Ref. Ares(2022)5951056 - 26/08/2022

BIM SPEED

A set of calibrated BEMs for real demonstration cases and proposed standardisation

Deliverable Other 3.4

Deliverable Report: D3.4, issue date on 28.08.2022

BIM-SPEED
Harmonised Building Information Speedway for Energy-Efficient Renovation

This research project has received funding from the European Union's Programme H2020-NMBP-EEB-2018 under Grant Agreement no 820553.

A set of calibrated BEMs for real demonstration cases and proposed standardisation

Deliverable 3.4

Issue Date	23/08/2022
Produced by	Università Politecnica delle Marche
Main author	Gianluca Maracchini (UNIVPM)
Co-authors	UNIVPM
Version:	V 03
Type:	Other
Reviewed by	Timo Hartmann (TUB), Agnieszka Łukaszewska (FAS)
Approved by	Timo Hartmann (TUB)
Dissemination	Public

Colophon

This research project has received funding from the European Union’s Programme H2020-NMBP-EEB-2018 under Grant Agreement no 820553.

Change log

Description	Author	Date
Initial version for internal review	Gianluca Maracchini (UNIVPM)	20.05.2022
Addressed internal review comments	Gianluca Maracchini (UNIVPM)	02.06.2022
Finale editorial review	Timo Hartmann (TUB), Agnieszka Łukaszewska (FAS)	01.07.2022
Finalization	Gianluca Maracchini (UNIVPM)	23.08.2022
Final editorial check	Samaneh Rezvani (DMO)	23.08.2022

Publishable executive summary

In the European Union, 40% of the overall EU energy consumption (EC) and about 35% of the total greenhouse gas (GHG) emissions are attributable to the building sector. This is mainly due to the low energy performance of most of the building stock [1–3], which uses half of this energy for heating the households [4]. Since in the next 10 years the energy demand is expected to increase by more than 20% [5], improving the energy performance of these buildings represents an urgent need and opportunity to significantly reduce the European EC and GHG emissions, and then to reach the European sustainability and energy efficiency targets.

Fully detailed and dynamic building energy models (BEM) are increasingly being used throughout the building's lifecycle to calculate the building's energy performance and occupants' thermal comfort considering different weather conditions, building geometry, internal loads, HVAC systems, and operational schedules. In energy renovation of existing buildings, BEMs are also used to identify the best energy retrofit strategy among different available options in terms of energy-saving and/or comfort conditions, and to verify its compliance with the requirements set by the National Standards.

However, a significant discrepancy called the "energy performance gap", is often found between simulated and measured energy use, reaching values up to 250%. This discrepancy is becoming more and more evident with the rapid deployment of smart energy meters and the internet of things (IoT) and can be traced back to the difficulty in obtaining the exact values of all the thousands of inputs needed for characterizing a BEM. A BEM with inaccurate input data and/or inaccurate energy predictions may lead to the design of erroneous Energy Conservation Measures (ECM). For example, if a coefficient of performance (COP) of the heating system lower than the actual one is set in the model, an ECM concerning the retrofit of the HVAC might be recommended. However, if put in place, this ECM will likely provide lackluster results in terms of energy-saving.

To minimize this risk and, then, better design ECMs, it is of paramount importance to increase the accuracy and reliability of BEM. At this aim, a BEM calibration is generally undertaken, consisting of fine-tuning model input parameters to minimize the discrepancy between simulated and measured data. However, to date, there is still no universal consensus on which is the best calibration procedure to be used. Indeed, while there are standard criteria for validating a calibrated model, there is still a lack of formal and recognized methodology or guidelines for BEM calibration, which makes the BEM calibration processes highly dependent on the user's skills and judgments.

To overcome this issue, based on a literature review on BEM calibration and BEM calibration Standards, a state-of-the-art automated BEM calibration procedure has been developed and presented, aimed at facilitating the use of BEM calibration in engineering practice by minimizing the number of inputs required from the practitioners and allowing to reduce the dependency of its efficiency from the expertise of the energy modeler through its automatization. In particular, the procedure assists the energy modelers over the entire calibration process, from the data gathering to the BEM optimization process, passing through model enrichment and sensitivity analysis technique. It consists of two main phases: a data-gathering, which is aimed at obtaining the model inputs (weather conditions, schedule information, etc.) and outputs (e.g., energy consumption) required for the calibration process; and an automated calibration phase, aimed at speeding up and simplifying the calibration process for the end-user, and at increasing the predictive accuracy of the calibrated model with respect to manual approaches.

