

## A Robust Approach For The Calibration Of the Material Properties in an Existing Wall

Alessandro Prada<sup>1</sup>, Andrea Gasparella<sup>2</sup>, Paolo Baggio<sup>1</sup>

<sup>1</sup> University of Trento, Trento, Italy

<sup>2</sup> Free University of Bozen/Bolzano, Bolzano, Italy

### Abstract

Calibration of building energy simulation model is essential to obtain an accurate model able to emulate the actual energy behavior of existing buildings.

There are two different approaches in the literature that are deterministic and Bayesian calibrations. The deterministic formulation of the calibration problem is ill-posed and consequently the pursuit of model discrepancy reduction may result in an over-fitting. On the contrary, Bayesian calibration assumes uncertain inputs and tries to reduce the uncertainty consistently with observed data but its implementation is the first stumbling block in its adoption.

This paper describes a new framework for systematically adjusting the parameters of a building simulation model using a deterministic approach but considering also the a-priori information about the model parameters. The new calibration procedure is tested on the calibration of a finite difference scheme applied to the solution of the heat equation in a timber wall. The calibrated model is compared against other calibrated models obtained by different procedures. The results highlight a better performance of the model and especially a higher reliability in emulating the wall behavior in other independent set of measurements.

### Introduction

In Europe, there is a clear long-term objective to decarbonise the energy system. Currently, the residential sector is responsible for 25.4% of the final energy consumption in the EU28 (Eurostat (2016)), but it also represents 63% of the potential building energy savings in 2050 (IEA (2010)). For these very reasons, the increase in the number of existing building retrofits will be a breaking point in meeting the European targets. Developing building energy simulation models (*BES*) for existing buildings is an essential part of evaluating many alternative energy saving measures (*ESM*). The reliability of the “cost-optimal approach” for identifying the optimal *ESM* mix and

reducing the energy demand of the European building stock relies on the accuracy of *BES* in emulating the actual energy performance of the existing building. Hence, the calibration of *BES* against experimental data is becoming more and more important.

The calibration of *BES* is an inverse problem complicated by the non-uniqueness of the solution since it is an ill-posed problem with an ill-conditioned solution. Therefore, a large number of calibrated models can be identified and not all of them are representative of the actual building behavior outside of the calibration period. As a results, the identified solution may not be close to the true solution and a validation, and often a refinement of the calibrated model, are required.

A detailed literature review of papers dealing with calibrated simulation of *BES* has been published (Reddy (2006); Coakley et al. (2014); Lee et al. (2015)). According to Coakley et al. (2014) the model calibration can be classified in manual or automated. In the first approach the calibration process involves tuning the initial simulation. This procedure can be time-consuming since the probability of identifying the actual parameters is low and the evaluation of the parameter validity is often left to the energy modeler experience (Reddy et al. (2007); Raftery et al. (2011); Angelotti et al. (2018)). On the contrary, automated approaches use computer code to assist or complete the model calibration. For instance, the calibration process can be formulated as an optimization problem by using the sum of discrepancy between *BES* and monitoring data (Raftery et al. (2011); Tahmasebi et al. (2012); Garrett et al. (2013); Penna et al. (2015); Sun et al. (2016)). This approach speeds up the calibration process, however the calibrated parameters may often reach the boundaries of the predefined variation range, thus assuming less plausible values (Penna et al. (2015)). This is an over-fitting of data since the *BES* is forced to recover the random observation errors rather than the true system state. Besides, each *BES* is an approximation of the actual system and it can neglect some

physics (Rouchier et al. (2018)) and, consequently, it introduces an epistemic uncertainty (Paté-Cornell (1996)). To a large extent, the formulations of calibration problem are based on the assumption of unbiased *BES*. For this very reason, there exists a set of parameters ensuring the model accurately emulate the actual system (Rouchier et al. (2018)) and, consequently, in the deterministic approach the main goal is to ensure that the model outputs fit the observed data as much as possible. Nonetheless, the identified set of parameters can never be an exact one, even when the model error is not considered, due to the observation errors. On the contrary, in the Bayesian approach inputs are assumed to be uncertain and the main goal is to reduce the uncertainty in the inputs in a consistent way with the observed data (Kennedy and O'Hagan (2001); Heo et al. (2012, 2015); Sadeghi et al. (2017); Menberg et al. (2017, 2019); Chong and Menberg (2018)). Knowledge about parameter uncertainty is incorporated by defining an a-priori distribution for each *BES* parameter. The a-priori information is about the parameters most likely values and their variability ranges resulting in a probability density function (*PDF*). Bayes's theorem is then used to compute a posterior parameter distribution that represents the solution of the calibration process. Learning the Bayesian theory and understanding how to implement is the first stumbling block in the adoption of this approach.

