OPTIMIZATION-BASED SIMULATION MODEL CALIBRATION USING SENSITIVITY ANALYSIS

Farhang Tahmasebi, Ardeshir Mahdavi

Department of Building Physics & Building Ecology, Vienna University of Technology, Vienna, Austria
e-mail: farhang.tahmasebi@tuwien.ac.at

Abstract

This paper examines the potential of sensitivity analysis-assisted optimization-based simulation model calibration. Toward this end, a university campus office area was selected and equipped with a monitoring infrastructure. Various streams of data were collected, including indoor climate, outdoor weather conditions, energy delivery via the building's heating system, and energy use for lighting and equipment. One of the primary objectives of this comprehensive monitoring campaign was to use monitored data to both populate the initial thermal simulation model of the office area and to maintain its fidelity through a calibration process. The present contribution also addresses a specific problem faced by an optimization-based simulation calibration approach: In many realistic circumstances, a large number of model input variables could be subjected to the optimization process. This large number of candidate input variables can be reduced to a certain extent via heuristically-based considerations pertaining, for example, to the knowledge domain captured in building physics. We argue, however, that this process could be further rationalized, if we make use of sensitivity analysis to identify a subset of the input variables most likely to influence the simulation results. Distinguishing this subset from the entire set of input variables will reduce the computational cost of the subsequent calibration process.

Keywords: Building performance simulation, calibration, optimization, sensitivity analysis

Abstrakt

Článek se zabývá potenciálem využití citlivostní analýzy při kalibraci simulačního modelu. V tomto případě byla zvolena administrativní část univerzitního kampusu, která byla vybavena systémem měření. Tento systém sbírá následující data - údaje o vnitřním prostředí budovy, klimatické podmínky, údaje o energii potřebné na vytápění, o spotřebě energie na osvětlení a pro provoz zařízení. Jedním z hlavních cílů tohoto rozsáhlého monitorovacího projektu bylo využití monitorovaných dat pro počáteční tepelný simulační model administrativní části univerzitního kampusu a udržení jeho věrohodnosti v průběhu kalibračního procesu. Příspěvek se zabývá specifickým problémem kalibračního přístupu při optimalizačních simulacích. V případě příznivých okolností, velký počet vstupních proměnných, by mohl být součástí optimalizačního procesu, avšak tohle množství vstupních hodnot může být zredukováno pomocí heuristicky podložených předpokladů vztahujících se například ke znalostem problematiky z pohledu stavební fyziky. My tvrdíme, že i přesto by mohl být tento proces ještě více racionalizován, kdybychom využili citlivostní analýzu pro identifikování podskupiny těch vstupních dat, které nejvíce ovlivňují výsledky simulace. Odlišením této podskupiny od ostatních vstupních hodnot bude snížen výsledný výpočetní čas následného kalibračního procesu.

Klíčová slova: simulace efektivity budov, kalibrace, optimalizace, citlivostní analýza

INTRODUCTION

Research on performance simulation deployment opportunities in the building operation phase has recently gained on momentum. Specifically, simulation routines have been successfully applied in the conception and implementation of predictive methods for building systems control [1]. Needless to say, the quality of such a predictive control system depends on the reliability of the integrated simulation models.

Thus, to ensure that predictions are dependable, applied simulation models must be calibrated. Moreover, given the dynamic nature of building operation, some input parameters of the model may have to be subjected to calibration more frequently. This circumstance implies that the calibration task cannot be approached as an ad hoc or one-time activity. Rather, it needs to be conducted on a systematic basis. Consequently, the entire calibration process should be preferably automated to ensure efficiency and consistency. Given this background, the present contribution explores the potential of an optimization-based approach to simulation model calibration that is intended to be employed in a building's monitoring and systems control environment.

The present contribution also addresses a specific problem faced by an optimization-based simulation calibration approach: In many realistic circumstances, a large number of model input variables could be subjected to the optimization process. This large number of candidate input variables can be reduced to a certain extent via heuristically-based considerations pertaining, for example, to the knowledge domain captured in building physics. We argue, however, that this process could be further rationalized, if we make use of sensitivity analysis to identify a subset of the input variables most likely to influence the simulation results. Distinguishing this subset from the entire set of input variables will reduce the computational cost of the subsequent calibration process.

METHODOLOGY

The monitored building

To explore the potential of optimization-based calibration in a realistic setting, we selected an actual office in a building of the Vienna University of Technology, which is equipped with a monitoring infrastructure. The monitoring infrastructure provides various streams of data, including indoor climate, weather conditions, energy delivery via the heating system, energy use for lighting and equipment, occupancy presence, and the opening state of windows and doors. Data are regularly collected with a variable frequency depending on the magnitude of changes in successive recordings.

