sensitivity Analysis Example

Here we demonstrate sensitivity analysis for 217 uncertain parameters with respect to 16 outputs in DOE benchmark building model. The outputs are annual energy consumptions for different end uses in the building. A total of 4,500 quasi Monte Carlo simulations was run, using TRNSYS as the simulation engine. First order sensitivity analysis (Figure 3) shows that there are few parameters that the system is sensitive to; there are only 19 parameters that contribute to 10% or more of overall sensitivity for at least one output.

For example, annual natural gas consumption sensitivity is almost entirely dominated by the parameter representing the efficiency of the boiler (67%) followed by the hot water temperature (8%). All other parameters combined contribute to less than 25% of the gas consumption sensitivity. The sum of all first order Sobol indices for gas consumption is 99.5%, suggesting that higher order sensitivities are negligible and no further sensitivity analysis is needed. From this one can further deduce that model for gas consumption can be reasonably well calibrated by adjusting only two uncertain parameters.

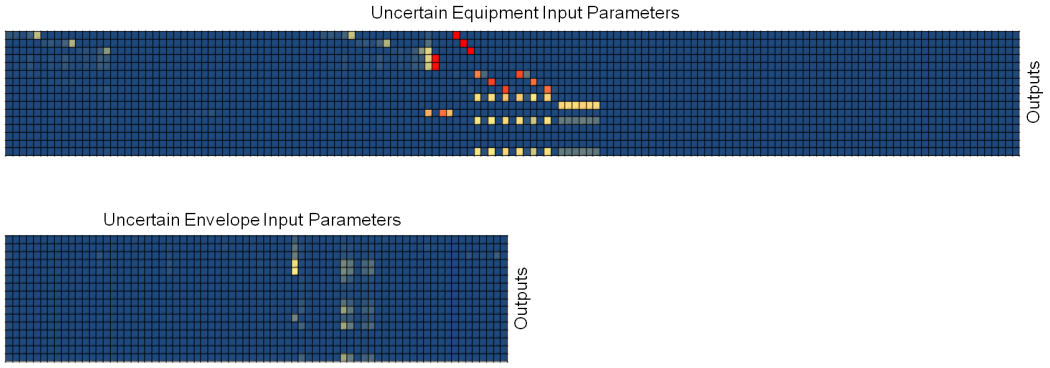


FIG. 3: First order Sobol indices for 217 uncertain input parameters calculated for 16 system outputs. Color code denotes value of the Sobol index with dark blue indicating small and red large value of the index. Results for equipment parameters and building envelope parameters are shown on separate panels. Parameter and output labels cannot be shown on this scale.

In order to assess accuracy and reliability of results, computations of sum of first order Sobol indices $\sum_i S_i$ and relative error of the least square fit

$$\epsilon = \frac{\|Ax - B\|}{\|B\|} \tag{19}$$

were included in the sensitivity analysis. These two quantitative measures help determine whether the analysis was successful and help decide if and what additional analysis needs to be performed. Figure 4 shows these measures evaluated for all system outputs.

Large error in response surface fit may indicate that there were too few probabilistic samples to get an accurate fit or that the form of the response surface is not suitable for the problem at hand. The latter means that either some important uncertain parameters were omitted from the analysis or that higher order sensitivities (those due to combined effects of several parameter perturbations) contribute significantly to the overall sensitivity.

If the sum of first order Sobol indices is significantly smaller than one, that suggests higher order sensitivities need to be investigated, and possibly some important parameters were not included in the analysis. If the sum is very close to zero, then it is quite certain that analysis needs to be expanded to cover additional parameters. It is highly unlikely that sensitivity is contained in higher orders, and at the same time single parameter perturbation do not affect the system significantly.

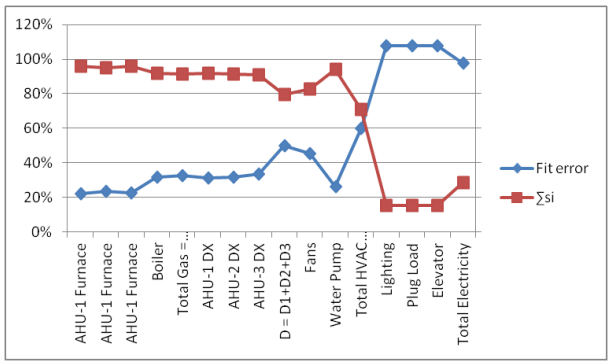


FIG. 4: Quality of sensitivity analysis is checked by evaluating response surface fit error and sum of all Sobol indices.

