Build Simul (2022) 15: 1883–1898
https://doi.org/10.1007/s12273-022-0900-5

A building energy models calibration methodology based on inverse modelling approach

Vicente Gutiérrez González, Carlos Fernández Bandera (✉)

Installations, construction and structures, School of Architecture, University of Navarra, Campus Universitario, Pamplona, 31009, Navarra, Spain

Abstract

Nowadays, building energy models (BEMs) are widely used, particularly in the assessment of energy consumption in buildings to address the potential savings that can be generated. The realisation of a dynamic energy model based on high-fidelity physics (white-box models) requires a tuning process to fit the model to reality, due to many uncertainties involved. Currently some research trends try to reduce this performance gap by modulating different types of experimental parameters such as: capacitances or infiltration. The EnergyPlus simulation software, in its latest versions, has implemented an object: HybridModel:Zone that calculates the infiltration and internal mass of buildings using an inverse modelling approach that employs only the measured indoor temperature data to invert the heat balance equation for the zone under study. The main objective of this paper is to reduce the execution time and uncertainties in the development of quality energy models by generating a new calibration methodology that implements this approach. This uses, as a starting point, a research created by the authors of this study, which was empirically and comparatively validated against the energy models developed by the participants in Annex 58. It is also worth highlighting the empirical validation of the HybridModel:Zone object, since it was activated in all scenarios where its execution is possible: periods of seven days or more of free oscillation and periods in which the building is under load. The findings are promising. The data generated with the new methodology, if compared with those produced by the baseline model, improve their resemblance to the real ones by 22.9%. While those of its predecessor did it by 15.6%. For this study, the two dwellings foreseen in Annex 58 of the IEA ECB project have been modelled and their real monitoring data have been used.

Keywords

internal thermal mass;
energy simulation;
EnergyPlus;
inverse model;
Annex 58;
infiltration;

Article History

Received: 09 February 2022
Revised: 16 March 2022
Accepted: 21 March 2022

1 Introduction

There is a growing interest in full-scale validation of specific elements and whole building models to characterise energy performance and efficiency (Strachan et al. 2016). This validation will reduce the gap between the post-construction behaviour of the building and the theoretically designed performance, as discussed by Subbarao et al. (Subbarao 1988; Subbarao et al. 1990) or Burch et al. (1990). Building energy simulation is at the heart of the energy assessment of buildings. Reliable quantification of energy conservation measures (ECMs) will increase confidence in simulation as an effective tool to increase retrofit investment (Siddharth et al. 2011; Fernández Bandera 2018). However, the simulation of building performance is a very loose science. The industry needs that the BEM casts confidence into the market (Gutiérrez González et al. 2019).

The 2013 ASHRAE fundamentals handbook (ASHRAE 2013) classifies the different types of energy models into two categories: physical or white-box models generated by direct modelling and data-driven or black-box models created by inverse modelling.

Developing a building energy model does not ensure that its performance resembles reality. As a general rule, what is called the "Building performance gap" occurs, that is, the distance between the behavior of the model and the

E-mail: cfbandera@unav.es

List of symbols

A	constant term
A_i	area of the i-th internal surface of the zone (m^2)
B	temperature term ($°C^{-1}$)
C	velocity term (s/m)
C_{pi}	specific heat capacity of the air from the i-th nearby zone (J/kg)
C_z	current zone air sensible heat capacity multiplier
D	velocity square term (s^2/m^2)
F_{sch}	infiltration schedule
h_i	convective heat transfer coefficient of the i-th internal surface of the zone ($W/(m^2 \cdot K)$)
I	infiltration
I_{design}	design infiltration rate, flow/zone, flow/floor area, flow/exterior area, flow/exterior wall area, or air changes/hour
$\dot{m}_{inf}$	mass flow rate of air from the i-th nearby zone (kg/s)
$\dot{m}_{i\_zone}$	mass flow rate of air from the i-th nearby zone (kg/s)
$\dot{Q}_i$	the i-th internal sensible heat gain rate of the current zone (W)
$\dot{Q}_{sys}$	sensible heat transfer rate due to HVAC system supply (W)
T_∞	outdoor air temperature (K)
T_{obd}	outside temperature (K)
T_{si}	temperature of the i-th internal surface of the zone (K)
T_z	current zone air temperature (K)
T_{zi}	temperature of the air from the i-th nearby zone (K)
T_{zone}	indoor temperature per thermal zone
WS	wind speed (m/s)
ρ	Spearman's rank correlation coefficient

Abbreviations

ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
BCE	Energy in Buildings and Communities Programme
BEM	building energy model
CTF	conduction transfer function
ECMs	energy conservation measures
EnEV	Energieeinsparverordnung (German Energy Saving Ordinance)
HVAC	heating, ventilation, and air conditioning
IEA	International Energy Agency
LBNL	Lawrence Berkeley National Laboratory
LOSC	free oscillation period
NSGA-II	non-dominated sorting genetic algorithm
ROLBS	randomly ordered logarithmic binary sequence
TFM	transfer function method

building. To correct this discrepancy, both the white-box and black-box models must be subjected to an adjustment or calibration process.

There are different approaches to the calibration process of white box models. Soebarto (1997) proposed a manual calibration by modifying different types of parameters through the observation of graphs. Royapoor and Roskilly (2015) manually calibrated an office building by comparing the data collected by sensors placed in the building. Yang and Becerik-Gerber (2015) adjusted the parameters they considered most sensitive in their detailed model at different levels. Robertson et al. (2015) selected the most sensitive parameters from the model and then used them in an optimization that helped minimize the discrepancy with the real building. Tüysüz and Sözer (2020) calibrated a residential building using a detailed model and short-term monitoring data. Ascione et al. (2020) used a Pareto optimization, calibration and modeling technique to realistically fit the detailed energy model of an industrial building. Martínez et al. (2020) conducted a performance comparison of approaches based on multi-objective optimization to calibrate white-box energy models.

To fit data-driven or black-box models to real data, inverse modeling is the most common solution. In recent years, interest in running BEMs based on this technique has been growing because they can provide a good estimate for measurement and verification. Different approaches in the field of inverse modelling have been implemented. Dhar et al. (2019) developed an inverse modelling methodology based on the Fourier series where temperature, as an input, was used to predict the energy demand of the building. Dong et al. (2005) used a support vector regression to calculate the building demand. Thomas et al. (2008) introduced in his inverse modelling methodology four linear regression methods to estimate building demand. Yu et al. (2010) incorporated a novel decision tree methodology into their process to ensure that it adequately responded to the required requirements. Kwok et al. (2011) used artificial neural networks trained by measured data to predict the energy demand of a building. Heo and Zavala (2012) and Zhang et al. (2013), created a predictive model based on inverse modelling using a Gaussian regression process. Fan et al. (2014) developed a model composed of eight learning algorithms to improve the predictions of their BEM. Robinson et al. (2017) used a gradient increase regression model to predict annual energy consumption in real estate.

Wang et al. (2018) proposed to apply an inverse modelling methodology based on an ensemble bagged tree. Sadaei et al. (2019) transformed the input data that will feed the energy model into multichannel images, converting it into a CNN (convolutional neural network) model. Yuan et al. (2019) applied a partial least squares regression (PLSR) and random forests to classify the main features to become input data in the model and thus predict, in this case, the heating demand of the building under study.

The inverse modelling technique is not unique to black box models. Detailed models can also be subjected to this adjustment process. Heo et al. (2012) used it by performing a Gaussian approach to the procedure. Giuliani et al. (2016) developed a methodology where the set points of the HVAC systems in the energy model were overridden by a special program to use the measured building surface temperatures in the adjustment process. Tahmasebi and Mahdavi (2012) introduced dynamic data into the white box model, which in combination with uncertainty indices generates the process to bring the BEM to reality. Chen et al. (2019) developed an energy model calibration methodology that combines white box models with metamodels in an office building. Nagpal et al. (2019) worked out a methodology to obtain the energy performance of a university campus where the models created are generated from the combination of black and white box BEM. Ramos Ruiz et al. (2016) used the building's indoor temperatures to guide the NSGA-II genetic algorithm towards finding the detailed energy model that fulfils the proposed objective function.

Physical or white box models characterize the dynamic thermal behavior of the building (Al-Homoud 2001; Crawley et al. 2008; Foucquier et al. 2013). And programs such as EnergyPlus can consider the thermal mass of the building envelope using the transfer function method (TFM), as well as the finite difference method according to the EnergyPlus Engineering Reference Book (US DOE 2020).

