ABSTRACT

From a modeling perspective, the thermal characteristics of buildings must be described with the help of parameters that often cannot be estimated with high accuracy. Results from simulations relying on erroneous parameter values can lead to inaccuracies which are hard to quantify, and small perturbations to a sensitive parameter can influence the results significantly. By examining the impact of uncertainties, it is possible to increase simulation quality, and thus inferences from the results. In this paper, the uncertainty associated with model parameters of a building using a solar thermal collector for heating and domestic hot water is analyzed. The influence of the uncertain parameters and variables on the solar fraction is quantified through the use of Monte Carlo simulations. The relationship between uncertainty analysis and sensitivity analysis is examined, as well as a practical methodology suggested to support the building design process and related decision making.

INTRODUCTION

Energy efficient buildings and advanced plant equipment commonly necessitate a more complex design process. Employing simulations in this process has the advantage that the part load behavior as well as the interactions between plant equipment and the building can be fully examined. The results of the analysis are very useful to designing and optimizing energy systems or to perform critical plant sizing. Although simulations are often used in building research and practice, uncertainties are hardly ever quantified. One reason for that could be the lack of simple tools and methodologies which are applicable to this specific problem. A mismatch between detailed and computationally expensive simulations on the one hand and crude parameter assumptions according to some rules of thumb on the other can often be found in literature and practice. This is misleading as the high accuracy of the results is only simulated creating a false sense of validity and engineering rigor. Whereas in other fields of science uncertainty analysis is widely used (Saltelli, Ratto, Andres et al. (2008), page 5-6), Lomas et al. (1992) conducted one of the first studies about Monte Carlo analysis for building simulations (Lomas, and Eppel (1992)). MacDonald conducted further research on uncertainty analysis. He analyzed in his dissertation Quantifying the Effects of Uncertainty in Building Simulation the influence of uncertain parameters in building simulations (Macdonald (2002)).

Parameters, notorious for their difficulty in estimation include building occupancy, air change rates, and domestic hot water consumption. Herkel et al. (2008) performed a literature review and analyzed user behavior in an office building. He came to the conclusion that a model of user behavior regarding window openings should depend on the season, the outdoor and indoor temperature, the time of the day and the presence of occupants (Herkel, Knapp, and Pfafferott (2008)).

Another way to handle uncertain knowledge are Bayesian networks. Dodier (1999) analyzed these belief networks in his work Unified Prediction and Diagnosis in Engineering Systems by Means of Distributed Belief Networks concerning their applicability to engineering systems (Dodier (1999)).

In this paper, a Monte Carlo technique is used to quantify the influence of uncertain parameters and variables in building simulation. The aim is to assign a probability density function (pdf) to each uncertain input of the simulation and to obtain a joint pdf for the result. In this way, it is possible to analyze the likely variation of the output given the uncertain set of inputs. It is also possible to calculate with which probability a target of a design process will be reached; an application that extends the classical building simulation analysis since the result gives not just the answer yes or no to a design question. In Figure 1 a comparison of the classical simulation approach and the analyzed approach is shown. Different to past building related studies, where sensitivity and uncertainty analysis were often analyzed separately, the link between both is examined.

SIMULATION

Analyzed building and plant equipment

The building simulated is a typical German building with a net floor area of 436 m². There is no air-
Figure 1: Classical building simulation approach versus building simulation approach with uncertainty analysis. The classical approach has single numbers as input and yields a single number (e.g. specific yearly energy consumption) with unknown accuracy which is often not acceptable given the uncertain inputs. The analyzed approach indicates how inputs might vary and quantifies the uncertainty in the result. The solid blue line in the probability density functions indicates the expected value and the dashed blue lines indicate the expected value plus/minus one standard deviation.

Table 1: Building parameters.