The monitored data was used to: i) create a weather file based on local data instead of using a predefined "typical" year; ii) populate the initial building model with dynamic data regarding internal loads, device states, and occupancy processes; iii) calibrate the initial model (see Table 1).

Tab. 1 - Use of monitored data in the calibration process

Use of data	Data point	Unit
Creating local weather data file	Global horizontal radiation	W/m^2
	Diffuse horizontal radiation	W/m^2
	Outdoor dry bulb temperature	^{o}C
	Outdoor air relative humidity	%
	Wind Speed	m/s
	Wind direction	degree
	Atmospheric pressure	Pa
Creating the initial model	Electrical plug loads	W
	State of openings (open/closed)	-
	State of the lights (on/off)	-
	Occupancy (presence/absence)	-
	Radiators' surface temperature	^{o}C
Calibration	Indoor air dry bulb temperature	^{o}C

The building model

The building was modeled in the building energy simulation tool EnergyPlus v7.0 [2]. It was assumed that the floor and ceiling surfaces of the office are adiabatic, as the office is situated between two occupied floors. In the zoning scheme, the open-plan south and north-oriented spaces were separated from the central corridor. However, using the nework-based multi zone airflow model of EnergyPlus [3], the airflow between these connected spaces was simulated. Figure 1 illustrates the building floor plan and the thermal zoning of the building model.

The monitored data was incorporated as simulation input information in terms of scheduled variables. Since writing schedules manually in EnergyPlus (and probably in any other simulation program with text-based input) is a time-consuming and error-prone process, a simple program was written in Matlab [4] to generate an event-based "compact schedule" for each data point.

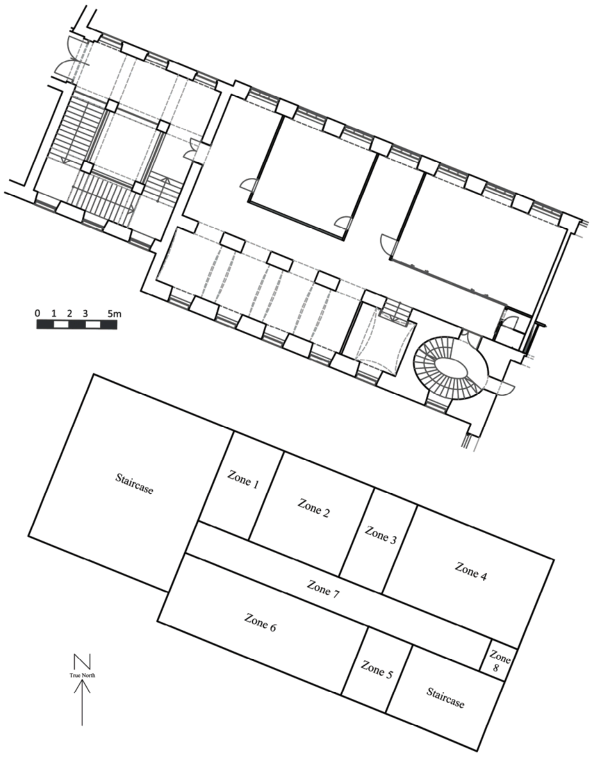

Fig. 1 - Building floor plan and thermal zoning of the model

The heating system model

To simulate the building's performance during the heating season, the heat delivery rate of the hydronic heating system had to be calculated and provided to the simulation model as input information. Toward this end, measured radiator surface temperatures were used. The heat emission rate of the radiators was obtained using the following equations:

$$q = q_R + q_C \tag{1}$$

$$q_R = \varepsilon \cdot \sigma \cdot A_R \cdot (T_S^4 - T_R^4) \tag{2}$$

$$q_C = h_C \cdot A_C \cdot (\theta_S - \theta_R) \tag{3}$$

$$h_C = 2 \cdot \left|\theta_S - \theta_R\right|^{0.25} + 4\varepsilon \cdot \sigma \cdot \left(\frac{T_S + T_R}{2}\right)^3 \tag{4}$$

Where:

q	heat delivery rate of radiators [W];
q_R	radiative component of heat delivery [W];
q_C	convective component of heat delivery [W];
ε	emissivity of the radiator [-];
σ	constant (5.67×10^{-8} W.m^{-2}.K^{-4});
A_R	effective radiator area for radiation [m^2];
T_S	surface temperature of radiators [K];
T_R	room temperature [K];
h_C	convective heat transfer coefficient [W.m^{-2}.k^{-1}];
A_C	effective radiator area for convection [m^2];
θ_S	surface temperature of radiator [^{o}C];
θ_R	room temperature [^{o}C].

