CALIBRATION OF DOMESTIC HOT WATER SYSTEM SIMULATIONS WITH STOCHASTIC PROFILES

A. Bres[1]
[1]AIT Austrian Institute of Technology GmbH, Center for Energy, Vienna, Austria

ABSTRACT

Calibration makes it possible to reduce uncertainties in simulations. While a number of techniques have been proposed to achieve this in the context of building and system simulation, they have mostly been applied to static building or system parameters, rather than parameters related to occupant-driven loads. This paper investigates calibration in the presence of stochastic behavioral models for domestic hot water demand. With such stochastic models, simulation outputs and measurements should be aggregated to facilitate comparison. The paper shows in which measure parameters used in stochastic behaviour models can be calibrated in a Bayesian setting using such aggregation.

INTRODUCTION

Occupant behaviour models

From the point of view of building performance simulation, energy consumption in buildings can be considered to be a function of weather and static parameters (e.g. related to building physics) but also parameters related to occupant behaviour. The representation of occupant behaviour in building simulation tools ranges from simple static schedules to a variety of sophisticated models (Yan et al., 2015). Stochastic methods have been proposed to account for randomness and dynamic interactions in occupant behaviour, for instance for the modelling of occupancy or plug loads in buildings. Stochastic occupant models are not effective in all cases (Tahmasebi, 2016), and the purpose of simulation should be taken into account when selecting an occupant behavior model (Gaetani et al., 2016). In the context of building performance optimization, a recent contribution showed that the use of stochastic occupant behaviour models could yield different, more robust results (Ouf et al., 2020). The simulation of system controls performance is arguably a use case in which stochastic models could be useful, by allowing the robustness of controls to be assessed better than with static profiles. What is more, hot water demand is a particularly variable aspect of occupant behavior, with discrete and mostly short events which may hardly be represented with static profiles.

Calibration methods

Calibration, which we define as the adjustment of input parameters for a better agreement of simulation with empirical data, is an important step to ensure or improve the quality of simulations when measurements are available. A distinction can be made between manual calibration methods, which mostly rely on user experience or intuition, and automated calibration methods (Coakley et al., 2014). Automated calibration methods may for instance use mathematical optimization to minimize a function corresponding to discrepancy between measurements and simulation (Sun & Reddy, 2006). An alternative to optimization-based calibration is Bayesian calibration, which makes it possible to quantify uncertainties in calibration parameters, as well as discrepancies between model predictions and observed values and observation errors (Kennedy & O'Hagan, 2001). This calibration method has increasingly been applied to building simulation in the last years (Chong et al., 2017; Heo et al., 2012). The measured values y are assumed to follow Equation (1), where ς is the true function of (known) input parameters x, ϵ is measurement error, θ is the true value of the calibration parameters t, η is the model function of x and t, and δ represents model discrepancy.

$$y(x) = \varsigma(x) + \epsilon = \eta(x, \theta) + \delta(x) + \epsilon \qquad (1)$$

When calibrating building simulation against hourly measurements in such as Bayesian setting, consideration should be given to the choice of which values y to use, especially as using all available hourly values may result in massive computational costs (Chong et al., 2017; Remmen et al., 2019).

Although a wide range of calibration techniques are available, their application as reported in the literature generally is mostly limited to the calibration of static parameters only.

Metrics, aggregation and uncertainty

Building performance simulation typically yields large amounts of results, mostly in the form of temporally (time steps) and spatially (zones)

DOI: 10.3217/978-3-85125-786-1-77
© Creative Commons BY-NC-ND

differentiated physical variables. The analysis of simulation results usually involves the aggregation of such results into meaningful indicators. In the context of calibration, the goodness of fit can be quantified with metrics such as the normalized mean bias error or the coefficient of variation of the root mean squared error defined in Equation (2), where s_i and y_i represent the simulated and measured quantities at different time steps, respectively.

$$CVRMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(s_i - y_i)^2}}{\frac{1}{n}\sum_{i=1}^{n} y_i} \qquad (2)$$

It is generally acknowledged that the difficulty of calibration increases with shorter time scales. This is reflected in the acceptance criteria defined in ASHRAE Guideline 14 (ASHRAE, 2002), which require calibrated models to have CVRMSE under 15% for monthly values or 30% for hourly values. Similarly, uncertainty of building performance with behaviour models has been shown to depend strongly on the spatio-temporal scale at which results are compared (Yao, 2020).