However, these are models that take a long time to generate, require a lot of experience to implement and have difficulties in adapting to socio-economic vicissitudes. It should not be forgotten that these models are composed of some three thousand adjustable parameters, and that the scientific community assures that the fewer parameters involved in the fitting process, the more robust the results obtained will be (Reddy 2006).

Following this assumption, the authors of this research developed a methodology for calibrating white box models using inverse modelling in its process, relying on public data generated by Annex 58 for its implementation.

Annex 58 describes an empirical validation study on two buildings in the German town of Horskilchen, as part of the AIE BCE project (Annex58 2016). The objective of Annex 58 was to test the correlation between the real data measured in the dwellings and the data generated by the energy models developed by the participants. 21 universities and research centres worked on the trial, creating their BEMs and providing the necessary data for the project (Strachan et al. 2016).

The requirements for approaching the experiment were made public and taking this opportunity the authors validated empirically and comparatively a new methodology for energy model adjustment (Gutiérrez González et al. 2020). This new approach achieves high quality results by reducing the number of sensors provided by the Annex 58 to be used in the process by more than 47%. In addition, a very significant reduction in the number of parameters involved in the process was implemented, leaving only 4: capacitances, infiltrations, internal mass and thermal bridges.

The calibration methodology implemented defines an objective function that will search for the best fit between real and simulated data using the uncertainty indices proposed in the experiment (MAE and ρ). Managing to generate robust and high quality models. For this purpose, the JePLUS package (JePLUS and JePLUS+EA) (Zhang and Korolija 2010) will be used, which allows the NSGA-II genetic algorithm (Deb et al. 2002) to be integrated with the EnergyPlus engine and thus used as a tool to find the best solution. The search space that the algorithm faces is defined by the possibility of combinations that can occur in the selected parameters.

Of the four parameters involved in the calibration process, some of them are difficult to quantify physically in a reliable way, such as capacitances and infiltrations. Seeing the difficulty in obtaining accurate data on these parameters, Hong and Lee presented a paper examining the EnergyPlus object: HybridModel:Zone (Hong and Lee 2019), which, through inverse modelling, inverts the heat equation of the air in the zone (US DOE 2016), using its internal temperature as input, to calculate analytically both the internal mass of the space (capacitances) and its infiltrations.

Lee and Hong (2019) validated the inverse modeling algorithm, which has the HybridModel:Zone object using the Facility for Low Energy Experiment in Buildings (FLEXLAB) at Lawrence Berkeley National Laboratory (LBNL). However, one of the limitations of this study was the fact that three days of free oscillation of the building were used instead of the seven necessaries for the optimal functioning of the algorithm. In addition, it was argued that if the measured energy in the time step of the HVAC systems were known, the inverse model could also be applied for those periods. But there was no empirical validation of this feature.

The main objective of this study is the implementation of the HybridModel:Zone object within the calibration

methodology developed by the authors of this research, as recommended by Lee and Hong (2019) in their work. Since they stated that the object could not replace an adjustment process, but rather should be used as a calibration aid. Thus, establishing a new technique for obtaining robust and quality calibrated models. Since the capacitance and infiltration values defined when activating the HybridModel:Zone object will be removed from the search space, thereby increasing the speed and accuracy of the process.

Due to the development of this new methodology, a second objective was implemented: the empirical validation of the object in all spaces in which it can be activated: load and free oscillation spaces.

Another novelty of the research is the development of a technique capable of transforming the specificity of the infiltrations produced by HybridModel:Zone into generic ones. This procedure is important because those calculated by the object are specific to the time period where it is activated and cannot be exported to others.

The organization of the document is structured as follows: Section 2 will describe the methodology used in the research. This part is made up of: Subsection 2.1, where a detailed of the dwellings used in the study will be made; Subsection 2.2 will provide how the energy model to be used has been generated; Subsection 2.3 will explain the time periods in which action will be taken; Subsection 2.4 will discuss the evaluation criteria used and Subsection 2.5 will summarize the methodology used by means of a flow chart. In Section 3, the results obtained in the different phases of the investigation will be presented, while Section 4 will show the conclusions reached.

2 Methodology

2.1 Description of the test site

To conduct this study, the real data generated by the experiment created to develop Annex 58 has been used and that, is freely available for any researcher to use in their investigations. This work has not been done specifically for the evaluation of the HybridModel:Zone object; however, the available data can be used for this purpose. The first step for the evaluation of the object was the development of the energy models of the N2 and O5 houses, built for the experiments. These houses are located south of Munich, in the Bavarian town of Holzkirchen. They are two single-family homes that share the morphology of a German detached house. And are separated by a few meters from each other without any element to give them shade (Figure 1).

Each house has 3 floors: basement, main and attic, with an approximate area of 100 m^2 per floor. The main floor consists of: living room; kitchen; children's room; bedroom; bathroom; entrance and corridor. Following the rules set out in Annex 58 for the BEM of dwellings, the rooms on the ground floor have been defined as separate thermal zones. The adjustment work could have been simplified by analyzing the whole ground floor as a single thermal zone. This option may be acceptable when the average temperature of the building is representative of that of the thermal zones. It would reduce costs, because it might only be necessary to account for an average temperature of the floor and the overall energy consumed by the dwelling, but even so, the precision in the adjustment with the measured data will be affected.

With regard to the construction systems used, it should be noted that the insulation of the exterior facades was made in accordance with the recommendations of the German energy code EnEV 2009. The window panes placed in both homes have a U of 1.1 W/(m^2·K) while the frames have 1.0 W/(m^2·K).

The houses are equipped with a mechanical ventilation system that introduces air into the living room for a value of 120 m^3/h. This air is exhausted through two outlets

Fig. 1 Situation of housing

located in the bathroom and in the children's room. Each room ejects 60 m^3/h.

The authors of the Annex 58 experiment equipped the rooms of the houses, except for the corridor, with fast-response electric heaters as a heating system. The baseboard heater model used was the Dimplex AKO K810/K811 with a radiative/convective split of 30%/70%.

Numerous sensors and measuring systems were installed in the dwellings to collect data from the building envelope, the HVAC systems, ventilation, etc. For the realization of this study, it was very important to have reliable data on the interior temperature of the analyzed thermal zones, since the HybridModel:Zone object needs to be fed with them. Therefore, this experiment is appropriate for this research, because each room was equipped with an indoor temperature sensor which was carefully installed. All of them were placed at a height of 125 cm from the floor and were protected against radiation that could be caused by adjacent surfaces or systems. In the living room, two extra interior temperature sensors were placed 67 and 187 cm above the floor separately. As with the other temperature sensors, these were also protected from possible radiation.

Table 1 makes a compilation of all of them showing the degree of precision.

To determine the infiltration of the houses, a blower door test was made on the main floor of both homes. The result showed that dwelling N2 had an infiltration value of n50 of 1.62 ac/h, while for the O5 house the value of n50 was 1.54 ac/h. These values were entered into the base model using the EnergyPlus Design Flow Rate object.

Table 1 Implemented sensors and their accuracy

	Sensor	Accuracy
Interior sensors	Air temperature in all rooms at height of 125 cm	±0.12 K
	Air temperature in living room at height of 67 cm and 187 cm	±0.14 K
	Air temperature in the basement and attic	±0.14 K
	Relative humidity in the living room	±2.30 %
	Fresh, supply and exhaust air temp. measured in the cellar	±0.04 K
	Heating power of the heated rooms	±1.50 %
	Supply and exhaust fan power	±1.50 %
	Ventilation flow rate	±3.50 m^3/h
	Heat flux at the west facade	±0.65 W/m^2
	West wall temp.: internal, external and between layers	±0.14 K
Weather station sensors	Dry bulb temperature	±0.10 K
	Ambient relative humidity	±2.00 %
	Ground temp. depth of 200, 100, 50, 0 cm	±0.10 K
	Wind speed at 10 m height	±0.10 m/s
	Wind direction at 10 m height	±1.00
	Global, diffuse and vertical solar radiation	± 2.00 %
	Long-wave radiation (horizontal, west vertical)	< 34.00 W/m^2

2.2 Development of the EnergyPlus model

Once the public data available in the Annex 58 had been analysed, the energy models were generated (Figure 2). In these, the thermal zones that make up the buildings were introduced and the materials of the different constructions of the houses were detailed. All HVAC systems were introduced into the BEM so that the model would reflect reality as closely as possible.