The use of the a-priori information of model parameters, such as feasible ranges, data sheets or physical values, is also helpful in the regularization of the calibration problem avoiding the over-fitting in the deterministic approach. The regularization provides a general framework for increasing the stability of calibration process with the possible drawback of lower accuracy. In a pioneer work, Carroll and Hitchcock (1993) used the optimization of a weighting function that was constructed as a sum of discrepancies between simulated and measured energy consumptions but also with the values of model parameters and their typical preferred values. In this way, the authors tried to maintain reasonable parameters during the calibration process. A similar approach was also used by Sun and Reddy (2006), which used the distance from the initial guess as a penalty function limiting the deviation of calibrated parameters from the most likely value. More recently, Miletic et al. (2017) have followed a similar approach by minimizing a weighted sum of *BES* discrepancy and a penalty function used as a regularization term. Differently from previous work, the penalty function is a monotonically increasing function depending on the difference between the probability of the new parameter with respect to the probability of the initial guess. The reliability of this approach is strongly linked to the choice of weights. A high weight assigned to the distance between *BES* outcomes and observations can lead again to over-

fitting. On the contrary, increasing the weight to the penalty function can induce to over-believing the a-priori information. The weight choice can easily be verified with the model validation by using an independent set of data observation. However, an adjustment of the weight set implies a new optimization function and, consequently, the necessity to solve again the optimization problem.

In this work, we propose a calibration framework based on a bi-criterion optimization problem aiming at fitting the *BES* outcomes to measured data and, simultaneously, a penalty function incorporating the a-priori information coming from manufacturer datasheet. A Pareto front is obtained from the solution of the optimization problem and a Post Pareto analysis is then proposed for the identification of the calibrated values of the parameters. This procedure is tested on the calibration of a finite difference scheme applied for the solution of the heat equation in the test case proposed by Angelotti et al. (2018). Finally, a performance comparison is conducted on different validation periods in order to compare the different calibrated models obtained by using either manual and automatic process.

## Methods

The deterministic calibration procedure is fairly consolidated in the literature especially in the early stages. The initial steps are the energy audit, in order to obtain hierarchically high information, and the data acquisition. Following on from this point, the process continues with the definition of the initial model using the initial guess of model parameters. Then, the dimension of the parameter space is reduced by means of a sensitivity analysis aiming at eliminating the parameters that have a limited impact on the model outcomes. Hereafter, there is the actual calibration phase in which the energy modeler tries to reduce the discrepancy between the data observations and the model outcomes. Finally, the last step concerns the validation of the calibrated model using independent sets of measurements. This work focuses mainly on the last two phases, starting from the manufactured declared data and not reducing the parameter space given the simplicity of the test case.

## Test Case

The case study is a prototype timber building (Angelotti et al. (2018)) located in Trento, in northern Italy (Figure 1). The ground floor level is 60 cm above the ground level and the building has a foundation characterized by concrete plinths. The external walls are highly insulated platform-frame structure. The U-value of the external wall is  $0.117 \text{ W m}^{-2} \text{ K}^{-1}$  thanks to four layers of insulation materials. Behind the external timber coating and a ventilated cavity, there are 10 cm of high density mineral wool. Then, behind 6.5 cm of a sandwich made of Lami-



Figure 1: Prototype building used as test case

nate Veneer Lumber panels with a low density mineral wool, there are two layers of low density mineral wool thick respectively  $9.5\text{ cm}$  and  $5.5\text{ cm}$ . Finally, two plasterboard layers are the internal finish of the wall. The building is used for experimental activities, hence its environmental variables as well as the HVAC operations are continuously monitored through a monitoring system that integrates the *KNX* and *BACnet* communication protocols. These protocols allow communication of building automation and control systems with heating, ventilating, air-conditioning, lighting and pieces of equipment. The *BACnet* protocol provides mechanisms for computerized building automation devices to exchange information, regardless of the particular building service they perform.