<table>
<thead>
<tr>
<th>parameter</th>
<th>value</th>
<th>unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A/V ) (area to volume ratio)</td>
<td>0.38</td>
<td>m(^{-1})</td>
</tr>
<tr>
<td>( U )-value (mean U-value)</td>
<td>0.53</td>
<td>W m(^{-2}) K(^{-1})</td>
</tr>
<tr>
<td>( A_{\text{win}} ) (total window area)</td>
<td>106</td>
<td>m(^{2})</td>
</tr>
</tbody>
</table>

Simulation model

Monte Carlo (MC) simulations require many simulation runs and are therefore computationally expensive. In order to reduce the computing time, it is necessary to find an appropriate simple model for thermal building simula-

tion. Furthermore, a model with many parameters is often not better than one with just a few parameters (Déqué, Ollivier, and Poblador (2000)). In this context it is imperative to determine the parameters with the greatest influence. In the case of building simulation, important parameters are the occupancy and the control parameters. For a simple zone model the simple hourly method (SHM) ac-

Figure 2: 3D-plan of the building.
According to the ISO 13790 standard is used (ISO 13790 (2007)). This zone model is based on five resistances and one capacity. The model was calibrated for this building; it showed a good agreement with the measured room temperature and the heating power of the building (Burhenne, and Jacob (2008)).

The object-oriented and equation-based modeling language Modelica is used to describe the system (Elmqvist (1997)) and the simulations are conducted using the software Dymola 6 (Dynasim AB (2004), Dynasim AB (2007)).

In actuality, the building is an office building heated by a gas boiler. For this analysis, however, it is assumed that it is a residential building with 12 occupants. A solar thermal collector with 19 m² area and a 1000 liter storage tank are modeled. The collector model was implemented in Modelica using the plug flow model description of Isaksson and Eriksson (Isaksson, and Eriksson (1993)). The collector flow rate is controlled by an on/off controller. The tank is modeled as a simple one capacitor / one resistor network and an ideal boiler is used for keeping the tank temperature at 60°C as long as the maximum power is sufficient. The radiation processor is implemented according to an equation-based model (written in the modeling language Neutral Model Format; Sahlin (1996)) from the simulation software IDA-ICE (Sahlin, Eriksson, Grozman et al. (2004)). Due to the similar structure, it is straightforward to translate other equation-based modeling languages into the Modelica language. The solar thermal system is designed for domestic hot water and heating. When the heat from this system is not sufficient, a gas boiler meets the load of the building. Figure 3 shows the graphical representation of the models.

Monte Carlo simulations

In a Monte Carlo analysis, a large number of evaluations of the model is performed with randomly sampled model inputs (Saltelli, Chan, and Scott (2000), p. 20-24). It contains the following main steps:

1. Selection of probability density functions (pdf) for each uncertain input ($X_i$).
2. Generation of a sample from each pdf.
3. Evaluation of the model for each element of the sample.
4. Result analysis.

It is assumed that the mass flow rate of the domestic hot water ($\dot{m}$) and the air change rate ($\text{ACH}$) are uncertain. Furthermore, the times ($k$) when people leave or come back to the building and how many people are present ($\text{occ}$) at a particular time cannot be determined exactly. Therefore, these times and the number of people who are in the building are sampled. Figure 4 shows the basis of the occupancy schedule with the distributions which indicate which values are sampled. The domestic hot water flow rates are generated with a program which was developed in the Solar Heating and Cooling Program of the International Energy Agency (IEA-SHC), Task 26: Solar Combisystems (Jordan, and Vajen (2003)). The sampling was done by multiplying a sampled factor with the mass flow value, which is generated with the step size of 60 s. The infiltration air change rates are implemented according to a schedule (Figure 5) and the value is multiplied with a sampled factor ($\text{ACH}$).

![Figure 4: Schedule for the occupancy of the building. The distributions indicate which values are sampled.](image)

![Figure 5: Schedule for the infiltration rate (air change rate) in the building.](image)

The sampling (step 2) generates the input matrix (Equation 1). The parameters for the distributions which were used for the sampling can be found in Table 2.
In this paper the solar fraction (SolFrac) is used to analyze the design of the solar thermal system. Its definition is

\[
\text{SolFrac} = \frac{Q_{\text{collector}}}{Q_{\text{total}}}.
\] (2)

Once the model is evaluated for each sample set (step 3), the result vector is obtained (Equation 3).

\[
Y_{\text{Output}} = \begin{bmatrix}
\text{SolFrac}^{(1)} \\
\vdots \\
\text{SolFrac}^{(n-1)} \\
\text{SolFrac}^{(n)}
\end{bmatrix}
\] (3)

Other results which could be analyzed include the yearly end or primary energy consumption of the gas boiler, life cycle costs for the plant system, and CO₂-emissions.