A weather station located nearby collected data on: dry bulb temperature; global and diffuse solar radiation; wind speed and direction and relative humidity. The data from these sensors was used to generate a valid weather file (.epw format) that could be entered into EnergyPlus.

One of the main difficulties faced by modellers when generating detailed energy models are the uncertainties that can be found in the building, which are very difficult to quantify (Heo et al. 2012; Hong et al. 2015; Lee et al. 2015a, b). If one refers to the physics of the building, there are many uncertainties such as: the physical characteristics of materials produced in-situ during the construction process; the characteristics and composition of the ground (Gutiérrez González et al. 2022) among others. Infiltration can be detected by means of a blower door, but this task, which is already costly in residential housing, is very difficult to manage in the case of tertiary buildings.

Some of the parameters that are difficult to quantify in BEMs are also: internal mass, generated by interior furnishings, partitions, etc. Simulation tools usually collect the mass of the envelope (Wang and Xu 2006; Zeng et al. 2011; Johra and Heiselberg 2017), but establishing the value of the internal mass of the building is a very complicated task and significantly affects the expected final result of an energy model.

Reducing this type of uncertainty, knowing the real value of capacities or infiltrations, is an important issue in an energy model. Thus reflecting the reality of the building and therefore better serving its final objective: saving energy consumption or improving user comfort (Ramos Ruiz and Fernández Bandera 2017a,b).

2.3 Experiment periods

The original Annex 58 experiment was conducted between August 21, and September 30 of 2013. This time interval was divided into 5 periods in which the houses were subjected to different energy conditions (Table 2):

- Period 1 (initialisation): During this three-day period, from 21 to 23 August, the houses were heated to a

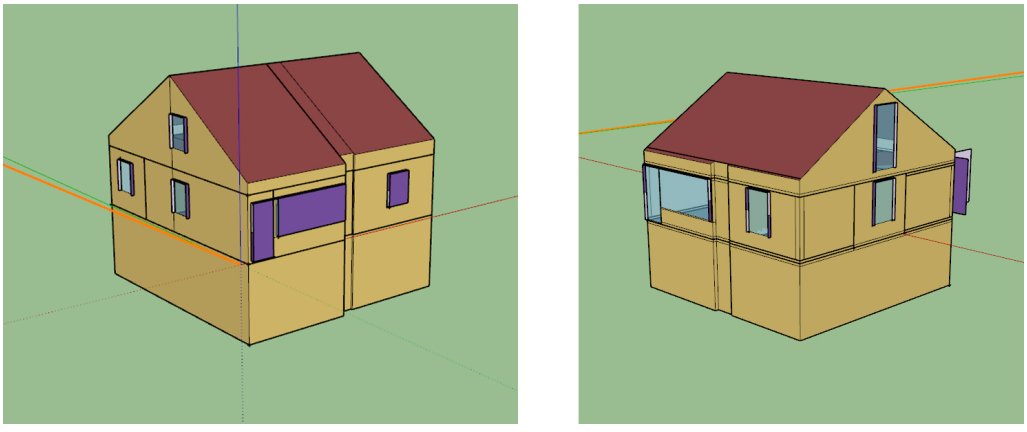

Fig. 2 Schematic view of the energy models (house O5 and N2 respectively) generated with the Open Studio plug-in for SketchUp

Table 2 Data periods, data provided and requested

Period	Date	Configuration	Data provided	Data requested
Period 1	2013/8/21 to 2013/8/23	Initialization (constant temperature)	Temperature and heat inputs	—
Period 2	2013/8/23 to 2013/8/30	Constant temperature (nominal 30 °C)	Temperature and heat inputs	Heat outputs
Period 3	2013/8/30 to 2013/9/14	ROLBS heat inputs in living room	Temperature and heat inputs	Temperature outputs
Period 4	2013/9/14 to 2013/9/20	Re-initialization constant temp. (nominal 25 °C)	Temperature and heat inputs	Heat outputs
Period 5	2013/9/20 to 2013/9/30	Free float	Temperature inputs	Temp. outputs

constant temperature of 30 °C, and thus obtain initial conditions with the same characteristics in both houses.

- Period 2 (set point 30 °C): This period goes from the 23rd to the 30th of August. The houses are kept at a constant temperature of 30 °C. The indoor temperature data can be entered into the energy model, while the energy required to reach these temperatures will be requested. This energy will be compared with what the real building consumes and thus be able to see the degree of adjustment that the model has with respect to reality.
- Period 3 (ROLBS): This period begins on 30 August and ends on 13 September. The heat input was introduced into the model by means of a randomly ordered logarithmic binary sequence (ROLBS). One of the objectives, when feeding the heat inputs through the ROLBS sequence, is to make sure that there is no relationship between the energy injections and the solar radiation. The energy contribution is the data provided to be entered into the model. The interior temperatures that the energy model was capable of producing with the given energy were requested to be contrasted with the real ones.
- Period 4 (set point 25 °C): period from 14 September to 19. It is a period very similar to the 2nd. But instead of heating the houses to 30 °C they are now warmed to 25 °C. The indoor temperature data is known and the energy that the BEM must consume to reach it is requested.
- Period 5 (free oscillation): the duration of this period is 10 days, from 20 to 30 September. The buildings are left in free oscillation, they will not have any energy contribution, except for what can come from the solar gains. The energy models will be asked for the internal temperatures of the thermal areas analysed and compared with the real ones. This will show the degree of adjustment that the BEM has with respect to reality.

The experiment was conducted between August and September when temperatures are not as extreme as in the winter period. To remedy this lack of a thermal gap between the indoor and outdoor temperatures in the houses, those responsible for the experiment in Annex 58 raised the thermostat setting of the heating systems to 25 and 30 °C in periods 2 and 4 respectively, well above the requirements of state regulations, which normally set it at 21 °C in the winter season. The aim is to simulate the possible temperature difference that could occur in the coldest months of the year.

One of the key elements to obtain optimal results using the calibration methodology on which this new approach is based, is to have a temperature difference between the outside and the inside of the building. Therefore, the same calibration methodology can be applied in both the winter and summer periods, obtaining the same degree of adjustment, as shown by Pachano et al. (2022), where the model developed using the methodology on which this study is based, generated one that was calibrated in the summer

period and analyzed in the winter months and vice versa.

In order to conduct this research, the HybridModel:Zone object was activated in periods 3 and 5, and thus be able to know the capacitances and infiltrations of the building. Once found, they will be introduced in the calibration process, generating the *Calibrated Hybrid model*. These models: *Base*, *Calibrated* and *Calibrated Hybrid* were tested in all the periods previously described.

In the fixed set-point periods (2 and 4) the object can not be activated, as the internal temperatures are practically constant, so the temperature differential between adjacent periods, which acts as a quotient in Eq. (1) for the calculation of the capacitance value, will be very close to 0. And for the equation to be solved optimally, dt has to be greater than 0.05 °C (US DOE 2020). This condition is necessary to avoid anomalies in the calculation of the internal mass of the thermal zone.

$$C_z \frac{\mathrm{d}T_z}{\mathrm{d}t} = \sum_{i=1}^{N_{si}} (\dot{Q}_i) + \sum_{i=1}^{N_{surfaces}} (h_i A_i (T_{si} - T_z)) + \sum_{i=1}^{N_{zones}} (\dot{m}_{i\_zone} C_{pi} (T_{zi} - T_z)) + \dot{m}_{inf} C_p (T_\infty - T_z) + \dot{Q}_{sys} \quad (1)$$

At the same time, following the recommendation given by Lee and Hong (2019), values of a thermal gap between the outside and inside temperature of a zone below 5 °C were discarded for the calculation of infiltration. To make the result more reliable.

2.4 Evaluation criteria

To check the degree of agreement obtained by the energy models, when comparing their data with the real ones, the same goodness-of-fit indices that were proposed by Annex 58 study have been used to evaluate their results: main absolute error (MAE) and Sperman's correlation coefficient (ρ) (Strachan et al. 2016). The MAE index will measure the magnitude of fit between the real data and those obtained by the simulated model and at the same time, the magnitude of fit of the shape will be dimensioned by the ρ index.

It has also been decided to classify the results obtained following the same evaluation as the one performed in the Annex 58 for both temperature and energy values. The calibration quality level is summarised in Table 3. Models highlighted in green represent those that achieve an excellent fit with the real data; yellow: models with a moderate adjustment; orange: represents an insufficient match and a model marked in red is considered not to have a good fit with the data obtained from the real building.

