SIMULATION MODEL CALIBRATION: AN OPTIMIZATION-BASED APPROACH

Farhang Tahmasebi, Robert Zach, Matthias Schuß, and Ardeshir Mahdavi
Department of Building Physics & Building Ecology,
Vienna University of Technology, Vienna, Austria

ABSTRACT

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. 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, the incorporated simulation models must be calibrated. Moreover, given the dynamic nature of building operation and the boundary conditions, some input parameters of the model may have to be subjected to calibration more frequently. Hence, the calibration task cannot be approached as a one-time activity. Rather, it needs to be conducted on a regular and systematic basis. Given this background, this paper examines the potential of recurrent optimization-based simulation model calibration of an office area. The results displayed a noticeable but not fully consistent improvement of the predictive potency of the calibrated model.

INTRODUCTION

Numeric performance simulation tools are conventionally used to predict the future performance of building designs (Hensen et al. 2011). More recently, the potential of simulation routines is being explored in the buildings' operation phase. Specifically, predictive systems operation approach has shown how simulation engines can be incorporated as an integral component of a building's control system (Mahdavi 2001). Thereby, to arrive at a preferable control option, implications of alternative control actions for the control task's objective function are evaluated proactively via parametric simulation.

Needless to say, the quality of such a system greatly depends on the reliability of the deployed simulation model. 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 a recurrent optimization-based approach to simulation model calibration that is intended to be employed in a building's monitoring and systems control environment.

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).

The building model

The building was modeled in the building energy simulation program EnergyPlus v7.0 (EnergyPlus 2012). 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 (Gu 2007), 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 (Matlab 2012) to generate an event-based "compact schedule" for each data point.

Table 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	°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	°C
Calibration	Indoor air dry bulb temperature	°C

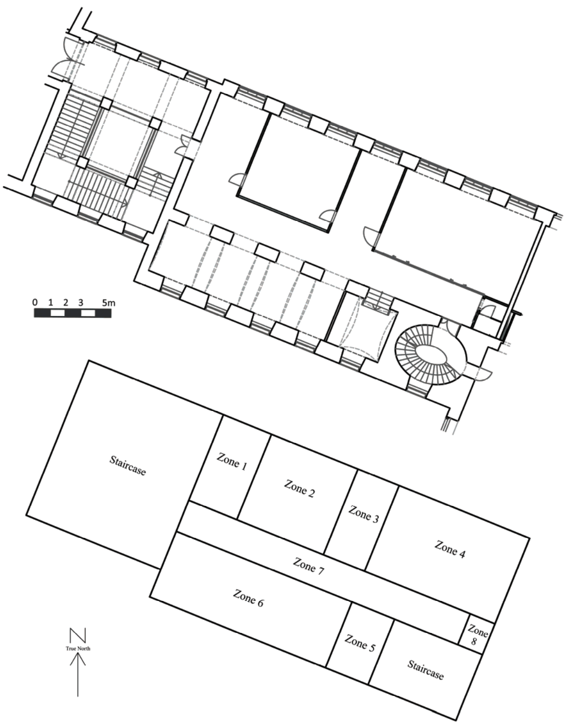

Figure 1
Building floor plan and thermal zoning of the model

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 [°C];
θ_R room temperature [°C].

Run periods

The model calibration and validation process involved a monitoring period of nearly six months consisting of four 44-day periods. Table 2 summarizes data on the run periods and Figure 2 shows the recurrent calibration process of the building performance model in these periods.

