UNCERTAINTY-WEIGHTED META-MODEL OPTIMIZATION IN BUILDING ENERGY MODELS

Bryan Eisenhower\textsuperscript{1}, Vladimir Fonoberov\textsuperscript{2}, and Igor Mezić\textsuperscript{1}

\textsuperscript{1}Department of Mechanical Engineering
University of California, Santa Barbara
\textsuperscript{2}AIMdyn Inc., Santa Barbara, California

ABSTRACT
In this paper we discuss a method to efficiently optimize whole-building energy models which contain a large number of candidate optimization variables and have discontinuous cost surfaces with multiple minima. The approach leverages uncertainty propagation and sensitivity analysis to identify critical parameters that are most effective for optimization. Large discontinuities in the cost function are identified using a filtering method and reduced order meta-models are created in partitioned subsets of the global feasible set that are separated by these discontinuities. A derivative-free optimization algorithm is used while employing an uncertainty-weighted cost function to obtain the best optimized solution for each feasible subset. This method is tested on a large office building modeled in EnergyPlus but can be adapted to other models and modeling software.

INTRODUCTION
Elaborate whole-building energy models, which predict a buildings energy consumption and comfort, are currently being used for design (of the building layout, construction, and operation) as well as to establish compliance for various accreditations. For many reasons, there has been a push for high performance buildings which require a highly integrated and optimized design and operation strategy.

Model-based optimization of a building design or operation strategy has been performed since the 1970’s. In most cases, a small set of parameters of a building are selected and a numerical optimizer is integrated with the whole-building energy model which provides a cost function which the optimizer attempts to minimize. Some of the more recent studies include Wang et al. (2005), Wetter and Wright (2003), Djuric et al. (2007), and Kämpf et al. (2010).

Unfortunately, using a whole-building energy model as a means to evaluate a cost function in an optimization experiment is computationally expensive. Often, the cost function contains metrics on the timescale of one year, which takes multiple minutes to calculate for a single set of optimization variables. In addition to this, the response of the model is not in closed-form and there exist discontinuities that challenge evolution of the optimizer (Wetter and Polak (2004)). Because of this, there has been a recent trend to utilize derivative-free optimization approaches for building optimization (for example Kämpf et al. (2010), or Hamdy et al. (2011), or the software in Wetter (2004)). In this study, we use a derivative-free optimization algorithm in conjunction with a pre-processor to identify large discontinuities and subsequently partition the feasible set into partially continuous subsets (we say partially because there may exist some small discontinuities in these smaller subsets).

Derivative-free optimization algorithms help to alleviate issues that arise in the cost surface which may lead an optimizer to converge to local minima, but the function evaluation of a whole-building energy model is still a computational burden. The use of a meta-model is one way to speed up the optimization process and alleviate this burden.

Meta-model based optimization is an approach which derives a model of a model from a data set which is either challenging to compute or has no closed-form with respect to its optimization variables, and produces a closed-form model which can be rapidly evaluated (Barton and Meckesheimer (2006), Knowles and Nakayama (2008), and Jin et al. (2003)). Meta-modeling has been used in the buildings community to predict energy usage (Dong et al. (2005)) and to perform sensitivity analysis on uncertain parameters (Mara and Tarantola (2008)), as well as for fault detection (Brown et al. (2011)), or for building optimization (Eisenhower et al. (2012)). In (Stavrakakis et al. (2012)) an Artificial Neural Network (ANN) meta-model was created from the output of a computational fluid dynamics model of a zone which was then used for numerical optimization.

One of the challenges with any approach that fits a model to data (whether the data is from a whole-building energy model, or an array of sensors) is the impact of uncertainty in the meta-model. In this paper we address this uncertainty in a multi-faceted way. First, we perform sensitivity analysis on the whole-building energy model. Once this is performed, we are able to identify the most critical parameters of the model (see Eisenhower et al. (2011)). It is essential to identify these critical parameters as fitting a model with hundreds of ancillary variables is numerically ill-posed. In addition to this, it is essential to identify critical parameters for the optimization process to minimize the expense of optimizing parameters which have little influence on the cost.