It is important to be able to define precisely which model most closely resembles the real data. For this reason, an evaluation criterion has been generated to show the degree of goodness of fit of the BEMs created. The uncertainty indices (MAE and ρ) obtained by all models and in all periods have been weighted equally, in order to compare them on equal terms. And thus, to find out which one fits best for the set of periods examined. Subsection 3.4 shows the results obtained.

2.5 Flow chart: summary of the methodology

The Figure 3 shows the flow chart followed in this study to achieve the results that will be discussed in Section 3. The research starts with the development of the *Baseline model* (1) using the Open Studio (Guglielmetti et al. 2011) plug-in for SkechUp (Ellis et al. 2008). The Open Studio software allows exporting the created energy model in .idf format, which is used by EnergyPlus. At the same time as the BEM is generated, the weather file (.epw) is created, consisting of the sensors that make up the meteorological station located next to the dwellings (2). With the BEM and the weather file created, the phases in which the experiment is to be carried out will be defined: Periods (3).

The base model is not subjected to any calibration process and will be simulated in all periods, without altering its characteristics, to be able to compare the degrees of adjustment achieved with the new models generated.

The proposed study tries to define the values of the capacitances and infiltrations of the buildings under consideration, to be introduced into the model and thus reduce the number of parameters when executing its calibration process. So once the BEM, the meteorological file and the simulation periods have been defined, the *Baseline Model* will be executed with the EnergyPlus HybridModel:Zone object activated to find these values by thermal zone and simulation period (4).

At the same time, the *Baseline Model* (1) will undergo an adjustment process following the methodology developed by the authors generating the *Calibrated Model* (5).

Table 3 Calibration criteria of Annex 58

MAE	Temp. magnitude fit	< 1 °C	1–2 °C	2–4 °C	> 4 °C
	Heat input magnitude fit	< 100 W	100–200 W	200–300 W	> 300 W
ρ	Temp./ heat input magnitude fit	> 90 %	80 %–90%	70 %–80%	< 70 %
	Calibration level	Excellent	Optimal	Medium	No calibrated

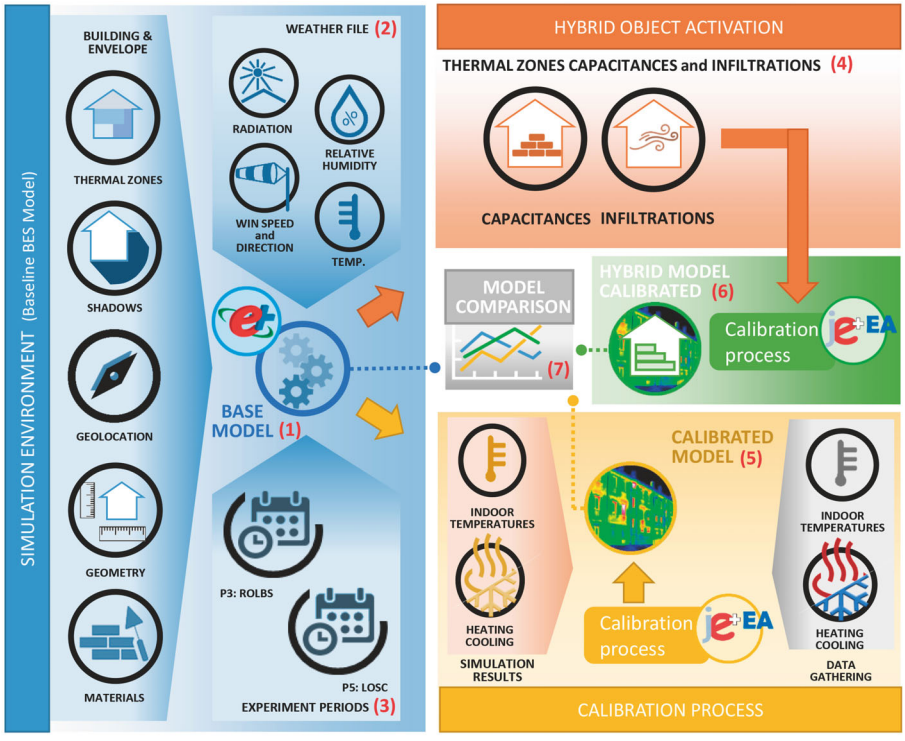

Fig. 3 Flowchart of the research performed

Finally, the results obtained of capacitances and infiltration by the execution of the object were introduced in the model that will be subjected to the calibration process, generating the *Hybrid Calibrated Model* (6).

The three models created: *Baseline Model* (1), *Calibrated Model* (5) and *Hybrid Calibrated Model* (6) were run in all the periods proposed for the Annex 58 experiment in view of analysing the results and comparing their degree of fit with the real data (7).

All these studies are conducted in an effort to reduce the adjustment times of the BEMs to reality and to increase the quality of the data entered into the model by obtaining values of capacitances and infiltrations through the EnergyPlus object.

3 Analysis of the results

3.1 Result of the capacitance and infiltration values obtained by the activation of the HybridModel:Zone object

To perform this study, the HybridModel:Zone object was activated in periods 3 (ROLBS) and 5 (free oscillation), in order to know the values of capacitances and infiltration that the BEM will provide for these periods in each thermal zone.

The obtained capacitance values (Table 4) were directly entered into the energy models.

A remarkable fact is that, for each analysed period different capacitance values are obtained. The value calculated by the HybridModel:Zone object depends on the degree of uncertainty of the energy model. Even so, the nearer the base model and the real building are, the closer the capacitance value generated by the object for each period analysed.

Analysing the capacitances found in the different thermal zones, it is observed that in all zones they are quite homogeneous. Their differences vary between 0.22 obtained in the kitchen of house O5 and 1.62 in the living room of house N2. The largest mismatch occurs in the bedroom of house N2 with 23% inequality.

When the object is activated for the infiltration calculation, the result is a time-step by time-step infiltration value for each thermal zone analysed. These values are specific to the time period under study and can only be used for that time frame.

In order to be able to use the infiltration values obtained in any time frame, a regression of the Design Flow Rate object of EnergyPlus was performed (Eq. (2)).

$$I=\left(I_{\text{design}}\right)\left(F_{\text{sch}}\right)\left[A+B\left|\left(T_{\text{zone}}-T_{\text{odb}}\right)\right|+C\left(WS\right)+D\left(WS^{2}\right)\right] \quad (2)$$

This regression allows the generation of a linear model in which the value of the dependent variable (infiltration) is determined from a set of independent variables (coefficients: *A*, *B*, *C* and *D*). The newly calculated coefficients (Table 5)

Table 4 Result of the value of the capacitance for the two houses and in all the analysed periods.

	Living room		Children room		Bedroom		Kitchen	
Period	N2	O5	N2	O5	N2	O5	N2	O5
ROLBS (Period 3)	11.21	10.12	4.89	3.93	5.41	6.05	13.95	5.23
Free oscillation (Period 5)	9.59	11.43	4.49	4.38	4.15	5.78	13.12	5.45

Table 5 Result of the values of the coefficients necessary for the calculation of the infiltration in the Design Flow Rate object of EnergyPlus

	Living room		Children room		Bedroom		Kitchen	
Coefficients	N2	O5	N2	O5	N2	O5	N2	O5
A	0.001573	0.000221	0.000295	0.003200	0.000297	0.001194	0.000168	0.000464
B	0	0.000136	0.000138	0	0.000068	0	0.000271	0
C	0.000045	0.000357	0.000468	0.000250	0.000367	0.000078	0.000141	0.000041
D	0	0	0	0	0	0	0.000012	0.000000

A: constant term; *B*: temperature term ($°C^{-1}$); *C*: velocity term (s/m); *D*: velocity squared term (s^2/m^2).

are fed into the energy model and can thus be used in any period.

Looking at the results that the EnergyPlus object shows in infiltration, it can be stated that it introduces more infiltration in the N2 dwelling than in the O5, where it practically cancels them in the bedroom and the kitchen.

The values of capacitances and infiltration shown in Tables 4 and 5 respectively were fed into the *Baseline Models*. This, in turn, was subjected to the calibration process developed by the authors of this study, resulting in the *Hybrid Calibrated Model*.