Research question

Using calibration techniques in the context of stochastic occupant models raises some questions which this paper proposes to address. Directly comparing time series of stochastically determined quantities appears questionable, if not hopeless, because of the high uncertainty of results at the temporal scale of a sub-hourly simulation time step, even for known parameters. In this context, how can simulation outputs and measurements be aggregated in order to make comparison and calibration possible, and in which measure can non-static parameters used in stochastic behaviour models be calibrated?

METHOD

Approach

Given the difficulty to compare time series directly, the approach taken here is to transform the measured or simulated time series in such a way as to reduce the randomness and then perform the comparison and calibration based on the transformed data.

Simulation model

The approach is applied to simulations combining a physical model of hot water storage and stochastic models for domestic hot water demand in single-family houses served by district heating. Reliable data about both the demand and the physical system are required when trying to optimize system behavior (for instance to limit return temperatures). The storage tank and its hysteresis control make it impossible to observe hot water demand directly, and consequently to validate demand profiles without considering the storage system.

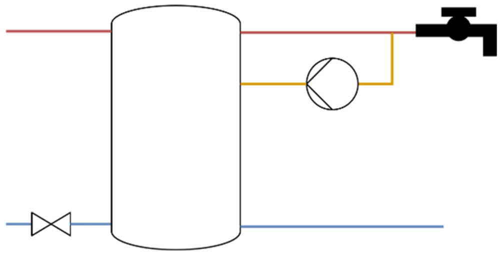

Figure 1: High-level diagram of the simulated system, with primary side on the left and secondary side on the right

The system is modelled with the general-purpose language Modelica, using the Modelica IBPSA library (*Modelica IBPSA library*, 2020) and the Modelica DisHeatLib library (*Modelica DisHeatLib library*, 2019). Central to this system is the hot water storage tank, which is modelled as a stratified tank with 4 volume segments. Simulations are carried out with the Dymola software, calling the simulation engine from Python using the BuildingsPy package (Wetter & USDOE, 2019). Both physical parameters (related for instance to the tank and its stratification) and stochastic demand parameters are subjected to uncertainties.

Stochastic hot water demand

While space heating loads to a large extent depend on weather conditions and building physics, loads from hot water consumption are much more dependent on occupant behavior, and characterized by a high time variability. They are accordingly difficult to predict and model. In addition to system-related information, inputs required for the modelling of loads from hot water consumption include mains water temperature, hot water temperature and the volume flow rates of hot water consumption. These can be specified at different time resolutions and following different methods, from fixed schedules to stochastically varying profiles. It has been argued that more realistic load profiles could be obtained with stochastic methods. Fischer et al. (Fischer et al., 2016) proposed a stochastic bottom-up model in which the number, times and durations of tappings are sampled from probability distributions for each activity requiring hot water.

In order to account for the high variability of hot water heating loads, stochastic profiles are generated following a method roughly similar to that proposed in the previously cited article, but with simplifications so as to use a minimum number of parameters. In particular, distinctions in activity types and in day type are not made. The number of

 © Creative Commons BY-NC-ND

tapping starts for each day is modelled as a Poisson distribution. Hot water temperature is assumed to be constant at 50 °C. The flow rate during each tapping is assumed to follow a normal distribution.

Calibration method

Since uncertainties cannot be ruled out from the system and even less from the hot water demand, it seems promising to use a Bayesian approach to calibration. A Bayesian approach inspired the reference work of Kennedy and O'Hagan (Kennedy & O'Hagan, 2001) is implemented. Uniform prior distributions are chosen for all the parameters summarized in Table 1. The calibration model is implemented in Python using PyMC3 (Salvatier et al., 2016). The state-of-the-art No-U-Turn Sampler (NUTS) is used to sample from the posterior distribution.

Calibration experiment

A calibration experiment is carried out with synthetic data in order to investigate the research question. The calibration parameters summarized in Table 1 are assumed to vary in the specified ranges. Latin hypercube sampling (McKay et al., 1979) is used to sample the corresponding parameter space effectively, for a number $n_{sim} = 200$ of simulations.