Once a subset of the parameters of a whole-building energy model...
energy model are identified, a reduced order meta-model is created for the optimization process. There is some uncertainty in this approximation to the full model and we therefore treat the cost function from this meta-model as an uncertain surface. Incorporating uncertainties into the cost surface provides a means to capture any model mismatch while it also helps to address any minor discontinuities as well as multiple minima in the cost surface.

In the next section we will discuss the general methods used in this paper by expanding on some of the topics in this introduction. We will then introduce a building model which is used as a case study and present specific results for each of the steps of the process, culminating in optimization results for a large office building.

**METHODOLOGY**

In this section we discuss the general methodology of the optimization process. We present each subject of the method in the same chronology that they are performed. Some of the specifics of each step are presented in the following section which implements the approach on a case study.

**Parameter sampling and uncertainty propagation:**

In order to calculate the meta-model, all possible optimization parameters (100’s) are identified in the full whole-building energy model, and feasible ranges are selected for each of these parameters. Numerous samples are chosen within these ranges and a deterministic (Quasi Monte Carlo) sampling approach (Eisenhower et al. (2011) or Burhenne et al. (2011)) is used to vary all optimization parameters at the same time creating many models (1000’s) that characterize the energy model around its baseline. The software used to calculate these samples is available at Aimdyn GoSUM Software (2012). Once the simulations are performed, a meta-model is fit to desired outputs of the model (e.g. yearly energy consumption, peak demands, average comfort, etc.).

We use a support vector regression (SVR) machine learning approach to generate the meta-model for key outputs of the original energy model (Smola and Scholkopf (2004)). The SVR approach attempts to find a function (a meta-model) whose deviation from data is at most a small constant (e.g. $\varepsilon$), which in turn defines a region where errors are accepted. An optimization problem is then formulated with a parameter that trades flexibility in the identified meta-model (the number of its support vectors) and tolerance to deviations greater than $\varepsilon$. The SVR approach is similar to learning in ANN, while one significant difference is that for ANN, many local optimal solutions may exist for the task of fitting the meta-model parameters, while in SVR, only one global solution exists.

**Discontinuity detection:** One of the most important steps to any optimization problem is defining the feasible region for which optimization variables are allowed to vary. In this paper, the global feasible region is defined by identifying realistic ranges during the parameter sampling step. However, as mentioned in the introduction, whole-building energy models often have discontinuities in them that are significant enough to challenge the performance of any optimization algorithm. To overcome this issue, and to identify discontinuous feasible regions, we pre-process the energy model data to identify large discontinuities. In order to do this, we rank-order parameter values and investigate the response of the outputs as these parameters are swept through their range. A filter is constructed to minimize noise and to identify large changes in outputs of the model as the parameter is varied. Once this has been performed, optimization in each of the feasible regions (which are each a subset of the original feasible set) is performed in parallel.

**Sensitivity analysis and model reduction:** As mentioned in the first step, we characterize the energy model with respect to all practical optimization parameters in the original energy model. Although meta-model based optimization is computationally very fast, for many reasons, it is not beneficial to optimize over all practical optimization variables. First, if the optimization project is to be implemented in real life, it is not likely that a building designer or operator will be able to alter 100’s of variable values - it is only cost effective to attack a few of the most critical variables in a building.

Secondly, the process of SVR calculates coefficients to Gaussian kernels which are a function of all the parameters in the desired meta-model. Each kernel depends on all parameters with the same weight, which in turn means that parameters that have little significance add noise to the calibrated kernel. In addition to this, using the sampling approach characterizes the model in center of the volume the best with less information on the tails of distributions (or corners of the volume). The number of corners (vertices) of a volume of this type increases as $2^N$ where $N$ is the number of optimization variables and the density of information in these corners goes to zero as the dimension increases.