3.2 Comparison, using MAE index, of the adjustment results obtained by the energy models

Tables 6 and 7 present the results obtained in the checking periods 3 (ROLBS) and 5 (LOSC) for the models studied: *Baseline*, *Hybrid Calibrated* and *Calibrated Model*, using the MAE index by thermal zone. These tables compare the real temperature with that produced by the models, thus showing the degree of agreement.

In turn, Tables 8 and 9 show the MAE index per thermal zone of the 3 energy models generated when are checked in periods 2 and 4. In these, the energy that the model needs to reach certain temperatures is compared with the energy that it really consumes to reach them.

The first column of the tables indicates the dwelling in question: N2 or O5. The second represents the type of energy model. The third defines the training period, the period where the HybridModel:Zone object has been activated or the model has been calibrated, and the fourth shows the period where the results have been checked. The remaining columns show the MAE index values obtained by the different models in each thermal zone studied.

Table 6 Result of the adjustment with the real data (in °C) produced by the energy models of both houses (N2 and O5) trained in period 3 and checked in periods 3 and 5 (MAE index)

House	Model	Training period	Checking period	Living room	Children's room	Bedroom	Kitchen
N2	Calibrated	Period 3	Period 3	0.32	0.23	0.22	0.55
	Hybrid calibrated	Period 3	Period 3	0.31	0.26	0.34	0.70
	Baseline	Period 3	Period 3	1.55	0.25	0.50	0.97
N2	Calibrated	Period 3	Period 5	0.40	0.28	0.23	0.51
	Hybrid calibrated	Period 3	Period 5	0.36	0.32	0.38	0.65
	Baseline	Period 3	Period 5	1.82	0.23	0.33	0.66
O5	Calibrated	Period 3	Period 3	0.45	0.29	0.19	0.28
	Hybrid calibrated	Period 3	Period 3	0.55	0.45	0.16	0.28
	Baseline	Period 3	Period 3	1.56	0.41	0.26	0.64
O5	Calibrated	Period 3	Period 5	0.66	0.34	0.20	0.30
	Hybrid calibrated	Period 3	Period 5	0.64	0.13	0.18	0.25
	Baseline	Period 3	Period 5	1.93	0.41	0.28	0.67

Green < 1 °C | Yellow 1–2 °C | Orange 2–4 °C | Red > 4 °C

Table 7 Result of the adjustment with the real data (in °C) produced by the energy models of both houses (N2 and O5) trained in period 5 and checked in periods 3 and 5 (MAE index)

House	Model	Training period	Checking period	Living room	Children's room	Bedroom	Kitchen
N2	Calibrated	Period 5	Period 3	0.36	0.27	0.34	0.59
	Hybrid calibrated	Period 5	Period 3	0.41	0.25	0.33	0.69
	Baseline	Period 5	Period 3	1.55	0.25	0.50	0.97
N2	Calibrated	Period 5	Period 5	0.34	0.22	0.21	0.51
	Hybrid calibrated	Period 5	Period 5	0.30	0.27	0.34	0.65
	Baseline	Period 5	Period 5	1.82	0.23	0.33	0.66
O5	Calibrated	Period 5	Period 3	0.46	0.53	0.30	0.47
	Hybrid Calibrated	Period 5	Period 3	0.67	0.48	0.17	0.46
	Baseline	Period 5	Period 3	1.56	0.41	0.26	0.64
O5	Calibrated	Period 5	Period 5	0.39	0.13	0.18	0.27
	Hybrid Calibrated	Period 5	Period 5	0.44	0.12	0.19	0.28
	Baseline	Period 5	Period 5	1.93	0.41	0.28	0.67

Green < 1 °C	Yellow 1–2 °C	Orange 2–4 °C	Red > 4 °C

Table 8 Result of the adjustment with the real data (in W) produced by the energy models of both houses (N2 and O5) trained in period 3 and checked in periods 2 and 4 (MAE index)

House	Model	Training period	Checking period	Living room	Children's room	Bedroom	Kitchen
N2	Calibrated	Period 3	Period 2	153	140	33	74
	Hybrid calibrated	Period 3	Period 2	142	129	62	51
	Baseline	Period 3	Period 2	76	175	85	46
N2	Calibrated	Period 3	Period 4	318	153	46	80
	Hybrid calibrated	Period 3	Period 4	314	150	64	44
	Baseline	Period 3	Period 4	102	188	81	63
O5	Calibrated	Period 3	Period 2	197	122	63	66
	Hybrid calibrated	Period 3	Period 2	190	98	63	64
	Baseline	Period 3	Period 2	170	149	73	65
O5	Calibrated	Period 3	Period 4	128	70	52	42
	Hybrid calibrated	Period 3	Period 4	209	90	37	39
	Baseline	Period 3	Period 4	129	145	49	54

Green < 100 W	Yellow 100–200 W	Orange 200–300 W	Red > 300 W

Table 9 Result of the adjustment with the real data (in W) produced by the energy models of both houses (N2 and O5) trained in period 5 and checked in periods 2 and 4 (MAE index)

House	Model	Training period	Checking period	Living room	Children's room	Bedroom	Kitchen
N2	Calibrated	Period 5	Period 2	85	132	22	81
	Hybrid calibrated	Period 5	Period 2	71	120	55	51
	Baseline	Period 5	Period 2	76	175	85	46
N2	Calibrated	Period 5	Period 4	242	152	42	81
	Hybrid calibrated	Period 5	Period 4	214	141	59	45
	Baseline	Period 5	Period 4	102	188	81	63
O5	Calibrated	Period 5	Period 2	219	87	29	59
	Hybrid calibrated	Period 5	Period 2	231	89	50	32
	Baseline	Period 5	Period 2	170	149	73	65
O5	Calibrated	Period 5	Period 4	457	76	43	126
	Hybrid calibrated	Period 5	Period 4	311	89	24	59
	Baseline	Period 5	Period 4	129	145	49	54

Green < 100 W	Yellow 100–200 W	Orange 200–300 W	Red > 300 W

A closer look at the results obtained for the MAE index by the BEMs shows that there are two clearly differentiated blocks when the degree of adjustment achieved with respect to the real measurements is examined. The *Baseline Model* obtains a degree of adjustment clearly further from reality than the *Calibrated Model* and the *Hybrid Calibrated Model*.

When the models are checked in periods 3 and 5 (Tables 6 and 7), where the temperature of the thermal zones are compared with the real temperature, the *Calibrated models* clearly achieve the best results, especially the difference in the living room is striking. When looking at the results shown in tables 8 and 9, where the energy produced by the models is compared to the real data, the opposite is true. The *Baseline models* manage a slightly better fit, especially in the living room. This may be the result of a slight over-fitting occurred during the calibration process. Since periods 2 and 4 have not been part of the calibration process.

3.3 Comparison, using ρ index, of the adjustment results obtained by the energy models

The same analyses performed for the MAE index have been done for the ρ index indicating the equality of the shapes of two curves. In this case, the temperature and energy curves produced by the energy models are compared with the curves produced by the real measurements. Tables 10 and 11 show the results derived from all the models analysed. As with the MAE index, the training periods are 3 and 5, periods in which the hybrid model has been activated. In turn, the models have been checked in all periods: periods 2 and 4, where energy is evaluated and periods 3 and 5 where temperature is checked.

When analysing the ρ index obtained by the energy models, it can be seen that, in general, the results are quite good. There is very little difference between the fit that one or the other models achieve.