To allow the easy application of automated calibration, a specific software tool (BEM-Calibration Tool) has been developed. This tool integrates, for the first time in the literature, expert knowledge, sensitivity analysis, and Artificial Intelligence (AI) optimization algorithms within the same software workflow, minimizing the number of inputs required from the practitioners in BEM calibration processes while maintaining an easy-to-use interface. The procedure has been applied to BIMSPEEED demo cases to produce calibrated BEMs.

List of acronyms and abbreviations

AI: Artificial Intelligence

BEM: Building Energy Model

BES: Building Energy Simulation

BIM: Building Information Model

BC: BEM Calibration

CV(RMSE): Coefficient of Variation of Root Mean Square Error

DHW: Domestic Hot Water

DoA: Description of Actions

EC: Evolutionary Computation

ECM: Energy Conservation Measure

EPW: EnergyPlus Weather File

GA: Genetic Algorithm

GPS-HJ: Generalized Pattern Search Hooke-Jeeves

GUI: Graphical User Interface

HVAC: Heating, Ventilation and Air Conditioning

IDF: EnergyPlus Input Data File

IoT: Internet of Things

NMBE: Normalized Mean Bias Error

NSGA-II: Non-dominated Sorting Genetic Algorithm-II

NVB: Naturally Ventilated Building

PSO: Particle Swarm Optimization

RMSE: Root Mean Square Error

SA: Sensitivity Analysis

SHGC: Solar Heat Gain Coefficients

CoP: Heating/Cooling system efficiency or Coefficient of Performance

SPEA: strength Pareto Evolutionary Algorithm

Definitions

Not applicable.

Contents

Table of Figures

Table of Tables

1. Introduction

1.1 Description of the deliverable content and purpose

The main contents of this report are:

- a brief state-of-the-art on Building Energy Model (BEM) calibration (Section 2);
- the description of the BEM calibration (BC) procedure compliant with BIM-to-BEM standards (Section 3);
- the description of the BEM-Calibration software tool to be used in the BC procedure for the automatic calibration, integrated into the BIM-SPEED platform as a third-party software (Section 4);
- an exemplary application of the BC procedure and tool on a BIM-SPEED demo case (Section 5).

1.2 Contributions of partners

All the partners assigned to Task 3.4 "A set of calibrated BEM for real demonstration cases and proposed standardisation" have contributed to the development of either the procedure or the exemplary demo case. More specifically, partners' contributions to the deliverable (type other) are specified in Table 1.

Table 1: Contribution of partners.

#	Partner	Contribution
1	UNIVPM	Leader of the deliverable; coordination of the overall task activity; general state-of-the-art overview; definitions of recommendations on BEM creation and data collection recommendations based on state-of-the-art; development of the calibration procedure and related automatic tool; data analysis of the Vitoria/Aldabe Demo case; Vitoria/Aldabe BEM calibration; report writing.
2	CSTB	Integration of the tool within the KROQI platform (third-party service card)
3	CARTIF	Indoor monitoring campaign of the Vitoria/Aldabe demo case for the exemplary application
4	CYPE	BEM creation of the Vitoria/Aldabe demo case from BIM for the exemplary application
5	STRESS	BEM corrections and adjustments for the exemplary application
6	MTB	Support for the use of optimization algorithms

1.3 Target group and relation to other activities

The main target group of this deliverable consists of designers and energy modelers that directly use energy simulations in building renovation projects. Task 3.4 "A set of calibrated BEM for real demonstration cases and proposed standardisation" is directly linked to other activities within the Energy Cluster, providing outputs useful for both WP4 "Conducting performance simulations of renovation scenarios" and WP8 "Demonstrating best practices of BIM for renovation" activities. Besides, it provides information for WP1 "Collecting and understanding BIM data from existing buildings" in the definition of which information must be collected about existing buildings to enable a correct calibration of BEMs. Figure 1 reports the scheme of interconnections of the whole Energy Cluster in which Task 3.4 "A set of calibrated BEM for real demonstration cases and proposed standardisation" represents the last step. Moreover, the developed calibration procedure and exemplary application should be considered as a lesson learned from BIMSPEED case studies and state-of-the-art to be used as input for standardization in Task 5.1 "Cooperation with standardisation bodies".