To account for both of these concerns, sensitivity analysis is performed on sampled data to identify critical parameters in the model. These critical parameters are identified by rank ordering sensitivity indices which identify the correlation between variance in parameters (optimization variables) and key outputs (see Saltelli et al. (2000) or Mara and Tarantola (2008)). Sensitivity analysis can be performed using different norms including a derivative-based global sensitivity index (Sobol’ and Kucherenko (2009))

$$
\mu_m = \int \left| \frac{\partial f}{\partial x_m} \right| dx, \quad (1)
$$

where the integration is performed over all dimensions of the sampling points. This is an estimate of the derivative of the meta-model ($f$), where $\mu_m$ is an $L^1$
Numerical optimization under uncertainty: Once an accurate meta-model has been created, function evaluations are computationally very efficient, and therefore rapid optimization experiments can be performed. To solve the optimization problem, we use a derivative-free method (NOMAD) which contains the Mesh Adaptive Direct Search (MADS) algorithm, which is a direct search algorithm with rigorous convergence properties (Le Digabel (2011)).

The second step that we mentioned above partitions the feasible set into subregions that are separated by large discontinuities. These partitioned subsets are free of large discontinuities but still contain smaller discontinuities and may possess a non-monotonic response to optimization variables. To account for this, as well as uncertainty in the accuracy of the meta-model, we treat the cost function as a sum of uncertain elemental costs

\[ C(p) = \alpha_1 C_1(p) + \alpha_2 C_2(p), \]

where \( C \) is a total cost function, \((C_i)\) are elemental costs, \( p \) are optimization variables, and \( \alpha_i \) are uncertainty constants. These uncertainty constants are sampled from a normal distribution and the optimal parameter sets are calculated numerous times to identify the best solution to the optimization problem.

CASE STUDY
To test this approach, a United States Department of Energy (DOE) EnergyPlus Benchmark Model was used (Deru et al. (2011)). The DOE benchmark model suite contains 16 models that represent approximately 70% of commercial building stock in the United States. The models are then organized so that each one can be simulated at one of 16 different locations in the US (using typical meteorological year weather data for each of these locations). Each model is also organized by construction type; new construction, existing construction - post 1980, and existing construction - pre 1980.

The model studied in this paper is a new construction large office building located in Las Vegas, Nevada. This location is subject to hot and dry summers and cool winters (relatively extreme in both the summer and winter). The large office building has 12 floors and a basement with 46,320 m² (498,588 ft²) of floor area. The building is a rectangular cube (aspect ratio 1.5), with 38% window to wall ratio, and is zoned using 19 zones (one central zone and one zone for each perimeter side of the building). A plenum is also modeled on top of each group of zones and a basement is included as a zone as well.

The construction of the building contains exterior mass walls (R-value = 1.17 [m²K/W]), and an insulated above deck roof (R-value = 2.79 [m²K/W]). The exterior windows have a U-Factor of 3.24 [W/m²K] and a solar heat gain coefficient of 0.25 [W/m²K], and the building sits on a 4 inch slab floor (with carpet) having an R-value of 0.54 [m²K/W]. For HVAC, the plant contains one gas-fired boiler and two chillers with an economizer integrated into the VAV reheat system. The specifics of the economizer, loads and usage schedules are based on ASHRAE guidelines (e.g., ASHRAE Standard 90.1-2004). The Energy Use Intensity (EUI) of this building is 514.54 [MJ/m²], with dominant contributions from interior equipment 181.00 [MJ/m²], interior lighting 121.68 [MJ/m²], and cooling 95.7 [MJ/m²] (the remaining usage is attributed to exterior equipment and lighting, as well as fans and pumps). All of the EnergyPlus simulations in this paper were performed with version 7.0.0.036.