Run periods

The model calibration and validation process involved a monitoring period of nearly three months consisting of two 44-day periods (Table 2). The sensitivity analysis was also performed in the calibration period.

Tab. 2 - Specification of run periods

Run periods	Start date	End date
Calibration	15.02.2011	30.03.2011
Validation	27.04.2011	09.06.2011

Optimization-based calibration approach

In an optimization-based approach, calibration is cast as an error-minimizing process. In this kind of optimization problem, the cost function addresses the difference between measured and simulated data (in the present case, indoor air temperature). The variables in the optimization algorithm include a number of model input parameters. The attributes of these variables will be varied toward minimizing the cost function.

To accomplish the optimization in a way that works smoothly with the simulation model, we used Genopt, which is a generic optimization program. Genopt has been developed to conveniently find the attribute range of relevant independent variables that would yield optimal system performance. Genopt optimizes a user-supplied cost function, using a user-selected optimization algorithm [5].

The algorithm used for the optimization was the hybrid generalized pattern search and particle swarm optimization algorithm. This is one of the recommended generic algorithms for problems, where the cost function cannot be simply and explicitly stated, but can be approximated numerically by a thermal building simulation program [5].

Selecting calibration variables via sensitivity analysis

The problem of large search space and multiple possible solutions has been addressed in previous research (see, for example [6,7]). In the present research, to identify a subset of the input variables most likely to influence the simulation results, first, the large number of candidate model parameters was reduced to a certain extent via heuristically-based considerations. This subset included 23 model input variables (Table 2). Secondly, these variables were subjected to a Monte Carlo-based sensitivity analysis.

The performed sensitivity analysis included four steps. In the first step, assuming uniform distribution of input varibles, a range was selected for each variable (Table 2).

In the second step, a sample of points was generated from the distribution of the inputs using the Latin hypercube sampling method, which is a particular case of stratified sampling [8]. The result was a sequence of 690 sample elements.

In the third step, the model was fed with the sample elements and a set of model outputs was produced. Since the sensitivity analysis was planned to be performed in the heating period, the building's total heat load during the run period was designated as the output.

Running 690 different models with randomly selected input parameters' values, a mapping was created from the space of the inputs to the space of the results that were used in the fourth step as the basis for sensitivity analysis. By solving a multiple linear regression model, using least squares [8], the absolute value of Standard Regression Coefficient (SRC) was calculated for the variables as a quantitative sensitivity measure. Table 4 shows the analysed variables in order of the absolute value of SRC.

Based on these results, the first four variables, which have SRC values higher than 0.1, were chosen to be subjected to optimization-based calibration in the next stage. These variables, their initial values and their allowed calibration ranges can be seen in table 5.

Calibration cost function

For the purpose of building performance analysis, error can be defined as the difference between a predicted value and a measured value [9]. In the present case, the error was calculated for the indoor air temperature averaged over all office zones. To minimize this error, and to maintain the "goodness of fit" of the model at the same time, a weighted function of two different indicators was defined as the cost function. The first indicator is the "Coefficient of Variation of the Root Mean Squared Deviations" (Equations 5 & 6) which serves to aggregate the individual time step errors into a single dimensionless number.

Tab. 3 – Variables subjected to SA and their ranges

Variables	Min. Value	Max. Value
White painted gypsum - Thermal conductivity	0.336	0.504
White painted gypsum - Density	960	1440
White painted gypsum - Thermal absorptance	0.82	0.93
White painted gypsum - Solar absorptance	0.24	0.36
White painted stucco - Thermal conductivity	0.576	0.864
White painted stucco - Density	1485	2227
White painted Stucco - Thermal absorptance	0.82	0.93
White painted Stucco - Solar absorptance	0.24	0.36
External walls brick layer - Thermal conductivity	0.56	0.84
External walls brick layer - Density	1360	2040
Wood parquet - Thermal absorptance	0.664	0.996
Wood parquet - Solar absorptance	0.48	0.72
Glazing - Solar transmittance	0.56	0.84
Glazing - Front side infrared emissivity	0.837	0.898
Glazing - Back side infrared emissivity	0.837	0.898
Glazing – Thermal conductivity	0.72	1.08
Windows frame - Thermal conductance	1.816	2.724
Outside windows discharge coeff. when open	0.64	0.96
Inside windows discharge coeff. when open	0.64	0.96
Outside closed openings air mass flow coeff.	0.00011	0.00017
Outside closed openings air mass flow exponent	0.52	0.78
Inside closed openings air mass flow coeff.	0.016	0.024
Inside closed openings air mass flow exponent	0.56	0.84