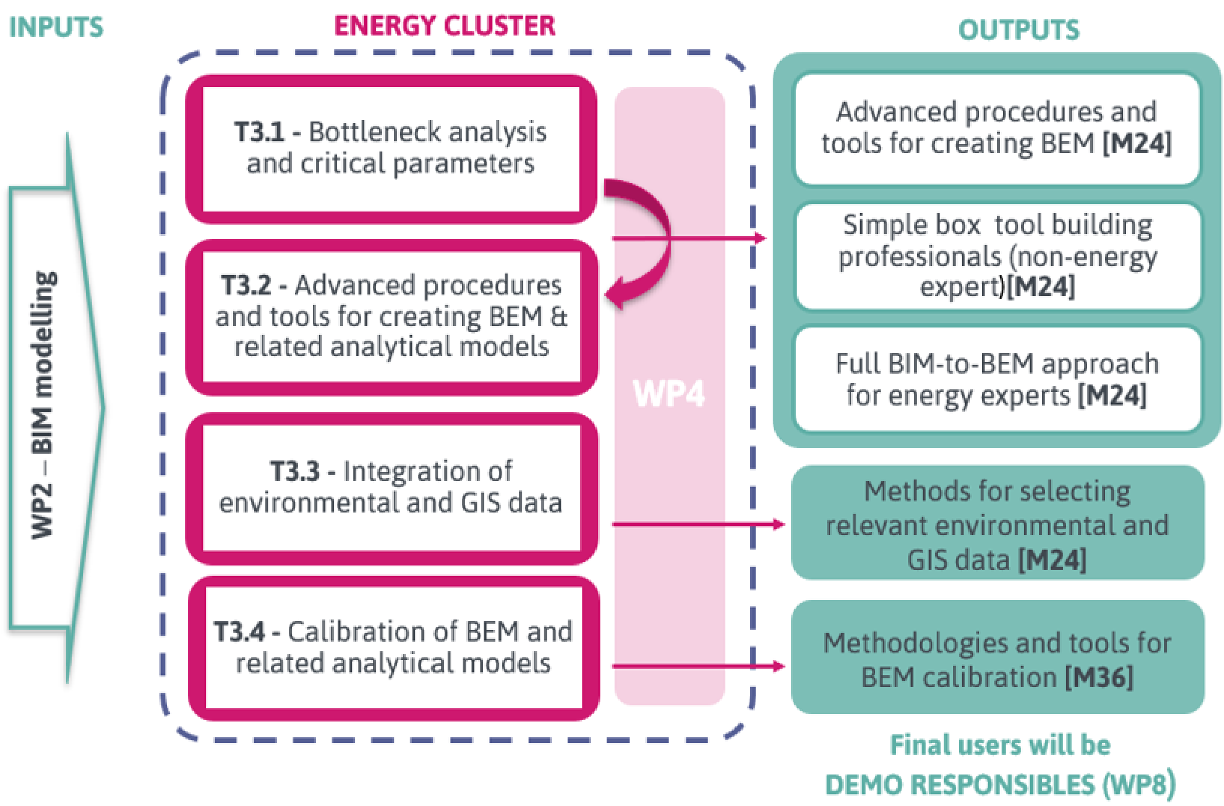


Figure 1: Scheme of Energy Cluster relations to other activities (from D3.1).

1.4 Relation with BIMSPEED Use Cases

In Table 2, the BIMSPEED Use Cases (UC) defined in D4.1 "Baseline and Use Cases for BIM-based renovation projects and KPIs for EEB renovation" that can benefit from UC12 "Calibration of the building energy model" are reported. In general, the UCs that use BEM simulation outputs to calculate specific KPIs will benefit from a calibrated BEM due to the importance in the decision processes of simulation results accuracy.

Table 2. List of the Use Cases developed in the BIM-SPEED project that will benefit from BEM calibration (UC12).