Table 1: Calibration parameters

PARAMETER	MIN	MAX	UNIT
Tank volume	0.15	0.25	m³
Tank height	1.0	2.0	m
Temperature of tank room	10	20	°C
Length of pipe to tank	0.5	4.5	m
Recirculation rate as a percentage of nominal flow rate	1.0	3.0	%
Mean number of taps per person day	2.0	4.0	
Mean volume flow rate of taps	0.01	0.02	l/s
Mean daily profile shift	-2.0	2.0	h

The simulation parameters summarized in Table 2 are assumed for all simulations.

Table 2: Simulation parameters

PARAMETER	VALUE
Simulation period	28 days
Output time step	10 minutes
Primary supply temperature	70 °C
Hot water demand temperature	55 °C

Result aggregation

The main result variable considered here is the heat power delivered to the system, as it would be measured by a district heating provider.
The raw results would be the time series formed by this variable at the output time steps $(\dot{q}_i)_{1 \le i \le n}$, which can be seen as a long vector (of length $n = 4032$ for 28 days simulation at a 10-minute time step). These results can be aggregated into shorter vectors or in scalar indicators. A simple example of indicator is the total supplied energy $Q = \frac{1}{1000} \sum_1^n \dot{q}_i \, \Delta t$ in kWh with Δt the output time step duration in h. Table 3 summarizes the aggregated values investigated in this paper.

Table 3: Processed outputs

NAME	SYMBOL	UNIT	LENGTH
Supplied energy	Q	kWh	1
Number of starts	n_{st}	1	1
Autocorrelation of supplied power	$R_{\dot{q}\dot{q}}(\tau)$	1	6
Time-of-day hourly averages of supplied power	$\dot{q}_{day,h}$	W	24

Autocorrelation at a given time lag τ provides a representation of the dynamics and of the typical duration of supply events. Time lags from one to 6 time steps are considered, as autocorrelation becomes almost null at time lags over one hour. The calibration experiments are repeated with different sets of processed outputs summarized in Table 4.

Table 4: Sets of processed outputs

NAME	A	B	C
Supplied energy	✓	✓	✓
Number of starts		✓	✓
Autocorrelation of supplied power		✓	✓
Time-of-day hourly averages of supplied power			✓

Surrogate models

The number of model evaluations necessary for sampling from the posterior makes it necessary to resort to surrogate models emulating the results of building simulation at much lower computational costs. Based on the n_{sim} simulations, fast surrogate models are trained to approximate processed outputs of the simulations.

 DOI: 10.3217/978-3-85125-786-1-77 © Creative Commons BY-NC-ND

Gaussian processes are used as surrogate models, taking advantage of their ability to learn nonlinear relations with high accuracy and to give an estimate of uncertainty at each point (Lim & Zhai, 2017; Rasmussen, 2003). The implementation of Gaussian processes in the scikit-learn library (Pedregosa et al., 2011) is used. One surrogate model is trained for each dimension of each one of the processed outputs described in the previous section.
Feature selection is carried out before training the surrogate models, using univariate linear regression tests to keep only significant calibration parameters for each surrogate model.

RESULTS AND DISCUSSION

Simulation results

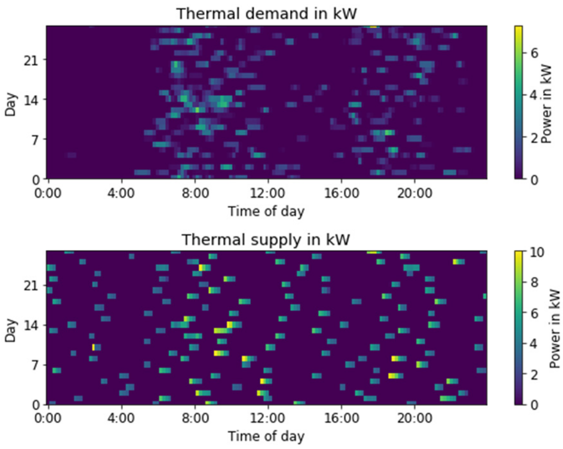

Figure 2: Day/time raster plots of demand and supply in a typical simulation run

Simulation results for a typical run are illustrated in Figure 2. These figures reveal the stochastic aspect of the results, with short spikes of heat supply occurring concentrated near times of higher hot water demand in the morning and evening, but rather randomly. Among the remarkable differences between thermal powers on the demand and supply side, one may note regular supply starts during the night and early afternoon at times of no or low demand, which may be ascribed to thermal losses in the storage.