PARAMETER SAMPLING
The first step in the optimization process is to explore the feasible set around the nominal model using parameter sampling. To accomplish this, all numerical parameters were initially identified in the model as candidate optimization variables. Parameters that define the architectural layout of the building were then omitted as well as some parameters in the model which were not appropriate as optimization variables. For instance, minimizing occupancy may reduce energy consumption, but this would not make the building very useful. With this in mind, we omitted uncertain parameters that are related to clothing, occupant activity, as well as occupancy, work effectiveness, and elevator usage (note that elevator energy consumption was left as an uncertain parameter). What remains was 682 parameters, the largest three types of parameters are: schedules (120), material properties (106), and parameters of the cooling coils (81). Each of these parameters was varied by 25% of their nominal value using a uniform distribution if the parameter has a non-zero nominal value and an exponential distribution if the nominal value is zero. In some cases the 25% range had to be reduced for physical reasons (constraints), or model stability.

Upon defining the ranges for uncertain parameters, 5000 deterministic samples were chosen by varying all parameters simultaneously. What results is 5000 EnergyPlus models, which were simulated in parallel on a 16 core desktop computer. Each simulation took on the order of 3 minutes to simulate, which means the whole batch took less than one day to compute. Once computed, total facility energy was calculated, as well as predicted mean vote (PMV) comfort (Fanger (1970)) for occupied zones during occupied hours (averaged over both). Distributions of these two outputs are presented in Figure 1.
DISCONTINUITY DETECTION

The distributions in Figure 1 seem well behaved, while in fact there does exist significant nonlinearity and discontinuity within the response that makes up these distributions. To identify these discontinuities, a pairing was performed between each uncertain parameter and each of the two outputs. The vector of 5000 parameter values was sorted in an increasing fashion, and the outputs were resorted to match this ordering. The outputs were then filtered of noise while amplifying its derivative using the filter

$$F(s) = \frac{s}{(s + a)^2},$$

where $F$ is an operator in the Laplace domain ($s = \sqrt{-1} \cdot \omega$, where $\omega$ is frequency) similar to a band-pass filter in signal processing. In terms of tuning, we found that $a = 1/100$ worked well for this sampling approach and highlights discontinuities very well (see Figure 2).

This filtering approach takes only seconds and was performed on all variables to identify two parameters that introduced significant discontinuity in the output variables. The response from both parameter 176 (Cooling setpoint, weekdays 06:00-22:00) and 427 (cooling sizing factor) are illustrated in Figure 3 and 4.

By identifying these significant discontinuities, we have partitioned the original global feasible set into four feasible subsets. The interior boundaries for these four sets are presented in Table 1, the exterior boundaries are as identified in the original definition of the ranges of the uncertain parameters.

**Table 1: Criteria for subsets identified by large discontinuities ($\delta_{176} = 23.69$, and $\delta_{427} = 1.29$).**

<table>
<thead>
<tr>
<th>Quadrant</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A</td>
<td>$p_{176} &lt; \delta_{176}$ and $p_{427} &lt; \delta_{427}$</td>
</tr>
<tr>
<td>1B</td>
<td>$p_{176} &lt; \delta_{176}$ and $p_{427} &gt; \delta_{427}$</td>
</tr>
<tr>
<td>2A</td>
<td>$p_{176} &gt; \delta_{176}$ and $p_{427} &gt; \delta_{427}$</td>
</tr>
<tr>
<td>2B</td>
<td>$p_{176} &gt; \delta_{176}$ and $p_{427} &lt; \delta_{427}$</td>
</tr>
</tbody>
</table>

Figure 1: Comfort and total annual energy after 5000 samples of the uncertain parameters. The red dot on the horizontal axis is the nominal simulation results.

Figure 2: Example discontinuity of calculated comfort with respect to parameter 176.

Figure 3: Partitioned simulation results based on discontinuity in parameter 176.

Figure 4: Partitioned simulation results based on discontinuity in parameter 427.
Meta-model Construction

Once the feasible subsets have been identified, a series of reduced order meta-models are calculated for each subset. The sensitivity indices, which identify how influential each parameter is on the outputs, is then calculated. An example of the indices for the comfort output variable are presented in Figure 5. In this figure, the influence of parameters 176 and 427 (as well as others) is evident. In addition to this, it is obvious once the parameters are rank ordered that only a few parameters (less than 10%) has significant influence on the output. This is important as it means that a reduced order model should capture most of the parametric influences and the full model and that noteworthy optimization results can be obtained using only a few automatically selected parameters.