In the example given in Figure 4 it is shown that sensitivity analysis was successful for all but three outputs, since in those cases response surface error was small. Furthermore, first order Sobol indices sum up almost exactly to one, suggesting that no further analysis is required. The three outputs where sensitivity analysis failed are energy consumptions of lighting, plug loads and the elevator. The sum of first order sensitivity indices for each of those is zero, suggesting that parameters that affect these three outputs were not included in the analysis. Upon closer look one can verify that this is indeed the case. Energy consumption in these cases depends almost exclusively on usage schedules (e.g. how long during the day the lights are on) rather than on any static parameters. Currently our code supports only static parameter sensitivity analysis, so lighting, plug load and elevator schedules were not included in the analysis. This is what the two quality measures indicated. Time varying uncertainty and sensitivity analysis will be addressed in another publication.

5. MODEL CALIBRATION

The sensitivity analysis and response surface (reduced-order model) of the original function $f(x)$ can then be used to calibrate the full-order model; that is, given a set of field or experimental measurements, the analysis could be used to compute the parameters that maximize the agreement of the model with the data. The response surface of the function $f(x)$ was described as

$$y = \tilde{f}(x) = f_0 + \sum_{j=1}^{d} \sum_{k=1}^{P} a_{jk} \phi_k(x_j) \approx f(x). \tag{20}$$

In general, the original function can be a vector $\mathbf{f}(x) = \big(f_1(x), f_2(x), \ldots, f_N(x)\big)$, in which case the response function is

$$y_n = \tilde{f}_n(x) = f_{0n} + \sum_{j=1}^{d} \sum_{k=1}^{P} a^n_{jk} \phi^n_k(x_j) \approx f_n(x), \quad n = 1, 2, \ldots, N. \tag{21}$$

The calibration procedure is described as follows: The significant parameters indicated by large Sobol indices are retained in the response function (21), while the rest are omitted. Using the measurement

data of the outputs y_n, which typically represent physical quantities such as monthly/annual electricity/gas consumption, we formulate an optimization problem to estimate parameters that minimize the difference between the model predictions and available data. The steps involved in the calibration process are now described in detail.

As a first step, for each output, order the parameters in the decreasing order of their Sobol indices. Compute their cumulative Sobol indices, and retain the parameters $\tilde{x}$ with a cumulative index smaller than a cut-off, say 0.95. Alternatively, retain all parameters with corresponding Sobol indices greater than a certain user-defined cut-off, say 0.05. Redefine the response surface (21) by retaining only terms containing the reduced set of parameters.

Then, given measurements c_i, the calibration problem can be formulated as:

$$\min_{\tilde{x}} J[\tilde{x}] = \sum_{i=1}^{N} \alpha_i (y_i(\tilde{x}) - c_i)^2, \tag{22}$$

$$0 < \tilde{x}_i < 1, \qquad i = 1, 2, \ldots, d. \tag{23}$$

The constants α_i are used to normalize the outputs with respect to certain nominal values; here, these are chosen to be the values of the function at the center of the domain $[0, 1]^d$. The optimization problem (23) can be solved using standard off-the-shelf solvers such as IPOPT [12] or NLOpt [13]. In this work, we just present a demonstration of the capability that is to be built into the tool-chain in the future. For that purpose, we use standard algorithms built into MATLAB®; in particular, we use the function fmincon, with an interior-point algorithm to solve (23).

Finally, to test the calibration, we compare the predictions of the calibrated model to the measurement data. That is, if the solution of the optimization problem (23) is $\bar{x}$, compute the error $e = \|\mathbf{f}(\bar{x}) - \mathbf{c}\|/\|\mathbf{c}\|$.

5.1 Calibration Example

Here, we describe the process involved in model calibration, using the response surface and results of sensitivity analysis described in section 4. The process itself was outlined for an arbitrary model in section 5, and here we describe its application to a TRNSYS model of the DOE benchmark building.

Recall, from section 4, that the response surface $\tilde{\mathbf{f}}(x)$ of the model $\mathbf{f}(x)$ is a polynomial function of all the original 217 parameters. The Sobol indices, which are an indicator of the importance of the parameters to the measured outputs, can be used to reduce the number of parameters defining the response surface. Here, we retain only the parameters with Sobol indices $S_i > 0.05$ for any of the given outputs, which sharply reduces the number of parameters from 217 to 29. Next, define the response surface in terms of the reduced number of parameters, $\tilde{\mathbf{f}}(\tilde{x})$. Given the measurements of the 16 outputs c_i, we then define a cost function similar to that in (23). The resulting optimization problem is solved using Matlab function fmincon and results in parameters that calibrate the model.