A crucial point in applying Monte Carlo techniques is the sample size. Macdonald analyzed this problem with respect to building simulation and stated that simple random sampling with a sample size of 100 should be used (Macdonald (2009)). However, in this paper a sample size of 1000 is used to generate the samples and it is done by a simple random sampling algorithm. The language and environment R (R Development Core Team (2009)) for statistical computing is used to draw a sample, change the
Table 2: Distribution parameters.

<table>
<thead>
<tr>
<th>parameter</th>
<th>distribution</th>
<th>µ</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>m (scaling factor)</td>
<td>normal</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>ACH (scaling factor)</td>
<td>normal</td>
<td>1</td>
<td>0.2</td>
</tr>
<tr>
<td>k₁ (time occupancy)</td>
<td>normal</td>
<td>7</td>
<td>0.5</td>
</tr>
<tr>
<td>k₂ (time occupancy)</td>
<td>normal</td>
<td>8</td>
<td>0.5</td>
</tr>
<tr>
<td>k₃ (time occupancy)</td>
<td>normal</td>
<td>12</td>
<td>0.5</td>
</tr>
<tr>
<td>k₄ (time occupancy)</td>
<td>normal</td>
<td>13</td>
<td>0.5</td>
</tr>
<tr>
<td>k₅ (time occupancy)</td>
<td>normal</td>
<td>17</td>
<td>0.5</td>
</tr>
<tr>
<td>k₆ (time occupancy)</td>
<td>normal</td>
<td>18</td>
<td>0.5</td>
</tr>
<tr>
<td>occ₁ (number of occupants)</td>
<td>normal</td>
<td>12</td>
<td>0.5</td>
</tr>
<tr>
<td>occ₂ (number of occupants)</td>
<td>normal</td>
<td>8</td>
<td>0.5</td>
</tr>
<tr>
<td>occ₃ (number of occupants)</td>
<td>normal</td>
<td>4</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Simulation input file, call the simulation program, analyze, and visualize the result.

**DISCUSSION AND RESULT ANALYSIS**

**First set of Monte Carlo simulations**

For each simulation, one solar fraction is obtained. Figure 6 shows the histogram for the result vector.

This gives a first indication how the result varies given the uncertainties in the inputs. A practical design question might be:

- **What is the probability to reach a solar fraction of ≥ 20%?**

After normalizing the histogram, one obtains the probability density function of the result (Figure 7).

The pdf can be described with a normal distribution:

\[
p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{x - \mu}{\sigma}\right)^2\right).
\]  

(4)

The parameters for the normal distribution can be found in Table 3.

<table>
<thead>
<tr>
<th>result</th>
<th>distribution</th>
<th>µ</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>SolFrac</td>
<td>normal</td>
<td>0.1847</td>
<td>0.01309</td>
</tr>
</tbody>
</table>

With this function it is possible to calculate the answer to the defined design question with

\[
P(SolFrac \geq 20\%) = \int_{0.2}^{\infty} p(SolFrac) \, dSolFrac \\
\approx 0.121 \approx 12.1\%.
\]

(5)

With such an answer available, decision makers can decide whether the design of the plant equipment should be changed or not. However, with the current design it is not very probable to achieve the design target of 20% or greater solar fraction.

**Second set of Monte Carlo simulations**

The tank size in the simulated example was 1000 L and the collector had an area of 19 m². Now we assume that the decision maker (e.g. client) formulates a new design question. The question is:

- **What is the probability to reach a solar fraction of ≥ 20% given a tank size of 2000 L and a collector area of 25 m²?**
After running the Monte Carlo simulations with the same sample set but with different parameters for the tank size and the collector area, a new PDF can be computed (Figure 8). Using the same sample set is important to make sure that the result does not change because of the sampled input. This, however, will have no significant influence when a large sample size is used.