But a closer examination of the results shows that the

Table 10 Result of the adjustment with the real data produced by the energy models of N2 house, trained in periods 3 and 5 and checked in all of them (ρ index)

House	Model	Training period	Checking period	Living room	Children's room	Bedroom	Kitchen
N2	Calibrated	Period 3	Period 2	0.96	0.58	0.91	0.72
	Hybrid calibrated	Period 3	Period 2	0.96	0.53	0.91	0.86
	Baseline	Period 3	Period 2	0.97	0.56	0.92	0.87
N2	Calibrated	Period 3	Period 3	0.99	0.98	0.98	0.92
	Hybrid calibrated	Period 3	Period 3	0.99	0.97	0.98	0.96
	Baseline	Period 3	Period 3	0.97	0.99	0.97	0.97
N2	Calibrated	Period 3	Period 4	0.94	0.62	0.87	0.67
	Hybrid calibrated	Period 3	Period 4	0.94	0.57	0.87	0.68
	Baseline	Period 3	Period 4	0.93	0.59	0.92	0.87
N2	Calibrated	Period 3	Period 5	1.00	1.00	0.99	0.95
	Hybrid calibrated	Period 3	Period 5	1.00	1.00	0.99	0.93
	Baseline	Period 3	Period 5	0.92	1.00	0.98	0.96
N2	Calibrated	Period 5	Period 2	0.96	0.56	0.91	0.74
	Hybrid calibrated	Period 5	Period 2	0.96	0.53	0.91	0.86
	Baseline	Period 5	Period 2	0.97	0.56	0.92	0.87
N2	Calibrated	Period 5	Period 3	0.99	0.97	0.97	0.92
	Hybrid calibrated	Period 5	Period 3	0.99	0.98	0.98	0.96
	Baseline	Period 5	Period 3	0.97	0.99	0.97	0.97
N2	Calibrated	Period 5	Period 4	0.95	0.60	0.85	0.67
	Hybrid calibrated	Period 5	Period 4	0.94	0.57	0.87	0.68
	Baseline	Period 5	Period 4	0.93	0.59	0.92	0.87
N2	Calibrated	Period 5	Period 5	1.00	1.00	0.99	0.95
	Hybrid calibrated	Period 5	Period 5	1.00	1.00	0.98	0.93
	Baseline	Period 5	Period 5	0.92	1.00	0.98	0.96

Green >0.90	Yellow 0.80–0.90	Orange 0.70–0.80	Red < 0.70

Table 11 Result of the adjustment with the real data produced by the energy models of O5 house, trained in periods 3 and 5 and checked in all of them (ρ index)

House	Model	Training period	Checking period	Living room	Children's room	Bedroom	Kitchen
O5	Calibrated	Period 3	Period 2	0.84	0.79	0.75	0.83
	Hybrid calibrated	Period 3	Period 2	0.84	0.79	0.75	0.83
	Baseline	Period 3	Period 2	0.90	0.75	0.78	0.82
O5	Calibrated	Period 3	Period 3	0.99	0.99	0.99	0.99
	Hybrid calibrated	Period 3	Period 3	0.98	0.98	0.99	0.98
	Baseline	Period 3	Period 3	0.99	0.99	0.99	0.99
O5	Calibrated	Period 3	Period 4	0.96	0.29	0.97	0.95
	Hybrid calibrated	Period 3	Period 4	0.95	0.42	0.95	0.93
	Baseline	Period 3	Period 4	0.96	0.42	0.96	0.93
O5	Calibrated	Period 3	Period 5	0.98	1.00	0.99	0.99
	Hybrid calibrated	Period 3	Period 5	0.97	1.00	0.99	0.99
	Baseline	Period 3	Period 5	0.98	1.00	0.98	0.98
O5	Calibrated	Period 5	Period 2	0.89	0.76	0.84	0.87
	Hybrid calibrated	Period 5	Period 2	0.90	0.91	0.87	0.91
	Baseline	Period 5	Period 2	0.90	0.75	0.78	0.82
O5	Calibrated	Period 5	Period 3	0.99	0.99	0.98	0.96
	Hybrid calibrated	Period 5	Period 3	0.98	0.99	0.99	0.96
	Baseline	Period 5	Period 3	0.99	0.99	0.99	0.99
O5	Calibrated	Period 5	Period 4	0.96	0.37	0.82	0.91
	Hybrid calibrated	Period 5	Period 4	0.96	0.41	0.96	0.95
	Baseline	Period 5	Period 4	0.96	0.42	0.96	0.93
O5	Calibrated	Period 5	Period 5	0.98	1.00	0.99	0.99
	Hybrid calibrated	Period 5	Period 5	0.98	1.00	0.99	0.98
	Baseline	Period 5	Period 5	0.98	1.00	0.98	0.98

Green >0.90	Yellow 0.80–0.90	Orange 0.70–0.80	Red < 0.70

children's room for both houses is the one that achieves the worst adjustment in all the models and especially in the periods where energy is the data to be compared. This may be due to certain errors or uncertainties that might occur in the data provided or in the physics of the thermal zone. Since in the rest of the zones and thermal periods the degree of adjustment achieved by the new methodology is of great quality.

3.4 Selection of the energy model with the best global adjustment to the real data

Table 12 shows the sum of the weighted uncertainty indices obtained by the BEMs in all the periods and thermal zones analysed. In order to distinguish which model has the best overall performance with respect to reality. The model with the most similar behaviour to the physical building is the one that obtains the lowest value in the total weighted sum (highlighted in blue).

When studying the results, the model that achieves the best similarity with the real data is the *Hybrid Calibrated Model*, improving this fit by 22.9% compared to that achieved by the base model. The *Calibrated Model*, obtained with the methodology on which the one developed in this paper is based, only manages to improve the degree of fit with respect to the *Baseline* model by 15.6%. The new technique

Table 12 Ranking of the degree of fit of energy models by weighted MAE index

		N2 house		O5 house			
	Training:	Period 3	Period 5	Period 3	Period 5	Total	
Model	Checking:	All periods	All periods	All periods	All periods	Sum	%
Calibrated		35.9	35.8	34.0	33.9	139.8	15.6
Hybrid calibrated		34.5	32.4	31.4	29.3	127.6	22.9
Baseline		38.9	40.1	45.8	40.7	165.6	0.0

achieves a 7.3% improvement over its predecessor in terms of similarity to the real building.

3.5 Comparison of the data obtained by the generated energy models against the real ones by means of a time sequence

The following Figures 4 to 7 present the temperature and energy results obtained by the analyzed models in a time sequence. The results of the models: *Baseline*, *Hybrid Calibrated* and *Calibrated* are plotted together with the real data (dotted line) to see graphically how well they fit these data.

As has been seen and commented when analysing the results in Tables 6 to 11, the *Hybrid Calibrated Model* is the one that achieves the best fit with respect to the real data in the set of periods analysed. Figures 4 and 5 illustrate this graphically. These, also clearly depict the two blocks mentioned above. On one hand we have the *Calibrated and Hybrid Calibrated Models* and on the other are the rest of the models. Figures 6 or 7 are a good example of this.

The graphs corresponding to the rest of the periods analyzed can be found in the Appendix, which is available in the Electronic Supplementary Material (ESM) from the online version of this paper.

4 Conclusions

The main objective of this research has been the development and evaluation of a calibration methodology based on that conducted by the authors of this research. Introducing the novelty of the activation of the EnergyPlus HybridModel:Zone object in the process.

At the same time, an empirical validation of the object

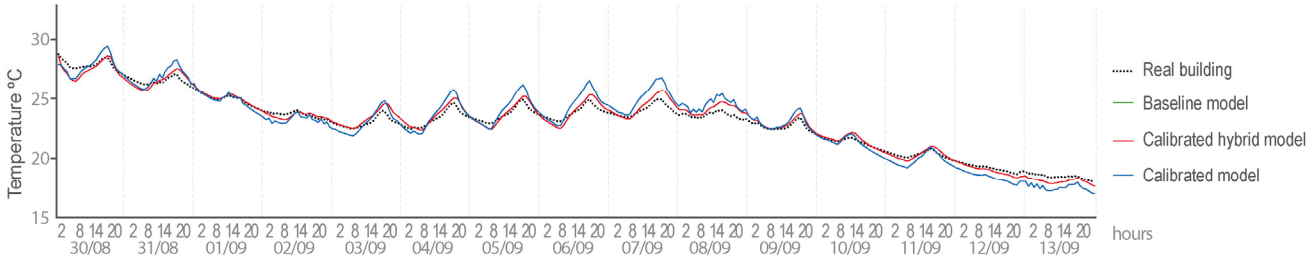

Fig. 4 Temperature (in °C) produced by the N2 house BEMs, trained in period 5 and checked in period 3