Besides, the south facing wall is equipped with heat flow meters (*HFM*) and temperature probes positioned on the surfaces (*TS*) and on the internal layers in order to register the envelope performances (Figure 2). The probes and the data acquisition chain accuracy is estimated as  $\pm 0.5^\circ\text{C}$  for the thermocouples (i.e. *T3A*, *T2A* and *T1A*) placed on the internal layers and  $\pm 0.3^\circ\text{C}$  for the Platinum Thermo-Resistances on the internal (*TSi*) and external surfaces (*TSe*) of the opaque components. The certified accuracy of the heat flow meters is  $\pm 3\%$  at  $100\text{ W m}^{-2}$ . However, due to the high insulation level of the external wall, the measured heat flux are much lower and, consequently, the uncertainty in measurement increases due to the low signal to noise ratio. Four different measurement periods are used in this study (Table 1). According to Angelotti et al. (2018) we used *P1* as calibration period. This series is characterized by a free floating internal temperature since the HVAC were turned off. Hereafter three different periods are used for validation. *P2* is another free floating period but with a greater temperature variation between day and night. On the contrary, *P3* is a period in which the heating system was turned on with a constant set-point temperature of  $20^\circ\text{C}$ . Finally, *P4* is a period comprising both a free floating regime and a conditioned period. The analysis target concerns the cali-

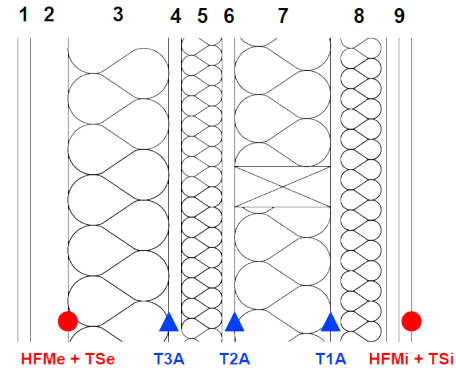


Figure 2: Wall cross section and probe positions

Table 1: Series of experimental observations used for calibration and validation

Code	Period	Data	Purpose
P1	29/5 - 7/6	2554	Calibration
P2	25/6 - 4/7	2560	Validation
P3	17/2 - 24/3	10399	Validation
P4	2/11 - 18/11	4572	Validation

bration of a finite difference scheme used to solve numerically the heat equation of the monitored external wall. A fully implicit finite difference scheme were implemented by means of a Fortran code (Prada et al. (2014)). Then, a Dirichlet boundary condition is set both at the external and internal surface of the wall by using the measured temperature profile of *TSe* and *TSi*. Afterward, the simulated internal and external heat fluxes are compared against the measured values. Table 2 shows the initial guess (*IG*) values

Table 2: Initial guess values (*IG*) and standard deviations of thermophysical properties from a-priori knowledge.

Layer	$\rho$ [ $\text{kg m}^{-3}$ ]	$\lambda$ [ $\text{W m}^{-1} \text{K}^{-1}$ ]	$c$ [ $\text{J kg}^{-1} \text{K}^{-1}$ ]
3	$135 \pm 27$	$0.04 \pm 0.008$	$1030 \pm 206$
4	$530 \pm 159$	$0.2 \pm 0.06$	$1600 \pm 480$
5	$40 \pm 8$	$0.035 \pm 0.011$	$1030 \pm 309$
6	$530 \pm 159$	$0.2 \pm 0.06$	$1600 \pm 480$
7	$40 \pm 8$	$0.035 \pm 0.011$	$1030 \pm 309$
8	$40 \pm 8$	$0.035 \pm 0.011$	$1030 \pm 309$
9	$900 \pm 270$	$0.21 \pm 0.063$	$1000 \pm 300$

of the thermophysical properties from the manufacturer data-sheets. These data are used in the initial model and they are considered as the most probable values. Table 2 shows also the standard deviations we assumed based on the source uncertainties and on the variability of the material properties, for example due to the material heterogeneity or to the compression level. These pieces of information are used for the definition of the a-priori *PDF* of each calibration parameter. This is a key aspect in the calibration procedure since a mistaken distribution can undermine the calibration procedure and over-believing of the initial guess values.