Table 2
Specification of run periods

Run periods	Start date	End date
1	15.02.2011	30.03.2011
2	27.04.2011	09.06.2011
3	10.06.2011	23.07.2011
4	24.07.2011	05.09.2011

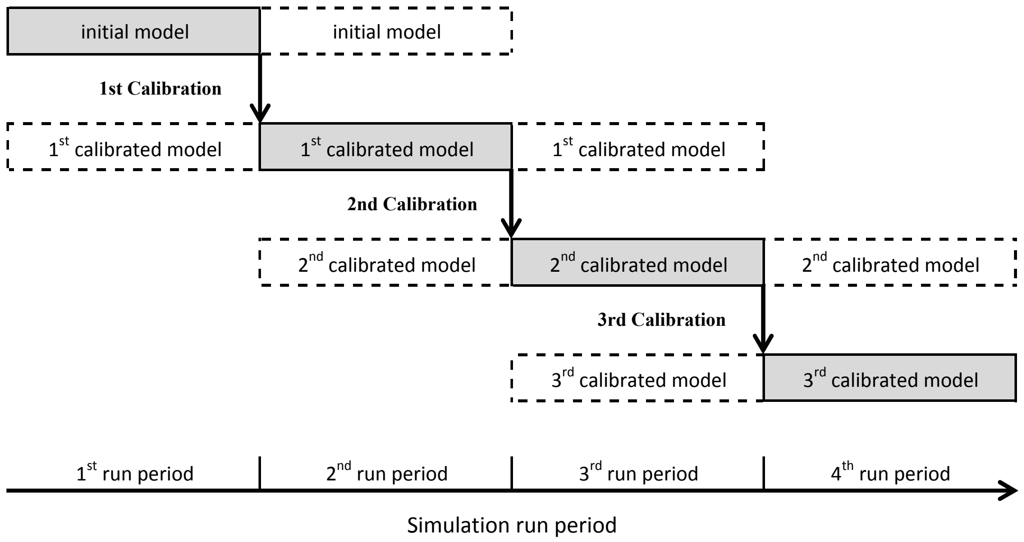

Figure 2 The process of recurrent calibration

As illustrated in Figure 2, first, the initial model was calibrated based on the monitored data from the first period (1st calibration). The resulting calibrated model was evaluated in the second period by comparing its predictions with the monitored data as well as the predictions of the initial model.

Subsequently, this calibrated model was re-calibrated twice (2nd and 3rd calibrations) based on monitored data from the second and third periods respectively. Finally the predictive performance of the 3rd calibration was evaluated using monitored data from the fourth period.

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 (LBNL 2011).

Algorithm used for the optimization was the hybrid generalized pattern search algorithm with 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 (LBNL 2011).

Calibration variables

The problem of large search space and multiple possible solutions has been addressed in previous research (see, for example, Coffey 2008). Methods such as sensitivity analysis have been proposed to limit the number of variables in the optimization process (Reddy et al. 2007). In the present research, the authors selected the pertinent variables for calibration based on their previous experiences. In the first calibration, five input variables were selected (see Table 3), which address the heat transfer processes in the building, namely conduction, convection (air infiltration and ventilation), and solar radiation. In case of external walls of the building, only the thermal properties of this component's dominant layer (brick) were subjected to variance and a variation range of 20% was applied to these parameters. Note that all variables of Table 3 may be seen as independent variables with the exception of the last two variables, namely the thermal conductivity and the density of the brick layer. To prevent the optimization process to arrive at physically unrealistic combinations of these two variables, a simplified relationship between them was postulated, which is derived from information in the relevant literature (Gösele et al. 1996):

$$\lambda_B = 0.0005 \cdot \rho_B - 0.12 \quad (5)$$

In the above equation, λ_B is the thermal conductivity of brick in [W.m^{-1}.K] and ρ_B is the density of brick in [kg.m^{-3}].

For the "Air Mass Flow Coefficient" through cracks, a wider range was selected based on the template libraries introduced for EnergyPlus (DesignBuilder 2008). For external windows, the entire possible range of the "open factor" was allowed. The variables of the first optimization-based calibration and their variation ranges are given in Table 3.

In the subsequent model calibration runs, only the first two input parameters presented in table 3 were subjected to calibration. The obtained variable values at each calibration were used as the initial values of the variables in the following calibration.