Calibration results

Figure 3 shows a juxtaposition of the prior and posterior distributions of the calibration parameters, together with their true values for case. Starting from a uniform prior distribution, some calibration parameters keep a flat posterior distribution, meaning that these parameters cannot be calibrated based on the available data. For other parameters, such as tank volume, a much sharper posterior distribution can be obtained.

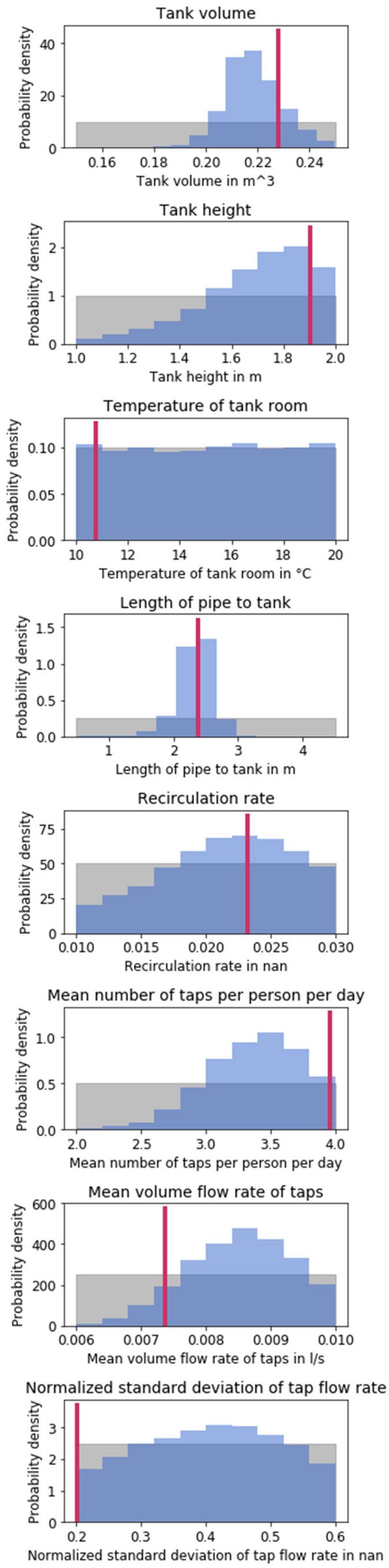

Figure 3: Prior distribution in grey, posterior distribution in blue, true value in red

 DOI: 10.3217/978-3-85125-786-1-77 © Creative Commons BY-NC-ND

The expected mean squared error from the true value of each parameter is used to quantify the distance from the prior and posterior distributions to the true value. The results summarized in Table 5 show the reduction in uncertainty made possible with the addition of additional output indicators from A to C. However, this only applies to some parameters, while for others the mean squared error from the true value remains almost constant, meaning that they remain "uncalibrated".

Table 5: Average mean squared distance from true value of normalized input parameters in prior and posterior distributions

PARAMETER	PRIOR	POSTERIOR		
		A	B	C
Tank volume	0.40	0.38	0.15	0.15
Tank height	0.49	0.41	0.32	0.30
Temperature of tank room	0.52	0.51	0.52	0.52
Length of pipe to tank	0.29	0.29	0.28	0.08
Recirculation rate	0.33	0.33	0.26	0.26
Mean number of taps per person per day	0.56	0.32	0.33	0.35
Mean volume flow rate of taps	0.33	0.41	0.37	0.35
Normalized standard deviation of tap flow rate	0.57	0.58	0.57	0.58
Mean daily profile shift	0.37	0.37	0.36	0.31

These results are confirmed when repeating the experiment with different choices of the ground truth, as illustrated in Figure 4. In these scatter plots, points on the diagonal mean that the distance from the truth remains equal for posterior and prior, meaning that calibration does not happen for the corresponding parameter. On the other hand, points below the diagonal point to an improvement in the parameter estimate and thus to successful calibration. For a parameter with limited impact like the temperature of the tank room, calibration fails in all the experiments. For the tank volume, calibration seems to succeed in cases B and C. For the mean daily profile shift, calibration fails in cases A and B, but succeeds with the additional output indicators in C.