To generate the reduced order meta-models, the sensitivity indices are rank ordered as in Figure 5 for both outputs, and the most influential parameters (largest sensitivity indices) are chosen based on the desired meta-model size. For this study, we investigate reduced order meta-models based on 6 to 20 parameters.

Numerical Optimization

Once the meta-models are created, the optimization is fairly straight forward. A cost function is created using a combination of comfort (PMV) and total yearly facility energy (Energy). The cost function is an uncertainty-weighted sum of normalized squares

\[ C = \alpha_1 \left( \frac{\text{PMV}}{\max(\text{PMV})} \right)^2 + \alpha_2 \left( \frac{\text{Energy}}{\max(\text{Energy})} \right)^2, \]

where \( \alpha_i \in N(1, 0.25^2) \) are chosen by simultaneous sampling the distribution 50 times using a deterministic sampling algorithm. That is, for each meta-model, 50 optimizations are performed (taking approximately 1 second each) to account for uncertainty in the meta-model. The optimized parameter values are then inserted into the original EnergyPlus model, replacing the 6-20 parameters that were chosen to be optimized. A EnergyPlus simulation was performed to calculate the expected optimized performance and best result for each optimization presented in Figure 6.

The results in Figure 6 illustrate a few important aspects of this process. First, it is obvious that optimization in the different feasible subsets provides different solutions. This is important to note because if the original feasible set was not partitioned, the optimizer may have converged in a non-desirable portion of the global feasible set. In addition to this, it is clear that there is a saturation-like quality with respect to the use of different numbers of optimization variables. In particular, the feasible subset 2B provides the best optimization results, while these results are obtained at approximately 12 parameters. Including any more optimization variables only improves the optimization results marginally.

The reason for this saturation in the optimization results can be explained by investigating where the optimized parameters lie in the feasible subset. The optimized parameter values for the feasible subset 2B are presented in Table 2. The bracketed parameters in this table indicate optimization variables that are very close to, or on the boundary. In almost all cases, the parameters that have the most influence on the output (e.g. as in Figure 5) are on the boundary. To obtain better optimization results (e.g. greater than 10% energy reduction), the boundary of the feasible set would need to be expanded beyond 25% of the nominal model.

SUMMARY

In this study we developed a method to optimize whole-building energy models that contain hundreds of candidate optimization variables. Parameters of the original energy model are sampled around the nominal design and discontinuities are systematically iden-
Table 2: Optimized parameter models for different size meta-models in quadrant 2B. The parameters in square brackets are those which are 0.02% from their imposed constraint (boundary). In this table, the following acronyms are used: Equipment Water Use (EWU), Reheat (RHC), Supply air humidity ratio (SAHR), Supply air temperature (SAT)