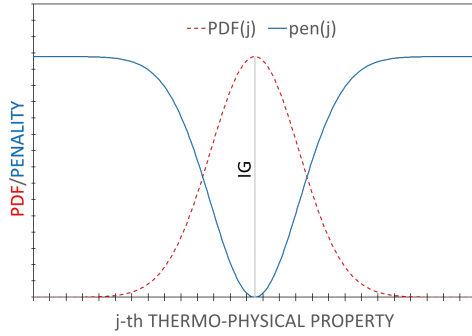


Figure 3: Definition of the  $pen_j$  function by using the parameter  $PDF$

### Calibration procedure

In the proposed approach, the calibration is carried out by means of an multi-objective optimization algorithm aimed at reducing the discrepancy between the model results and the measurements and, simultaneously, ensuring a high likelihood of the calibrated parameters by using the a-priori knowledge of the energy audit. The coefficient of determination ( $r^2$ ) is the first optimization objective with the purpose of quantifying the goodness of fit of the predicted internal heat fluxes and the measured ones, differently from Angelotti et al. (2018) which used the coefficient of variation of the root mean square error normalized by an effective heat flux value. The  $r^2$  indicates the extent to which simulated outcomes are close to regression line of the measured values. It is a statistical index commonly used since it varies in the range  $[0 \div 1]$ . ASHRAE (2013) and IPMVMP (2012) suggest to ensure as a rule of thumb a  $r^2$  greater than 0.75 in calibrated models.

The second objective is the minimization of a penalty function that represents the regularization term. From the a-priori knowledge we know the initial guess ( $IG$ ) of the three thermophysical properties of each layers. Then we assume a normal distribution of each parameter centered at the  $IG$  to take into account the random variability of properties. Other  $PDF$  can however be adopted in the same calibration framework to consider the other types of uncertainty, the asymmetry of the parameter distribution around the  $IG$  or the quality of the a-priori information. From this point, we define the penalty function of the parameter ( $v_j$ ) as the difference between the  $PDF$  of the  $IG_j$  and the  $PDF$  of the  $j$ -th calibration parameter (eq. 1) as illustrated in Figure 3.

$$pen_j = \frac{(1 - \exp(-\frac{v_j - IG_j}{2 \cdot Var(v_j)}))}{\sqrt{2 \cdot Var(v_j)}} \quad (1)$$

The  $PDF$  choice plays therefore a key role in the calibration process and in the reliability of the obtained calibrated models since an incorrect  $PDF$  may penalize good solutions or it may not penalize values far from the actual ones. The second optimization goal is defined as the total penalty function which is the

sum of all the  $pen_j$  of the proposed set of parameters.

$$Pen = \sum_j pen_j \quad (2)$$

### Optimization algorithm

A large number of evolutionary algorithms for solving multi-objective optimization problems have been developed in the literature, however the *NSGA-II* is to a large extent the most popular (Hamdy et al. (2016)). The *NSGA-II* uses elitism by maintaining the current and the previous population. Then, after the population mating, the population is sorted according to the non-domination concept and the best ranking solutions are selected as the next parent population. The optimization algorithm we developed in Matlab is a customization of the original *NSGA-II* and, especially, we modified the sampling and convergence criterion. Firstly, the possible *ESM* combinations in the variable domain are selected by a Latin Hypercube Sampling to increase the sampling performance by increasing the sample uniformity in the hyperspace. Finally, the hypervolume measure (a.k.a. *S-metric*), proposed by Zitzler and Thiele (1999), is adopted as a stopping criterion. A threshold of  $10^{-4}$  in the variation of the normalized *S-metric* between two consecutive generations has been adopted as the convergence criterion in the code.