Tab. 4 – Variables in order of absolute value of SRC

Variables	| SRC |
External walls brick layer - Thermal conductivity	0.7735
Outside windows discharge coefficient when open	0.4128
Glazing - Solar Transmittance at Normal Incidence	0.3660
Outside openings air mass flow coeff. when closed	0.1132
Glazing - Front Side Infrared Emissivity	0.0831
Inside openings air mass flow coeff. when closed	0.0760
Inside openings air mass flow exponent when closed	0.0663
Glazing - Back Side Infrared Emissivity	0.0626
White-painted Stucco - Solar absorptance	0.0592
Glazing - Thermal conductivity	0.0374
White painted gypsum - Thermal conductivity	0.0369
White painted Stucco - Thermal absorptance	0.0314
Brick - Density	0.0314
Windows frame - Thermal conductance	0.0285
White painted stucco - Thermal conductivity	0.0218
Outside openings air mass flow exponent when closed	0.0152
White painted gypsum - Thermal absorptance	0.0145
Inside windows discharge coefficient when open	0.0104
Wood parquet - Solar absorptance	0.0090
White painted gypsum - Solar absorptance	0.0058
Wood parquet - Thermal absorptance	0.0038
White painted gypsum - Density	0.0015
White painted stucco - Density	0.0010

Tab. 5 - The variables in the first calibration

Variables	Unit	Initial value	Lower band	Upper band
External walls brick layer thermal conductivity	$W.m^{-1}.K^{-1}$	0.70	0.56	0.84
Outside windows discharge coefficient when open	-	0.80	0.00	1.0
Glazing solar transmittance at Normal Incidence	-	0. 837	0.56	0.85
Outside openings air mass flow coeff. when closed	$kg.s^{-1}.m^{-1}$	1.4×10^{-4}	1.4×10^{-5}	0.003

$$RMSD = \sqrt{\frac{\sum_{i=1}^{n}(m_i - s_i)^2}{n}} \tag{5}$$

$$CV(RMSD) = \frac{RMSD}{\overline{m}} \cdot 100 \tag{6}$$

The other indicator used in the cost function is the "coefficient of determination" denoted by R^2. This indicator has been deployed because the main purpose of the developed model is the prediction of future outcomes and R^2 provides a measure of how well future outcomes are likely to be predicted by the model. In other words, R^2 is a statistic that will give some information about the goodness of fit of a model. The coefficient of determination ranges from 0 to 1. An R^2 of 1.0 indicates that the regression line perfectly fits the data. Therefore, it is preferable to maximize the R^2 value in the optimization process. While there are different definitions of R^2, here it has been calculated via Equation 7:

$$R^2 = \left(\frac{n\sum m_i s_i - \sum m_i \sum s_i}{\sqrt{\left(n\sum m_i^2 - (\sum m_i)^2\right)\cdot\left(n\sum s_i^2 - (\sum s_i)^2\right)}} \right)^2 \tag{7}$$

In Equations 5 to 7, m_i is the measured air temperature (averaged over all office zones) at each time step, s_i is simulated air temperature at each time step, n is the total number of time steps, and $\overline{m}$ is the mean of the measured values. The defined cost function f takes into account the CV(RMSD) and R^2 in an equally weighted manner (Equation 8).

$$f_i = 0.5 \cdot CV(RMSD)_i + 0.5 \cdot (1 - R_i^2) \cdot \frac{CV(RMSD)_{ini}}{(1 - R_{ini}^2)} \tag{8}$$

In Equation 8, $CV(RMSD)_i$ is the coefficient of variation of the RMSD at each optimization iteration, R_i^2 is the coefficient of determination at each optimization iteration, $CV(RMSD)_{ini}$ is the coefficient of variation of the RMSD of the initial model, and R_{ini}^2 is the coefficient of determination of the initial model.

To efficiently manage the repetitive process of varying the input variables' attributes, the calculation of the cost function was tightly integrated with the simulation application. To accomplish this, the monitored indoor air temperatures were incorporated into the input stream and the EnergyPlus runtime language [10] was used to calculate the cost function by the EnergyPlus engine after each run of the model.