In this case, since field or experimental data was not available, we compute this as follows. For the sensitivity and uncertainty analysis using the sampling method, we evaluated the TRNSYS model at multiple samples (4500 in number), and the resulting outputs define a distribution, with a mean μ_i and variance σ_i^2 that can be numerically computed. The distribution of total gas consumption, along with the mean and one standard deviation are shown in Figure 5. We assume that the calibration data are defined as $c_i = \mu_i + \beta\sigma_i$, where $\beta = 1.5$ is an arbitrarily chosen parameter. We considered a wide range of values of β for calibration and report a representative one. However, note that the calibration data should be within the spread of the output distribution for meaningful results; for data outside of this range, the response surface is no longer accurate.

The calibration parameters are finally substituted in the full TRNSYS model and the error in predicting all the outputs is computed. A comparison of the error, both before and after calibration, is shown in Figure 6. The calibration data, along with the model predictions (before and after calibration) are shown in Figure 7.

6. CONCLUSIONS AND FUTURE CHALLENGES

We demonstrated a systematic and scalable approach for sensitivity analysis with building energy simulation models. The reduced order model (or meta-model) obtained as part of the process is re-used for automated calibration of the high fidelity building energy model from which it was derived. The advantage of meta-model is that it returns function evaluation in a fraction of a second as opposed to the high fidelity

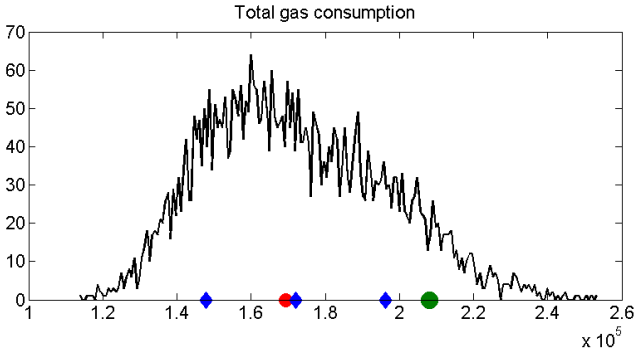


FIG. 5: Histogram of total gas consumption, obtained from the 4500 samples used for sensitivity analysis. Also shown are the calibration data (green, square), the nominal output (red, circle), and the mean and single standard deviations (blue, diamond)

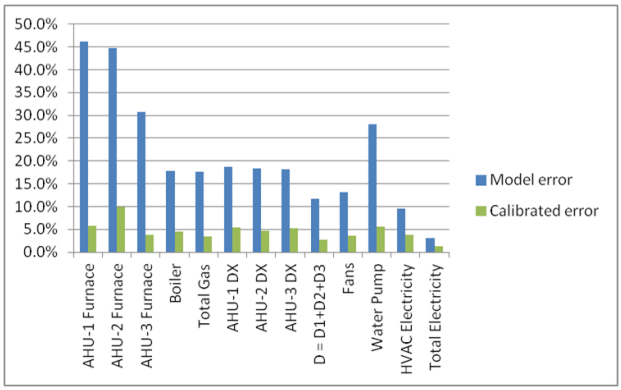


FIG. 6: DOE medium size benchmark building model calibration: error in predicting various outputs, before (blue) and after (green) calibration.

model, which takes several minutes to simulate. Furthermore, the meta-model is obtained in a simple analytical form, so it can be easily manipulated and used in various optimization algorithms. We demonstrated the successful calibration of the DOE benchmark building model (with higher fidelity simualtion-generated data), where we reduced model prediction errors that ranged from 10-40% to less than 5% each.