### Results

The results shown in this section are the two objectives adopted in the proposed calibration framework. In particular, the two's complement of the sum of the coefficient of determinations computed for the internal and external heat fluxes and the penalty index were minimized. The chart in Figure 4 shows the objectives of each combination of calibration parameters (*Sim*) while the light red marks identify the Pareto front. Note that two different points are highlighted with different marks on the Pareto curve. The square mark indicates the identified solution by using only the discrepancy minimization (*obj*). This point indicates the result of an automated calibration process in which no attention is paid to the distance from the initial guess values of the parameters (i.e. the thermophysical properties declared by the manufacturer). The star mark instead highlights the optimal point (*Opt*) selected after the Post Pareto analysis. In this work, we select the solution on the Pareto Front with the penalty index closest to the threshold but lower than it. The threshold has been defined as 1 because the a-priori information is considered very reliable since the wall is a prefabricated construction with declared thermophysical properties. Finally, the green squares indicates the two objective indexes obtained by using the calibrated models obtained in Angelotti et al. (2018). The three calibrated models were obtained by different energy modelers by using either automated or manual calibration processes.



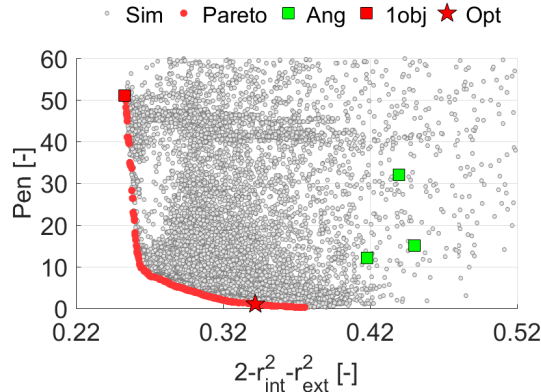


Figure 4: Results of calibration period (P1)

Several optimal solutions can be identified on the Pareto front during the calibration period, and the goodness of fit indexes range from 0.2532 to 0.3756. The calibrated model selected by the post Pareto analysis is characterized by a two's complement of the coefficients of determination equal to 0.3417. This solution is further from measurement with respect to the *1obj* model but it outperforms the calibrated models obtained in Angelotti et al. (2018) also because obtained by using a different goodness of fit index. Nonetheless, it should be emphasized that *Opt* is better also in term of penalty function. This means that *Opt* has the calibration parameters closer to the most probable values that are declared by the manufacturers. As a matter of fact, a deviation lower than 5% with respect to the nominal value is found in most of the parameters. Table 3 highlights in bold the parameters for which an absolute variation greater than 10% was noted. The most consistent changes concern the layer 7 where the thermal conductivity of the plasterboard is reduced by 13%, the specific heat increases by 19.9% and the specific mass increases by 59%. This latter property is affected by greater variations also in the other layers. For instance, it is reduced by 58% in the two *LVL* layers and increased by 26% and 27% respectively in the layers 7 and 8. Table 3 shows the figures concerning the calibrated model *1obj*, which is obtained by neglecting the distance from the likely values.

The results clearly show an over-calibration of the model *1obj* where the thermal capacity of the wall is considerably reduced in order to have a good consistency with the measurements. For this reason, the specific heats and the specific masses reach the low end of the variation range in the layers from 3 to 6, thus assuming little plausible values.

### Model validations

The different calibrated models, and those identified by the Pareto front, are validated using three independent sets of experimental observations. These periods are characterized by different boundary conditions with respect to the calibration set, as described

Table 3: Thermophysical properties of the calibrated models.

Layer	$\rho$ [kgm <sup>-3</sup> ]	$\lambda$ [Wm <sup>-1</sup> K <sup>-1</sup> ]	$c$ [Jkg <sup>-1</sup> K <sup>-1</sup> ]
<i>Opt</i> calibrated model			
3	142	0.040	1025
4	<b>220</b>	0.203	1613
5	36	0.035	1053
6	<b>220</b>	0.203	1613
7	<b>51</b>	0.035	1041
8	<b>51</b>	0.035	1045
9	<b>1431</b>	<b>0.181</b>	<b>1199</b>
<i>1obj</i> calibrated model			
3	<b>114</b>	0.053	<b>696</b>
4	<b>216</b>	0.193	<b>645</b>
5	<b>24</b>	<b>0.031</b>	<b>617</b>
6	<b>216</b>	0.193	<b>645</b>
7	<b>52</b>	0.033	1129
8	<b>50</b>	0.034	1191
9	<b>1439</b>	<b>0.087</b>	<b>1591</b>