![Figure 8: Probability density function of the result for the second set of Monte Carlo simulations. The blue area under the distribution represents the probability that the solar fraction is above 20%.
](image)

The parameters for the normal distribution of the second Monte Carlo simulation can be found in Table 4.

**Table 4: Distribution parameters for the result of the second Monte Carlo simulation.**

<table>
<thead>
<tr>
<th>result</th>
<th>distribution</th>
<th>$\mu$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SolFrac</td>
<td>normal</td>
<td>0.2253</td>
<td>0.01564</td>
</tr>
</tbody>
</table>

The answer of the newly defined design question is now:

$$P(SolFrac \geq 20\%) = \int_{0.2}^{+\infty} p(SolFrac) \, dSolFrac \quad (6)$$

$$\approx 0.947 \approx 94.7\%.$$  

Now the decision maker has been properly informed about the probability of achieving the design target. Together with cost data associated with the plant equipment he would even know how much he has to pay for the increased probability of solar fractions in excess of 20%.

**Sensitivity analysis**

Since a large number of simulation runs were conducted for the Monte Carlo analyses, it is possible to further exploit this available data. One way is to analyze the sensitivity of the sampled input parameters, giving the modeler information about the sensitivity of the different parameters and in which manner they influence the result. An interesting and simple method is a graphical sensitivity analysis. The analysis is conducted for the first design of the system (19 m² collector area and 1000 L tank volume). Each sample and its corresponding result (e.g. $ACH_1, Q_1; \ldots ; ACH_n, Q_n$) is plotted in a scatter plot. Figure 9 shows a matrix of scatter plots for the sampled inputs as well as for the solar fraction, the total energy demand of the building and the total collector energy over one year.

It can be seen that the total heat consumption of the building ($Q_{total}$) and the solar fraction ($SolFrac$) are very sensitive to the air change rate ($ACH$). Furthermore, the linear dependency between the variable and the results is visible.

The collector energy ($Q_{coll}$) is sensitive to the flow rate of the domestic hot water ($\dot{m}$). This dependency is not linear since at some point the area of the collector is not sufficient to meet the higher demand.

The other pairs do not show such a strong dependency. Nonetheless, it would be worthwhile to apply another sensitivity analysis method to further examine the dependencies in the model.

**CONCLUSION**

In this paper, a Monte Carlo based approach to quantify the influence of uncertain parameters and variables was discussed. The method is applicable to answer questions during the design process. It was demonstrated that statistical methods can extend the use of classical building simulations. Two designs were compared. The probability that the first design meets the defined design target was approximately 12% whereas the second design is more likely (≈ 95%) to meet the requirements. With a classical building simulation the answer would have been either yes or no depending on the assumptions. It appears to be desirable to account for the world’s uncertainty even – or better especially – in building simulation.

Furthermore, a way of presenting a sensitivity analysis was shown, which is very useful to determine the sensitive parameters. It offers insight into the influence of the input as well as to the model behavior under changing parameters or variables.

**Future work**

In this paper, the uncertain input variables were considered independent of each other. This is easy to implement and to handle within the complex simulation environment. Nonetheless, variables such as the occupancy, the air change rate and the domestic hot water flow rate are unfortunately not independent. Assuming independence for these variables leads to an overestimation of the influence brought upon by the uncertain input. Future work will address the problem of dealing with dependent in-
put variables and will include user models such as the one presented in Herkel, Knapp, and Pfafferott (2008) in the uncertainty analysis procedure.

Furthermore, the assumptions of the distributions from which the sample was drawn need to be verified. Therefore measured data from several buildings and building configurations should be analyzed to obtain reasonable assumptions for the prior distributions.

The analysis in this example was conducted on the basis of cumulative annual energy consumption. However, it would be interesting to develop a method which improves and automates the visualization of temporally resolved results (e.g. time series plot with uncertainty band).

ACKNOWLEDGMENT

This study was funded by the Reiner Lemoine Stiftung and the German Federal Ministry of Economics and Technology (BMWi) under the program “EnOB” (BMWi 0327410A). Special thanks to Robert H. Dodier for insightful discussions on statistical problems.

REFERENCES


Dymosim AB. 2004. Dymola Multi-Engineering Model-