Our analysis was limited to uncertainties with respect to static design parameters. Building performance is also significantly affected by uncertain processes such as weather or occupancy. These must be included to complete the building performance analysis. However, methods to quantify time varying uncertainties is

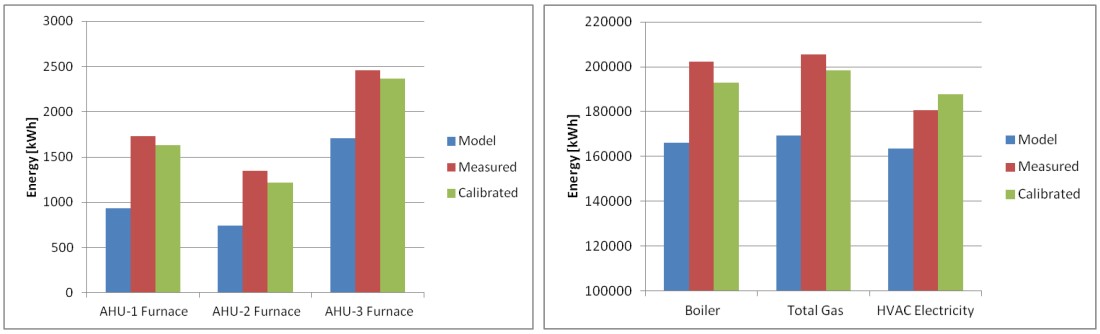


FIG. 7: DOE medium size benchmark building model calibration. The figure compares the model predictions of five different quantities, before (blue) and after (green) calibration, with the assumed measurements (red).

not as mature as methods for parametric uncertainty, and this is an active research area.

Our approach provides a reliable and fully automated way to assess first order parametric sensitivities in building models. Higher order sensitivities are computationally more costly to evaluate. For example, to perform second order sensitivity analysis for the same set of parameters we would need order of 10^7 numerical samples instead of 4500 we needed for the first order analysis. Therefore, to perform a higher order analysis with the same computational capability, the dimensionality of the problem needs to be reduced. There is still no automated way to do this. Nevertheless, first order analysis results and domain knowledge expertise may provide sufficient information to eliminate insignificant uncertain inputs and perform efficient higher order analysis on a subset of sensitive parameters. In the benchmark building example, first order analysis was sufficient for almost all of the analysis. We anticipate that for more elaborate building design problems, especially in cases with deep energy efficient retrofits, a systematic method for dimensionality reduction and system decomposition is needed.

ACKNOWLEDGEMENTS

This work is funded by Energy Efficient Buildings Hub, sponsored by the Department of Energy under Award Number DE-EE0004261.

[1] Genyuan Li, Carey Rosenthal, and Herschel Rabitz. High dimensional model representations. *J. Phys. Chem. A*, 105(33), 2001.

[2] I.M. Sobol. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and Computers in Simulation*, 55:271280, 2001.

[3] I.M. Sobol, S. Tarantola, D. Gatelli, S.S. Kucherenko, and W. Mauntz. Estimating the approximation error when fixing unessential factors in global sensitivity analysis. *Reliability Engineering & System Safety*, 92(7):957 – 960, 2007.

[4] Bryan Eisenhower, Zheng ONeill, Satish Narayanan, Vladimir A. Fonoberov, and Igor Mezic. A methodology for meta-model based optimization in building energy models. *Energy and Buildings*, 47:292 – 301, 2012.

[5] Zheng ONeill, Bryan Eisenhower, Shui Yuan, Trevor Bailey, Satish Narayanan, and Vladimir Fonoberov. Modeling and calibration of energy models for a dod building. In *Proceedings of ASHRAE Annual Conference*, Montreal, Quebec, Canada, 2011.

[6] H. Niederreiter. Random number generation and quasi-monte carlo methods. In *CBMS-NSF Regional Conference Series in Applied Math*, volume 63. SIAM, 1992.

[7] Thomas Gerstner and Michael Griebel. Numerical integration using sparse grids. *Numerical Algorithms*, 18(3-4):209–232, 1998.

[8] Xiu Yang, Minseok Choi, Guang Lin, and George Em Karniadakis. Adaptive anova decomposition of stochastic incompressible and compressible flows. *Journal of Computational Physics*, 231(4):1587 – 1614, 2012.

[9] P. Torcellini, M. Deru, B. Griffith, K. Benne, M. Halverson, D. Winiarski, and D.B. Crawley. Doe commercial building benchmark models. In *2008 ACEEE Summer Study on Energy Efficiency in Buildings*.

[10] Stephen Joe and Frances Y. Kuo. Constructing sobol sequences with better two-dimensional projections. *SIAM J. Sci. Comput.*, 30:2635–2654, 2008.

[11] Ladislav Kocis and William J. Whiten. Computational investigations of low-discrepancy sequences. *ACM Trans. Math. Softw.*, 23:266–294, June 1997.

[12] Andreas Wachter and Lorenz T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming*, 106:25–57, 2006.

[13] Steven G. Johnson. The nlopt nonlinear-optimization package. `http://ab-initio.mit.edu/nlopt`, 2011.