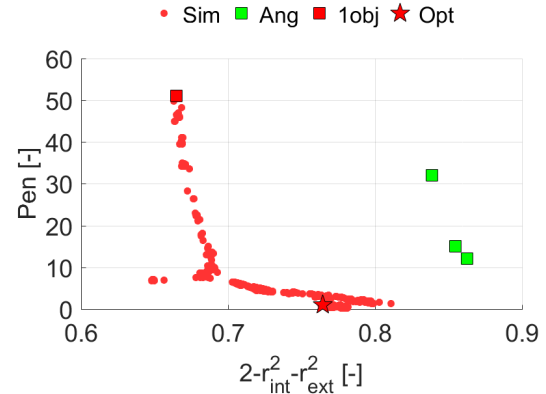


Figure 5: Results of validation (P2) period

in Table 1. Hence, these validations are an effective test to identify the possible calibration over-fitting. The graph in Figure 5 shows the validation during another period in which the building internal temperature is in free floating. An important result is the worsening of the accuracy index for all the models, with a two's complement of the coefficient of calibration in the range 0.6480 to 0.8628. In this test, there is a greater temperature variation between day and night, due to the climate characteristics, and this causes an inversion of the flow direction especially on the outer side of the wall. On the contrary, the internal heat fluxes is stable but very close to zero due to the high insulation of the wall and the limited temperature difference between the internal and the external environment. Hence, the noise has a greater impact on the low voltage signal of the thermopile (*HFMi*) and, consequently, the uncertainty in data observations is high. In these conditions, therefore, the models with a high penalty function continue to behave better than those closer to the most likely values of the manufacturer. This aspect clearly emerges

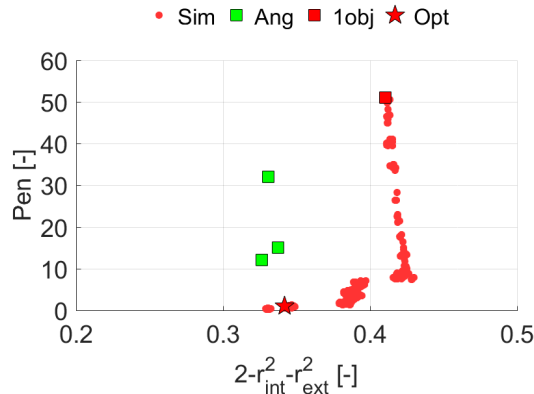


Figure 6: Results of validation (*P3*) period

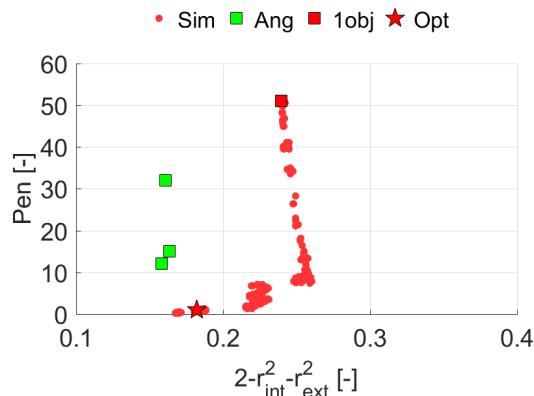


Figure 7: Results of validation (*P4*) period

by looking at the trends of the goodness of fit indexes in Figure 6. The HVAC systems are always turned on in the building during the heating period *P3* and the heat flow is always outgoing. This causes an important variation in the shape of the front since the calibrated models with a penalty index lower than 10 behave better than other solutions. The lower uncertainty in the measurements, due to the greater intensity of the heat flow meter signal, penalizes the over-fitted models. For this very reason, *Opt* outperforms *1obj*. It should also be noted that the *Ang* models have very good performance, since the calibrated parameters have been selected on the basis of the experience of the energy modelers. A similar result emerges from the *P4* validation test in which the building has a mixed operation, especially an internal free floating temperature of the room in the first part up to the switching on of the heating system. Notice that the models characterized by a lower penalty index are more accurate in predicting the actual behavior of the wall (Figure 7). The *Opt* solution, identified with the post Pareto analysis, performs better than *1obj* and slightly worse than the *Ang* models. Altogether, the *Opt* model provides good performance in all the validation periods, unlike the other calibrated models, that significantly deviate from the likely values of material properties to adapt the model predictions to the uncertain measures. However, these models tend to reduce their accuracy when the uncertainty is re-

duced due to a better signal-to-noise ratio of the heat flow meters.