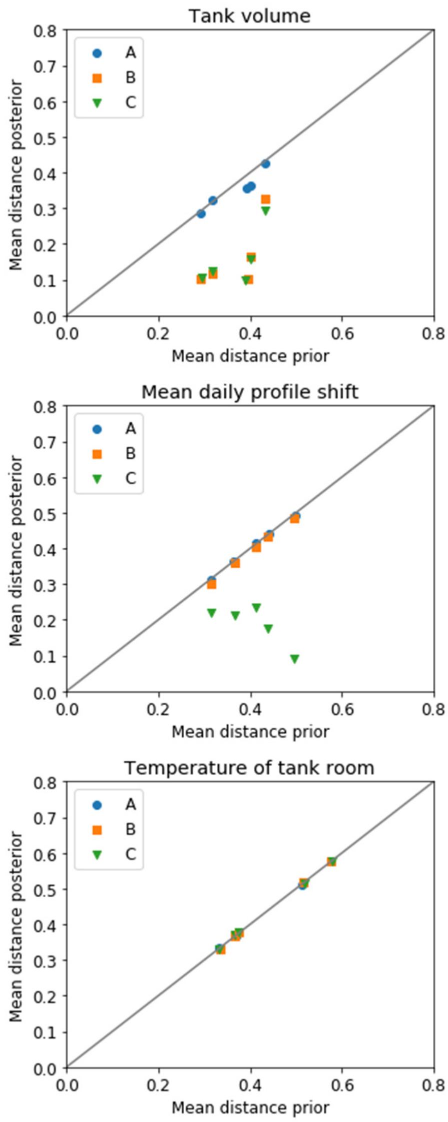

Figure 4: Mean distance from true value in prior versus posterior distributions for different "truths" and different sets of output indicators A, B and C

Discussion

Calibration and stochastic behavior models are two relevant components in the application of building performance simulation to support decisions in energy systems. Combining these two components raises the question adressed in this paper: if outputs at a given time step are stochastic, how can simulation and measurements be compared? We argue that simulation outputs should be processed adequately for such calibration, and carry out a calibration experiment with synthetic data to prove the approach.

The results show that processed outputs resulting from different ways of aggregating the raw simulation results make it possible to reduce the uncertainties in different input parameters. Different

 DOI: 10.3217/978-3-85125-786-1-77 © Creative Commons BY-NC-ND

input parameters affect simulation results differently, and thus have different impacts on different output indicators. For instance, parameters such as the profile shift have an influence on the temporal distribution of heating loads but not on their sum. Consequently, calibrating such parameters cannot be done using only the latter value, but becomes possible with output indicators reflecting temporal distribution, such as time-of-day hourly averages.

When selecting processed outputs for calibration, one should take into account their predictability, but also their "informativeness" for input parameters of interest. The predictability of processed outputs depends on the stochastic models involved and on the level of aggregation.

The exclusive use of synthetic data is arguably a limitation of the present study, as it neglects many of the complexities of real calibration problems. However, it is this use of synthetic data that makes it possible to investigate the ability of a calibration method to get closer to a ground truth otherwise not accessible. Therefore, further investigations with synthetic data may yield additional insight. For instance, the impact of the duration of a calibration period (here four weeks) on the quality of results would be of interest. It may be expected that longer calibration periods should lead to higher predictability for some aggregated results and thus more accurate calibration.

The possibly high number of uncertain parameters in typical occupant models is an issue which may deserve attention and has been avoided in the present paper by choosing a simple model with a limited number of parameters. The issue may be addressed with sensitivity analysis.

Several practical applications of the Bayesian calibration methodology may be considered, including fault detection and diagnosis, as well as improved controls. Probabilistic fault detection and diagnosis may be carried out with the present method by introducing uncertain parameters corresponding to faulty behaviour. For instance, an excessive recirculation rate may be diagnosed with the model described here. Calibrated simulation with stochastic profiles may also be used for the improvement of system controls. In this context, the consideration of uncertainties may be expected to result in more robust choices, as observed in published work on stochastic model predictive control.