RESULTS

The optimized values of the model input variables are given in Table 6. Table 7 presents the values of the indicators used in the weighted cost function, for the initial and calibrated models. Note that these results are based on the comparison of measured and simulated indoor temperatures as aggregated over all office zones.

DISCUSSION

As it can be seen from Table 7, the initial model generated outputs with relatively low R^2 values in both the calibration and validation periods. The automated calibration, however, could effectively increase the R^2 value and reduce the error in terms of CV(RMSD). Thus, the present study points to the promising potential of monitoring-based optimization-assisted simulation model calibration using sensitivity analysis.

Tab. 6 - The optimized values of the model variables.

Variables	Unit	Optimized value
External walls brick layer thermal conductivity	$W.m^{-1}.K^{-1}$	0.561
Outside windows discharge coefficient when open	-	0.284
Glazing solar transmittance at Normal Incidence	-	0.850
Outside openings air mass flow coeff. when closed	$kg.s^{-1}.m^{-1}$	4.15×10^{-4}

Tab. 7 - R^2 and RMSD of the initial & calibrated model.

Period	Initial Model		Calibrated Model	
	R^2	CV(RMSD)	R^2	CV(RMSD)
Calibration	0.35	4.21	0.85	3.34
Validation	0.69	8.07	0.87	2.68

The performance of the approach could be further improved via a more detailed process for the determination of the cost function and associated weights. Note that the convergence-based approach to the definition of the values of model input parameter in the course of the optimization process does not mean that "true values" for such parameter are found. Rather, optimization exploits the uncertainty potential in our knowledge of the exact values of such parameter to provide a better fit to the monitoring results. It is thus important, that care is taken while defining the permissible variations from the initial values of model input parameter.

CONCLUSION

We demonstrated the utility of sensitivity analysis in assisting the optimization-based calibration of the thermal performance model of an office building. Data obtained via the monitoring system was deployed to both populate the initial simulation model and to maintain its fidelity through a systematic optimization-based calibration process.

To perform the optimization-based calibration, a set of heuristically selected model input variables was subjected to sensitivity analysis in order to identify a subset of the input variables most likely to influence the simulation results. Distinguishing this subset from the entire set of input variables reduced the computational cost of the subsequent calibration process. Moreover, a cost function was proposed, which equally weighted an error and a goodness of fit indicator.

The results displayed a noticeable improvement of the predictive potency of the calibrated model even though only four input variables of the model were subjected to calibration. Hence, the sensitivity analysis-assisted optimization-based calibration represents a promising opportunity for performance enhancement in applications pertaining to building automation, diagnostics, facility management, and model-based systems control.

ACKNOWLEDGMENT

The research presented in this paper is supported by funds from the project "Control & Automation Management of Buildings & Public Spaces in the 21st Century" (CAMPUS21, Project Reference: 285729) as well as from the Austrian Science Foundation (FWF) I545-N22 project. The authors also gratefully acknowledge the contributions of Robert Zach and Matthias Schuß toward the automation of multiple models generation and simulation, which was used for the purposes of the present study.

REFERENCES

[1] Mahdavi, A. 2001. Simulation-based control of building systems operation. Building and Environment 36, 789–796.

[2] EnergyPlus 2012. http://apps1.eere.energy.gov/buildings/energyplus/.

[3] Gu, Lixing. 2007. Airflow Network Modelling in EnergyPlus, 10th International Building Performance Simulation Association Conference, Beijing, China.

[4] Matlab 2012. http://www.mathworks.com/products/matlab/.

[5] LBNL 2011. Genopt user manual, Lawrence Berkeley National Laboratory, Available at: http://SimulationResearch.lbl.gov.

[6] Coffey, B. 2008. A development and testing framework for simulation-based supervisory control with application to optimal zone temperature ramping demand response using a modified genetic algorithm, M.Sc. thesis, Concordia University.

[7] Reddy, T., Maor, I., Panjapornpon, C. 2007. Calibrating detailed building energy simulation programs with measured data-Part 1: general methodology, HVAC&R Research 13, 221-241.

[8] Saltelli, A., Ratto, M., Andres, T. (2011). Global Sensitivity Analysis - The Primer. John Wiley & Sons Inc., ISBN: 978-0-470-05997-5.

[9] Polly B., Kruis N., Roberts D. 2011. Assessing and improving the accuracy of energy analysis for residential buildings, U.S. National Renewable Energy Laboratory's (NREL), Sponsoring organization: U.S. Department of Energy, Report number: DOE/GO-102011-3243.

[10] DOE 2011. EnergyPlus application guide for EMS (a.k.a. The book of EnergyPlus runtime language). US Department of Energy.