Table 3
The variables in the first calibration

Variables	**Unit**	**Initial value**	**Lower band**	**Upper band**
Closed windows: Air mass flow coefficient	$kg.s^{-1}.m^{-1}$	1.4×10^{-4}	1.4×10^{-5}	0.003
External windows: Open factor	-	1.0	0.0	1.0
External windows: Solar transmittance	-	0.837	0.670	1.000
External walls (brick): Thermal conductivity	$W.m^{-1}.K^{-1}$	0.73	0.56	0.90
External walls (brick): Density	$kg.m^{-3}$	1700	1360	2040

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 (Polly et al. 2011). 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 6 & 7). CV(RMSD) serves to aggregate the individual (time interval-specific) errors into a single dimensionless number.

$$RMSD = \sqrt{\frac{\sum_{i=1}^{n}(m_i - s_i)^2}{n}} \qquad (6)$$

$$CV(RMSD) = \frac{RMSD}{\overline{m}} \cdot 100 \qquad (7)$$

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 8:

$$R^2 = \left(\frac{n\sum m_i s_i - \sum m_i \sum s_i}{\sqrt{\left(n\sum m_i^2 - (\sum m_i)^2\right)\cdot\left(n\sum s_i^2 - (\sum s_i)^2\right)}} \right)^2 \qquad (8)$$

In Equations 6 to 8, 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 9).

$$f_i = 0.5 \cdot CV(RMSD)_i + 0.5 \cdot (1 - R_i^2) \cdot \frac{CV(RMSD)_{ini}}{(1 - R_{ini}^2)} \qquad (9)$$

In Equation 9, $CV(RMSD)_i$ 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 (DOE 2011) 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 in the recurrent calibrations are given in Table 4. Table 5 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.

Table 4 The variables' values in initial and calibrated models

Variables	**Unit**	**Initial model**	**1st calibrated model**	**2nd calibrated model**	**3rd calibrated model**
Closed windows: Air mass flow coefficient	$kg.s^{-1}.m^{-1}$	1.4×10^{-4}	4.0×10^{-5}	1.5×10^{-5}	2.75×10^{-5}
External windows: Open factor	-	1.0	0.36	0.18	0.43
External windows: Solar transmittance	-	0.837	0.717	0.717	0.717
External walls (brick): Thermal conductivity	$W.m^{-1}.K^{-1}$	0.73	0.65	0.65	0.65
External walls (brick): Density	$kg.m^{-3}$	1700	1462	1462	1462

Table 5 Values of R^2 and CV(RMSD) in the initial and calibrated models

	1st run period		**2nd run period**		**3rd run period**		**4th run period**	
	R^2	CV(RMSD)	R^2	CV(RMSD)	R^2	CV(RMSD)	R^2	CV(RMSD)
Initial model	0.38	4.13%	0.68	6.97%	-	-	-	-
1st calibrated model	0.81	2.15%	0.89	3.30%	0.73	1.86%	-	-
2nd calibrated model	-	-	0.88	2.17%	0.60	2.53%	0.70	3.39%
3rd calibrated model	-	-	-	-	0.76	2.11%	0.83	2.78%

To further illustrate the performance of the calibrated model, Figures 2 and 3 depict monitored office temperatures together with both inital and first calibrated model simulation results. This is done here for 9-day periods within the first and second periods.

DISCUSSION

As it can be seen from Table 5, the first and third automated calibrations significantly improved the model performance (during the second and forth run periods respectively) in terms of R^2 and CV(RMSD). The second calibration, however, slightly decreased the predictive model performance.

From these results, we conclude that the optimization-based simulation model calibration has a promising potential toward improving the run-time performance of embedded simulation engines in buildings' control and automation systems. However, the calibration process requires further improvement in terms of efficiency and consistency. Toward this end, future research will explore the possibilities to further enhance the recurrent calibration process via a more detailed process for the determination and interpretation 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 a recurrent optimization-based calibration of the thermal performance model of an office building. Data obtained via the monitoring system is 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 cost function was proposed, which equally weighted an error and a goodness of fit indicator. The results displayed a noticeable but not fully consistent improvement of the predictive potency of the calibrated model. Hence, the optimization-assisted simulation model calibration method represents a promising opportunity for performance enhancement in applications pertaining to building automation, diagnostics, facility management, and model-based systems control.