## Conclusion

This paper has shown a calibration procedure including the a-priori information in a deterministic approach by adding a further goal in the optimization problem. This new objective has been defined as the difference between the probability densities function of the initial guess value and the parameters obtained by the optimization algorithm.

Following on from this point, the procedure is tested on the calibration of a simple test case, but with noisy experimental observations that increase the over-fitting risk. The results clearly show how the choice of a model with a low penalty index (*Pen*) tends to reduce the accuracy of the model, with respect to the *1obj* model, in the calibration period and in periods with similar characteristics. However, the models that are closer to the values of the declared properties behave better when they are validated on sets of measures with different operating characteristics. This, therefore, makes it possible to obtain a model that is more in keeping with reality and, consequently, more effective in predicting the real behavior of the wall. Additionally to the reduction of the over-calibration risk, this approach also allows to reduce the time required to refine the model parameters following the discrepancies pointed out in the validation phase.

It is important to employ reliable a-priori information to have a really effective procedure. Hence, the correct definition of the *PDF* of each calibration parameter is the key aspect and it have to consider the quality of the a-priori knowledge, in order to avoid the over-believing of the initial guess values.

## Acknowledgment

The authors would like to thank the Autonomous Province of Trento in supporting the construction of the experimental building, as part of the project "Dolomito Pro" funded by the EU EDR 2007 – 2013 found. This study is performed as part of the Project "BIM Methodologies for a New Industrialization of Energy Redevelopment Interventions of Existing Building Assets", funded by Fondazione Caritro.

## Nomenclature

$c$	Specific Heat [ $J kg^{-1} K^{-1}$ ]
$IG_j$	Initial guess value of parameter $j$
$\lambda$	Thermal conductivity [ $W m^{-1} K^{-1}$ ]
<i>PDF</i>	Probability Density Function [–]
<i>Pen</i>	Penalty function [–]
$Pen_j$	Penalty function of parameter $j$ [–]

$r^2$	Coefficient of determination [–]
$\rho$	Specific Mass [ $kg\ m^{-3}$ ]
$T$	Temperature [ $K$ ]
$v_j$	Value of parameter $j$
$Var(v_j)$	Variance of parameter $j$