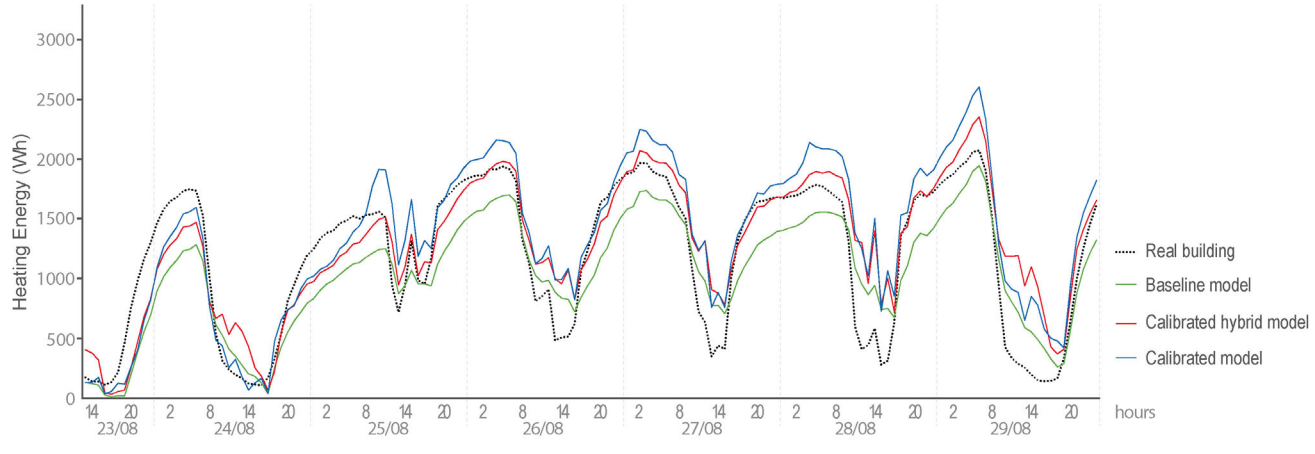

Fig. 5 Energy (in Wh) produced by the O5 house BEMs, trained in period 5 and checked in period 2

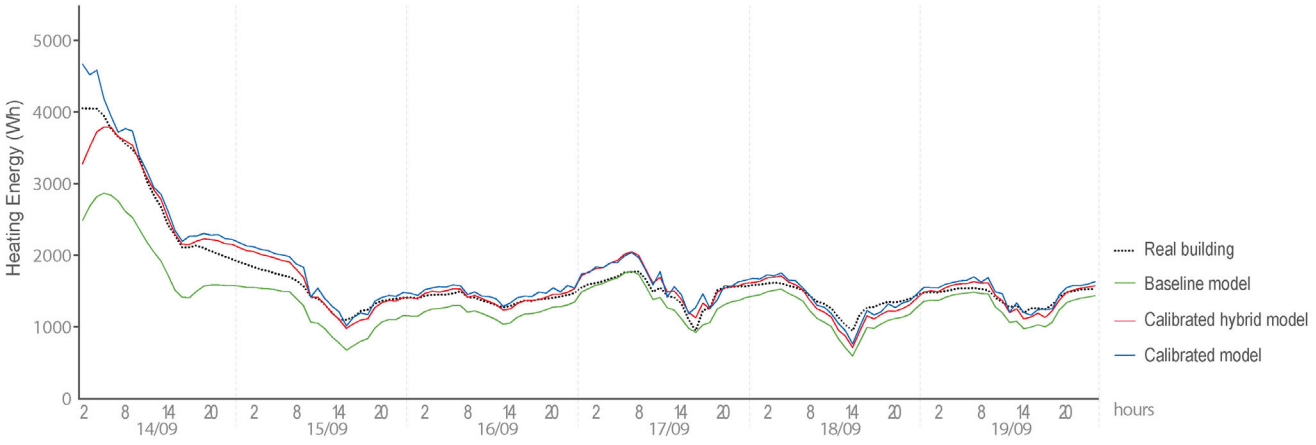

Fig. 6 Energy (in Wh) produced by the N2 house BEMs, trained in period 5 and checked in period 4

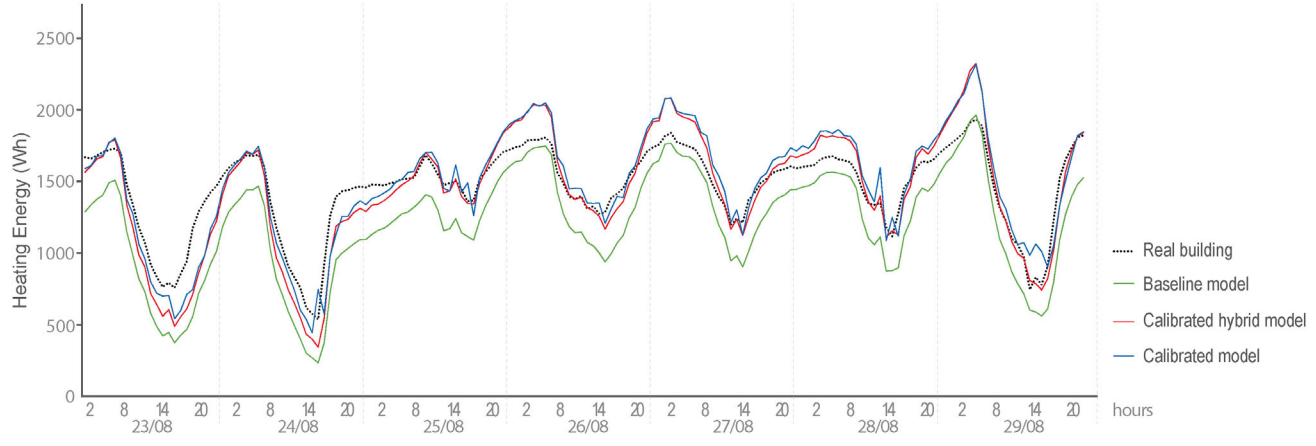

Fig. 7 Energy (in Wh) produced by the N2 house BEMs, trained in period 3 and checked in period 2

has been performed, testing it with real data and in all time periods where it can be activated.

To achieve these ends, the real data provided by Annex 58 of dwellings N2 and O5 in the German town of Holzkirchen have been used. These are reliable and robust data, made available to researchers for use in their studies.

The main objective achieved in this paper has been that the *Hybrid Calibrated model*, generated with the calibration methodology developed in this study, obtains the best degree of fit with respect to the real data than the rest of the models analysed. It improves by 22.9% that of the *Base model*, while the *Calibrated model* only does so by 15.6%.

Another noteworthy result is that the *Calibrated model* created with the method developed by the authors managed to improve the results obtained by the participants of the Annex 58 by an average of more than 42% in the degree of adjustment. And that this outcome has been surpassed by 7.2% by the proposed new model: *Hybrid Calibrated*.

It is also worth noting that the *Hybrid Calibrated model* adapts better to the checking periods than the others, reducing the risk of over-fitting that can occur in the process.

The results obtained are very encouraging because this calibration methodology is the one that provides the closest models to the real building. In future work it will be necessary to continue exploring this line of research in different types of buildings and locations, to consolidate the results obtained.

Electronic Supplementary Material (ESM): Complementary data on the results obtained by the four energy models analysed in different time periods, shown in time sequence are available in the online version of this article at https://doi.org/10.1007/s12273-022-0900-5.

Acknowledgements

This research was funded by the Government of Navarra under the project “From BIM to BEM: B&B” (ref. 0011-1365-2020-000227). We would like to thank the promoters of Annex 58, the possibility of accessing the data of the houses. Without this data, it would not have been possible to complete this work.

Funding note: Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

References

Al-Homoud MS (2001). Computer-aided building energy analysis techniques. *Building and Environment*, 36: 421–433.

Annex58 (2016). Annex 58: Reliable building energy performance characterisation based on full scale dynamic measurement. Available at https://bwk.kuleuven.be/bwf/projects/annex58/index.htm. Accessed 10-06-2020.

Ascione F, Bianco N, Iovane T, et al. (2020). A real industrial building: Modeling, calibration and Pareto optimization of energy retrofit. *Journal of Building Engineering*, 29: 101186.

ASHRAE (2013). ASHRAE Handbook: Fundamentals. Atlanta, GA, USA: The American Society of Heating, Refrigerating and Air-Conditioning Engineers.

Burch J, Subbarao K, Lekov A, et al. (1990). Short-term energy monitoring in a large commercial building. *ASHRAE Transactions*, 96(1): 1459–1477, 1990.

Chen J, Gao X, Hu Y, et al. (2019). A meta-model-based optimization approach for fast and reliable calibration of building energy models. *Energy*, 188: 116046.

Crawley DB, Hand JW, Kummert M, et al. (2008). Contrasting the capabilities of building energy performance simulation programs. *Building and Environment*, 43: 661–673.

Deb K, Pratap A, Agarwal S, et al. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6: 182–197.

Dhar A, Reddy TA, Claridge DE (1999). Generalization of the Fourier series approach to model hourly energy use in commercial buildings. *Journal of Solar Energy Engineering*, 121: 54–62.