<table>
<thead>
<tr>
<th>Parameter Names / Number of Parameters</th>
<th>6</th>
<th>8</th>
<th>10</th>
<th>12</th>
<th>14</th>
<th>16</th>
<th>18</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density of exterior wall insulation</td>
<td>319.02</td>
<td>308.96</td>
<td>331.13</td>
<td>272.31</td>
<td>[327.81]</td>
<td>318.65</td>
<td>284.04</td>
<td></td>
</tr>
<tr>
<td>Thickness of concrete slab</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>Specific heat of ext. wall concrete layer</td>
<td>[627.87]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Velocity of air in comfort schedule</td>
<td>0.19</td>
<td>0.21</td>
<td>0.22</td>
<td>0.23</td>
<td>0.23</td>
<td>0.23</td>
<td>0.23</td>
<td>0.23</td>
</tr>
<tr>
<td>Cooling setpoint, weekdays 06:00-22:00</td>
<td>[25.56]</td>
<td>[25.23]</td>
<td>[25.56]</td>
<td>[25.56]</td>
<td>[25.48]</td>
<td>[25.56]</td>
<td>[25.56]</td>
<td>[25.56]</td>
</tr>
<tr>
<td>Heating setpoint, weekdays 24:00-06:00</td>
<td>16.11</td>
<td>[14.59]</td>
<td>[14.59]</td>
<td>[14.59]</td>
<td>[14.59]</td>
<td>[14.59]</td>
<td>[14.59]</td>
<td>[14.59]</td>
</tr>
<tr>
<td>Seasonal reset SAT</td>
<td>[15.99]</td>
<td>[15.99]</td>
<td>[15.99]</td>
<td>[15.99]</td>
<td>[15.99]</td>
<td>[15.99]</td>
<td>[15.99]</td>
<td>[15.99]</td>
</tr>
<tr>
<td>HW loop temperature (schedule value)</td>
<td>[0.04]</td>
<td>[0.03]</td>
<td>[0.04]</td>
<td>[0.05]</td>
<td>[0.05]</td>
<td>[0.05]</td>
<td>[0.05]</td>
<td>[0.05]</td>
</tr>
<tr>
<td>EWU, all non-design days 15:00-21:00</td>
<td>0.04</td>
<td>[0.03]</td>
<td>[0.04]</td>
<td>[0.05]</td>
<td>[0.05]</td>
<td>[0.05]</td>
<td>[0.05]</td>
<td>[0.05]</td>
</tr>
<tr>
<td>Plug loads, fraction radiant - Core bottom</td>
<td>[0.38]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flow per Exterior Surface Area - Top plenum</td>
<td>0.0003</td>
<td>[0.0002]</td>
<td>[0.0002]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixed water heater, max. temp. limit</td>
<td>96.43</td>
<td>97.72</td>
<td>99.83</td>
<td>96.61</td>
<td>92.37</td>
<td>[102.67]</td>
<td>[102.67]</td>
<td>[102.78]</td>
</tr>
<tr>
<td>Cooling sizing factor</td>
<td>[1.00]</td>
<td>[1.06]</td>
<td>[1.00]</td>
<td>[1.00]</td>
<td>[1.00]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cooling design SAT - Core middle</td>
<td>12.45</td>
<td>10.88</td>
<td>11.25</td>
<td>12.90</td>
<td>12.92</td>
<td>12.81</td>
<td>14.30</td>
<td>14.30</td>
</tr>
<tr>
<td>Zone cooling design SAHR - Perim. mid. Zn1</td>
<td>(0.006)</td>
<td>0.009</td>
<td>0.008</td>
<td>0.008</td>
<td>0.011</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Central cooling design SAHR, VAV2</td>
<td>[0.011]</td>
<td>[0.01]</td>
<td>[0.01]</td>
<td>[0.01]</td>
<td>[0.01]</td>
<td>[0.01]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Central heating design SAHR, VAV2</td>
<td>[0.010]</td>
<td>[0.010]</td>
<td>[0.010]</td>
<td>[0.006]</td>
<td>[0.007]</td>
<td>[0.008]</td>
<td>[0.006]</td>
<td>[0.006]</td>
</tr>
<tr>
<td>Rated outlet water temp., Perim. top Zn4 RHC</td>
<td>71.75</td>
<td>[67.83]</td>
<td>[67.54]</td>
<td>[67.54]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant min. air flow fraction - Core bottom</td>
<td>[0.37]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant min. air flow fraction - Perim. top Zn1</td>
<td>[0.38]</td>
<td>[0.23]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

tified which defines inner boundaries of feasible subsets. Reduced order meta-models are obtained for each feasible subset and optimization is performed considering an uncertain cost function.

The method was tested on a large office building modeled in EnergyPlus and approximately 10% of the yearly energy was saved while increasing occupant comfort by altering predominately operational variables (e.g. parameters for how the HVAC system is operated). Better results are likely available if the external boundaries of the feasible subsets are expanded.

REFERENCES