## References

- Angelotti, A., M. Martire, L. Mazzarella, M. Pasini, I. Ballarini, V. Corrado, G. De Luca, P. Baggio, A. Prada, F. Bosco, and C. Cornaro (2018). Building energy simulation for nearly zero energy retrofit design: The model calibration. In *Proceedings - 2018 IEEE International Conference on Environment and Electrical Engineering*. Palermo (IT), 12-15 June 2018.
- ASHRAE (2013). *Handbook: Fundamentals*. American Society of Heating, Refrigerating and Air Conditioning Engineers (ASHRAE). Atlanta (USA).
- Carroll, W. and R. Hitchcock (1993). Tuning simulated building descriptions to match actual utility data. In *ASHRAE Transactions*, Volume 99 Part 2, pp. 429–439.
- Chong, A. and K. Menberg (2018). Guidelines for the bayesian calibration of building energy models. *Energy and Buildings* 174, 527 – 547.
- Coakley, D., P. Raftery, and M. Keane (2014). A review of methods to match building energy simulation models to measured data. *Renewable and Sustainable Energy Reviews* 37, 123 – 141.
- Eurostat (2016). *Complete Energy Balances—Annual Data 2016*.
- Garrett, A., J. New, and T. Chandler (2013). Evolutionary tuning of building models to monthly electrical consumption. In *ASHRAE Transactions*, Volume 119 Part 2, pp. 89–100.
- Hamdy, M., A.-T. Nguyen, and J. L. Hensen (2016). A performance comparison of multi-objective optimization algorithms for solving nearly-zero-energy-building design problems. *Energy and Buildings* 121, 57–71.
- Heo, Y., G. Augenbroe, D. Graziano, R. T. Muehleisen, and L. Guzowski (2015). Scalable methodology for large scale building energy improvement: Relevance of calibration in model-based retrofit analysis. *Building and Environment* 87, 342 – 350.
- Heo, Y., R. Choudhary, and G. Augenbroe (2012). Calibration of building energy models for retrofit analysis under uncertainty. *Energy and Buildings* 47, 550 – 560.
- IEA (2010). *Energy Technology Perspectives 2010—Scenarios & Strategies to 2050*.
- Efficiency Valuation Organization (2012). *International Performance Measurement and Verification Protocol: concepts and options for determining energy and water savings. Volume I*.
- Kennedy, M. and A. O'Hagan (2001). Bayesian calibration of computer models. *Journal of Royal Statistical Society* 63(3), 425–464.
- Lee, S. H., T. Hong, M. A. Piette, and S. C. Taylor-Lange (2015). Energy retrofit analysis toolkits for commercial buildings: A review. *Energy* 89, 1087 – 1100.
- Menberg, K., Y. Heo, and R. Choudhary (2017). Efficiency and reliability of bayesian calibration of energy supply system models. In *Proceedings of BS2017 International Conference on IBPSA*. San Francisco (USA), 7-9 August 2017.
- Menberg, K., Y. Heo, and R. Choudhary (2019). Influence of error terms in bayesian calibration of energy system models. *Journal of Building Performance Simulation* 12(1), 82–96.
- Miletić, M., C. Dipasquale, and R. Fedrizzi (2017). Calibration of the energy simulation models using tikhonov-type regularization: Application to a residential building apartment. In *Proceedings from BSA2017: Building Simulation Applications Conference*, pp. 175–184. Bolzano (IT), 8-10 February 2017.
- Paté-Cornell, M. E. (1996). Uncertainties in risk analysis: six levels of treatment. *Reliability Engineering and System Safety* 54, 95–111.
- Penna, P., F. Cappelletti, A. Gasparella, F. Tahmasebi, and A. Mahdavi (2015). Multi-stage calibration of the simulation model of a school building through short-term monitoring. *Journal of Information Technology in Construction* 20, 132–145.
- Penna, P., A. Gasparella, F. Cappelletti, F. Tahmasebi, and A. Mahdavi (2015). Optimization-based calibration of a school building based on short-term monitoring data. In *Proceedings of the 10<sup>th</sup> European Conference on Product and Process Modelling*. Vienna (AT), 17-19 September 2014.
- Prada, A., F. Cappelletti, P. Baggio, and A. Gasparella (2014). On the effect of material uncertainties in envelope heat transfer simulations. *Energy and Buildings* 71, 53 – 60.
- Raftery, P., M. Keane, and A. Costa (2011). Calibrating whole building energy models: Detailed case study using hourly measured data. *Energy and Buildings* 43(12), 3666 – 3679.

- Raferly, P., M. Keane, and J. O'Donnell (2011). Calibrating whole building energy models: An evidence-based methodology. *Energy and Buildings* 43(9), 2356 – 2364.
- Reddy, T. A. (2006). Literature review on calibration of building energy simulation programs: Uses, problems, procedure, uncertainty, and tools. In *ASHRAE Transactions*, Volume 112 Part 1, pp. 226–240.
- Reddy, T. A., I. Maor, and C. Panjapornpon (2007). Calibrating detailed building energy simulation programs with measured data - part i: General methodology. *HVAC&R Research* 13(2), 221–241.
- Rouchier, S., M. Rabouille, and P. Oberlé (2018). Calibration of simplified building energy models for parameter estimation and forecasting: Stochastic versus deterministic modelling. *Building and Environment* 134, 181 – 190.
- Sadeghi, S. A., N. M. Awalgaonkar, P. Karava, and I. Bionis (2017). A bayesian modeling approach of human interactions with shading and electric lighting systems in private offices. *Energy and Buildings* 134, 185 – 201.
- Sun, J. and T. A. Reddy (2006). Calibration of building energy simulation programs using the analytic optimization approach. *HVAC&R Research* 12, 177–196.
- Sun, K., T. Hong, S. C. Taylor-Lange, and M. A. Piette (2016). A pattern-based automated approach to building energy model calibration. *Applied Energy* 165, 214 – 224.
- Tahmasebi, F., R. Zach, M. Schuß, and A. Mahdavi (2012). Simulation model calibration: an optimization-based approach. In *Proceedings - BauSim 2012 German-Austrian IBPSA Conference*. Berlin.
- Zitzler, E. and L. Thiele (1999). Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. *IEEE Transactions on Evolutionary Computation* 3(4), 257–271.