Dong B, Cao C, Lee SE (2005). Applying support vector machines to predict building energy consumption in tropical region. *Energy and Buildings*, 37: 545–553.

Ellis PG, Torcellini PA, Crawley DB (2008). Energy design plugin: An EnergyPlus plugin for sketchup. Technical report, National Renewable Energy Lab. (NREL), Golden, CO, USA.

Fan C, Xiao F, Wang S (2014). Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques. *Applied Energy*, 127: 1–10.

Fernández Bandera C, Muñoz Mardones A, Du H, et al. (2018). Exergy as a measure of sustainable retrofitting of buildings. *Energies*, 11: 3139.

Foucquier A, Robert S, Suard F, et al. (2013). State of the art in building modelling and energy performances prediction: A review. *Renewable and Sustainable Energy Reviews*, 23: 272–288.

Giuliani M, Henze GP, Florita AR (2016). Modelling and calibration of a high-mass historic building for reducing the prebound effect in energy assessment. *Energy and Buildings*, 116: 434–448.

Guglielmetti R, Macumber D, Long N (2011). OpenStudio: An open source integrated analysis platform. Technical report, National Renewable Energy Lab. (NREL), Golden, CO, USA.

Gutiérrez González V, Álvarez Colmenares L, López Fidalgo JF, et al (2019). Uncertainty's indices assessment for calibrated energy models. *Energies*, 12(11): 2096.

Gutiérrez González V, Ramos Ruiz G, Fernández Bandera C (2020). Empirical and comparative validation for a building energy model calibration methodology. *Sensors*, 20: 5003.

Gutiérrez González V, Ramos Ruiz G, Fernández Bandera C (2022). Ground characterization of building energy models. *Energy and Buildings*, 254: 111565.

Heo Y, Choudhary R, Augenbroe GA (2012). Calibration of building energy models for retrofit analysis under uncertainty. *Energy and Buildings*, 47: 550–560.

Heo Y, Zavala VM (2012). Gaussian process modeling for measurement and verification of building energy savings. *Energy and Buildings*, 53: 7–18.

Hong T, Piette MA, Chen Y, et al. (2015). Commercial Building Energy Saver: An energy retrofit analysis toolkit. *Applied Energy*, 159: 298–309.

Hong T, Lee SH (2019). Integrating physics-based models with sensor data: An inverse modeling approach. *Building and Environment*, 154: 23–31.

Johra H, Heiselberg P (2017). Influence of internal thermal mass on the indoor thermal dynamics and integration of phase change materials in furniture for building energy storage: A review. *Renewable and Sustainable Energy Reviews*, 69: 19–32.

Kwok SSK, Yuen RKK, Lee EWM (2011). An intelligent approach to assessing the effect of building occupancy on building cooling load prediction. *Building and Environment*, 46: 1681–1690.

Lee SH, Hong T, Piette MA, et al. (2015a). Accelerating the energy retrofit of commercial buildings using a database of energy efficiency performance. *Energy*, 90: 738–747.

Lee SH, Hong T, Piette MA, et al. (2015b). Energy retrofit analysis toolkits for commercial buildings: A review. *Energy*, 89: 1087–1100.

Lee SH, Hong T (2019). Validation of an inverse model of zone air heat balance. *Building and Environment*, 161: 106232.

Martínez S, Eguía P, Granada E, et al. (2020). A performance comparison of multi-objective optimization-based approaches for calibrating white-box building energy models. *Energy and Buildings*, 216: 109942.

Nagpal S, Hanson J, Reinhart C (2019). A framework for using calibrated campus-wide building energy models for continuous planning and greenhouse gas emissions reduction tracking. *Applied Energy*, 241: 82–97.

Pachano JE, Peppas A, Bandera CF (2022). Seasonal adaptation of VRF HVAC model calibration process to a Mediterranean climate. *Energy and Buildings*, 261: 111941.

Ramos Ruiz G, Fernández Bandera C, Gómez-Acebo Temes T, et al. (2016). Genetic algorithm for building envelope calibration. *Applied Energy*, 168: 691–705.

Ramos Ruiz G, Fernández Bandera C (2017a). Analysis of uncertainty indices used for building envelope calibration. *Applied Energy*, 185: 82–94.

Ramos Ruiz G, Fernández Bandera C (2017b). Validation of calibrated energy models: Common errors. *Energies*, 10: 1587.

Reddy TA (2006). Literature review on calibration of building energy simulation programs:Uuses, problems, procedures, uncertainty, and tools. *ASHRAE Transactions*, 112(1): 226–240.

Robertson JJ, Polly BJ, Collis JM (2015). Reduced-order modeling and simulated annealing optimization for efficient residential building utility bill calibration. *Applied Energy*, 148: 169–177.

Robinson C, Dilkina B, Hubbs J, et al. (2017). Machine learning approaches for estimating commercial building energy consumption. *Applied Energy*, 208: 889–904.

Royapoor M, Roskilly T (2015). Building model calibration using energy and environmental data. *Energy and Buildings*, 94: 109–120.

Sadaei HJ, de Lima e Silva PC, Guimarães FG, et al. (2019). Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series. *Energy*, 175: 365–377.

Siddharth V, Ramakrishna PV, Geetha T, et al. (2011). Automatic generation of energy conservation measures in buildings using genetic algorithms. *Energy and Buildings*, 43: 2718–2726.

Soebarto V (1997). Calibration of hourly energy simulations using hourly monitored data and monthly utility records for two case study buildings. In: Proceedings of the 5th International IBPSA Building Simulation Conference, Madison, WI, USA.

Strachan P, Svehla K, Heusler I, et al. (2016). Whole model empirical validation on a full-scale building. *Journal of Building Performance Simulation*, 9: 331–350.

Subbarao K (1988). PSTAR: Primary and secondary terms analysis and renormalization: A unified approach to building energy simulations and short-term monitoring. Technical report, Solar Energy Research Inst., Golden, CO, USA.

Subbarao K, Balcomb JD, Burch JD, et al. (1990). Short-term energy monitoring: Summary of results from four houses. *ASHRAE Transactions*, 96(1): 1478–1483.

Tahmasebi F, Mahdavi A (2012). Monitoring-based optimization-assisted calibration of the thermal performance model of an office building. In: Proceedings of International Conference on Architecture and Urban Design, Tirana, Albania.

Thomas Ng S, Skitmore M, Wong KF (2008). Using genetic algorithms and linear regression analysis for private housing demand forecast. *Building and Environment*, 43: 1171–1184.

Tüysüz F, Sözer H (2020). Calibrating the building energy model with the short term monitored data A case study of a large-scale residential building. *Energy and Buildings*, 224: 110207.

US DOE (2016). EnergyPlus Engineering reference—EnergyPlus 8.5. The Reference to EnergyPlus Calculation.

US DOE (2020). EnergyPlus Engineering Reference. The reference to EnergyPlus calculations.

Wang S, Xu X (2006). Parameter estimation of internal thermal mass of building dynamic models using genetic algorithm. *Energy Conversion and Management*, 47: 1927–1941.

Wang Z, Wang Y, Srinivasan RS (2018). A novel ensemble learning approach to support building energy use prediction. *Energy and Buildings*, 159: 109–122.

Yang Z, Becerik-Gerber B (2015). A model calibration framework for simultaneous multi-level building energy simulation. *Applied Energy*, 149: 415–431.

Yu Z, Haghighat F, Fung BCM, et al. (2010). A decision tree method for building energy demand modeling. *Energy and Buildings*, 42: 1637–1646.

Yuan P, Duanmu L, Wang Z (2019). Coal consumption prediction model of space heating with feature selection for rural residences in severe cold area in China. *Sustainable Cities and Society*, 50: 101643.

Zeng R, Wang X, Di H, et al. (2011). New concepts and approach for developing energy efficient buildings: Ideal specific heat for building internal thermal mass. *Energy and Buildings*, 43: 1081–1090.

Zhang Y, Korolija I (2010). Performing complex parametric simulations with jEPlus. In: Proceedings of the 9th International Conference on Sustainable Energy Technologies (SET2010), Shanghai, China.

Zhang Y, O'Neill Z, Wagner T, et al. (2013). An inverse model with uncertainty quantification to estimate the energy performance of an office building. In: Proceedings of 13th International International IBPSA Building Simulation Conference, Chambery, France.