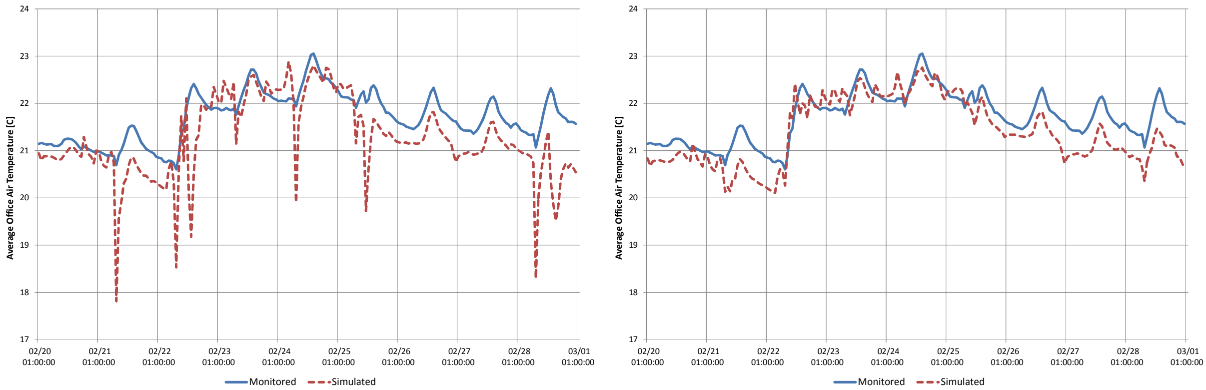

Figure 2 Monitored & simulated office temperature in 1st period: initial model (left), 1st calibrated model (right)

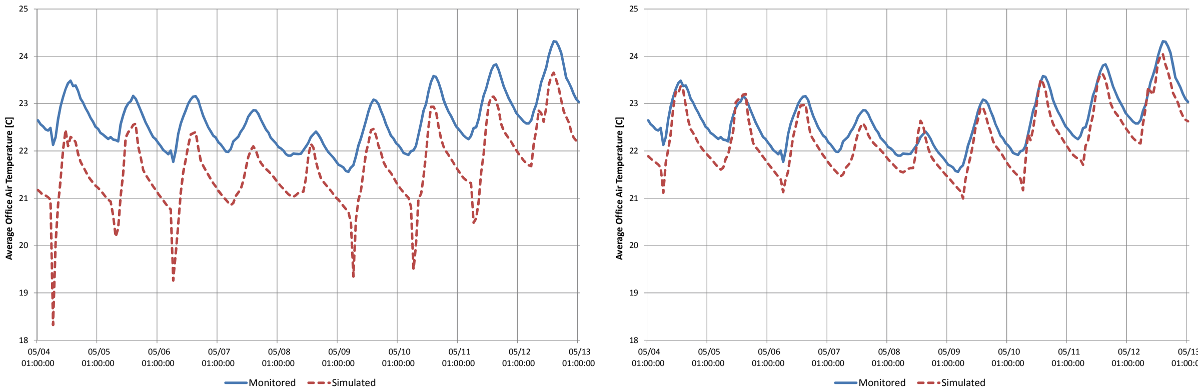

Figure 3 Monitored & simulated office temperature in 2nd period: initial model (left), 1st calibrated model (right)

ACKNOWLEDGMENT

The research presented in this paper is supported by funds from the program "Innovative Projekte" of the Vienna University of Technology as well as from the Austrian Science Foundation (FWF): Project: I 545-N22.

REFERENCES

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.

DesignBuilder 2008. DesignBuilder software Ltd. Simulation and CFD training guide, Available at: www.designbuilder.co.uk/.

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

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

Gösele, K., Schüle, W. 1983. Schall, Wärme, Feuchte. Bauverlag BV GmbH, ISBN: 978-3762527329.

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

Hensen, L. M., Lamberts, R. (ed.) 2011. Building performance simulation for design and operation. Spon Press, ISBN: 978-0-415-47414-6.

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

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

Matlab 2012. http://www.mathworks.com/products/matlab/.

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.

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.