CONCLUSION

This paper investigated calibration in the context of simulation models including physical components and stochastic behavioral models for domestic hot water consumption. In this context, simulation outputs and measurements should be aggregated in order to make comparison and calibration possible. The paper shows that, aggregating simulation results in various ways, parameters used in stochastic behaviour models can be calibrated in a Bayesian setting. A proof-of-concept of the approach is presented with synthetic data, using simulation with known input parameters as a ground truth. The degree of uncertainty reduction strongly depends on the parameters and on the output indicators used. For some parameters, calibration is impossible when looking only at aggregate energy consumption but becomes possible when looking at other output indicators. There is potential in investigating further output indicators and their combinations, considering the tradeoff between computation effort and quality of calibration results.

ACKNOWLEDGEMENT

This research was performed within the TEMPO project, funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No 768936.

REFERENCES

ASHRAE. (2002). *ASHRAE Guideline 14 – Measurement of Energy and Demand Savings* (ASHRAE Guideline 14). ASHRAE.

Chong, A., Lam, K. P., Pozzi, M., & Yang, J. (2017). Bayesian calibration of building energy models with large datasets. *Energy and Buildings*, *154*, 343–355. https://doi.org/10.1016/j.enbuild.2017.08.069

Coakley, D., Raftery, P., & Keane, M. (2014). A review of methods to match building energy simulation models to measured data. *Renewable and sustainable energy reviews*, *37*, 123–141.

Fischer, D., Wolf, T., Scherer, J., & Wille-Haussmann, B. (2016). A stochastic bottom-up model for space heating and domestic hot water load profiles for German households. *Energy and Buildings*, *124*, 120–128.

Gaetani, I., Hoes, P.-J., & Hensen, J. L. M. (2016). Occupant behavior in building energy simulation: Towards a fit-for-purpose modeling strategy. *Energy and Buildings*, *121*, 188–204. https://doi.org/10.1016/j.enbuild.2016.03.038

Heo, Y., Choudhary, R., & Augenbroe, G. A. (2012). Calibration of building energy models for retrofit analysis under uncertainty. *Energy and Buildings*, *47*, 550–560. https://doi.org/10.1016/j.enbuild.2011.12.029

Kennedy, M. C., & O'Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *63*(3), 425–464.

 DOI: 10.3217/978-3-85125-786-1-77 © Creative Commons BY-NC-ND

Lim, H., & Zhai, Z. J. (2017). Comprehensive evaluation of the influence of meta-models on Bayesian calibration. *Energy and Buildings*, *155*, 66–75.

McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, *21*(2), 239–245.

Modelica DisHeatLib library. (2019). https://github.com/AIT-IES/DisHeatLib

Modelica IBPSA library. (2020). https://github.com/ibpsa/modelica-ibpsa

Ouf, M., O'Brien, W., & Gunay, B. (2020). Optimization of electricity use in office buildings under occupant uncertainty. *Journal of Building Performance Simulation*, *13*.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Rasmussen, C. E. (2003). Gaussian processes in machine learning. *Summer School on Machine Learning*, 63–71.

Remmen, P., Schäfer, J., & Müller, D. (2019). Refinement of Dynamic Non-Residential Building Archetypes Using Measurement Data and Bayesian Calibration. *Proceedings of the 16th IBPSA internal conference*. Building Simulation 2019.

Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. (2016). Probabilistic programming in Python using PyMC3. *PeerJ Computer Science*, *2*, e55.

Sun, J., & Reddy, T. A. (2006). Calibration of building energy simulation programs using the analytic optimization approach (RP-1051). *HVAC&R Research*, *12*(1), 177–196.

Tahmasebi, F. (2016). *Exploring the effectiveness of occupant behavior models toward more reliable building performance simulation* [PhD Thesis]. Technische Universität Wien.

Wetter, M., & USDOE. (2019). *BuildingsPy*. https://doi.org/10.11578/dc.20190430.2

Yan, D., O'Brien, W., Hong, T., Feng, X., Gunay, H. B., Tahmasebi, F., & Mahdavi, A. (2015). Occupant behavior modeling for building performance simulation: Current state and future challenges. *Energy and Buildings*, *107*, 264–278. https://doi.org/10.1016/j.enbuild.2015.08.032

Yao, J. (2020). Uncertainty of building energy performance at spatio-temporal scales: A comparison of aggregated and disaggregated behavior models of solar shade control. *Energy*, *195*, 117079. https://doi.org/10.1016/j.energy.2020.117079

 DOI: 10.3217/978-3-85125-786-1-77 © Creative Commons BY-NC-ND