ID	USE CASE	GOAL
UC2	Assessing the Energy Performance of buildings with simulated data	Providing the building energy performance assessment using simulated data.
UC4	Assessing the as-built Thermal Comfort with simulated data	Assessing the occupants' thermal comfort before and/or after the renovation using simulated data.
UC8	Assessing the as-built Indoor Air Quality with simulated data	Assessing the level of indoor air quality before and/or after renovation by making use of simulations.
UC13	Optimization procedure for selecting the best EEB renovation scenario	Assessing the best renovation scenario for energy, cost, and comfort criteria from a pool of simulated alternatives
UC16	Assessing operational energy cost and payback (simulated)	Calculating the operational energy cost and renovation payback using a simulated approach
UC17	Assessing operational energy cost post-renovation and actual payback	Calculating the actual operational energy cost and renovation payback.
UC18	Assessing the fuel poverty condition	Assessing the fuel poverty indicator.
UC19	Assessing the actual energy savings	Assessing the actual energy savings by comparing pre and post-renovation data measured in the building

2. State-of-the-art on BEM calibration procedures and tools

2.1 BEM Calibration in building renovation projects

Fully detailed and dynamic building energy simulations (BES) are increasingly being used throughout the building's lifecycle to calculate the building's energy performance and occupants' thermal comfort considering different weather conditions, building geometry, internal loads, HVAC systems, and operational schedules. In energy renovation of existing buildings, BEMs are also used to identify the best energy retrofit strategy among different available options in terms of energy-saving and/or comfort conditions, and to verify its compliance with the requirements set by the National Standards.

However, a significant discrepancy is often found between simulated and measured energy use, reaching values up to 250%, which is becoming more and more evident with the rapid deployment of smart energy meters and the internet of things (IoT) [6]. The main causes of this discrepancy can be traced back to:

a) the lack of information to characterize the building energy model (BEM), which generates uncertainty also in terms of actual use and operation of the building systems;
b) the scenario uncertainties, such as those related to weather conditions and building surroundings;
c) model inadequacy arising from simplifications and abstractions of actual physical building systems.

In building renovation projects, a BEM with inaccurate input data and/or inaccurate energy predictions may lead to the design of erroneous Energy Conservation Measures (ECM). For example, if a coefficient of performance (COP) of the heating system lower than the actual one is set in the model, an ECM concerning the retrofit of the HVAC might be recommended. However, if put in place, this ECM will likely provide lackluster results in terms of energy-saving.

To minimize this risk and then to better design ECMs, it is of paramount importance to increase the accuracy and reliability of numerical results. At this aim, a BEM calibration (BC) is generally undertaken, consisting of fine-tuning model input parameters to minimize the discrepancy between simulated and measured data [7]. Despite uncalibrated models can still be useful for comparative analysis, the International Energy Agency's Energy in Buildings and Communities (IEA-EBC) Annex 53 highlighted the importance of calibrated models in BES for the correct design of the interventions and the accurate energy-saving estimation (and other energy performance indicators) achievable through the energy conservation measures (ECM) [8].

Over the past two decades, several studies and articles reviews on BC have been published in the literature [7,9–11]. However, to date, there is still no universal consensus on which is the best calibration procedure to be used. While there are standard criteria for validating a calibrated model, indeed, there is a lack of formal and recognized methodology or guidelines for BC, which makes the BC processes highly dependent on the user's skills and judgments. This notwithstanding, a common workflow of a state-of-the-art BC procedure can be identified, which can be subdivided into the following three main phases (Figure 2):

- a **data-gathering phase**, where information about observable model inputs (weather data, building operation information and occupation, building characteristics, etc.) and outputs (e.g. building energy performance, indoor thermal environment, etc.) is gathered through a detailed audit and/or a monitoring campaigns. The BEM model is then enriched based on the available input information and simplified or adapted based on available output information;
- a **model inputs screening phase**, where the dimensionality of the search space of the calibration problem is reduced to speed up the calibration process. This is made by identifying unobserved inputs that do not substantially affect the observed output(s) and that can be neglected in the calibration phase. Not influential parameters can be identified based on expert judgment or sensitivity analysis (SA) techniques, or both. Concerning SA, this can be also useful to determine the most important input parameters that need further investigation and, then, higher characterization efforts in the data-gathering phase;
- a **calibration phase**, carried out through manual or automated approaches (optimization-aided or Bayesian), to find the set of input values for the unobserved parameters identified in the screening phase, which minimizes the error between simulated and measured data.

In the following subsections, the techniques adopted in the literature to carry out these three main phases are briefly summarized and discussed. Section 2.2 briefly describes the main analytical techniques adopted in the first two phases, i.e. useful for data gathering, BEM characterization, and input screening. Section 2.3 discusses the main approaches adopted for the calibration phase. Finally, Section 2.4 reports a brief discussion on the software tools for automated BC available in the literature.

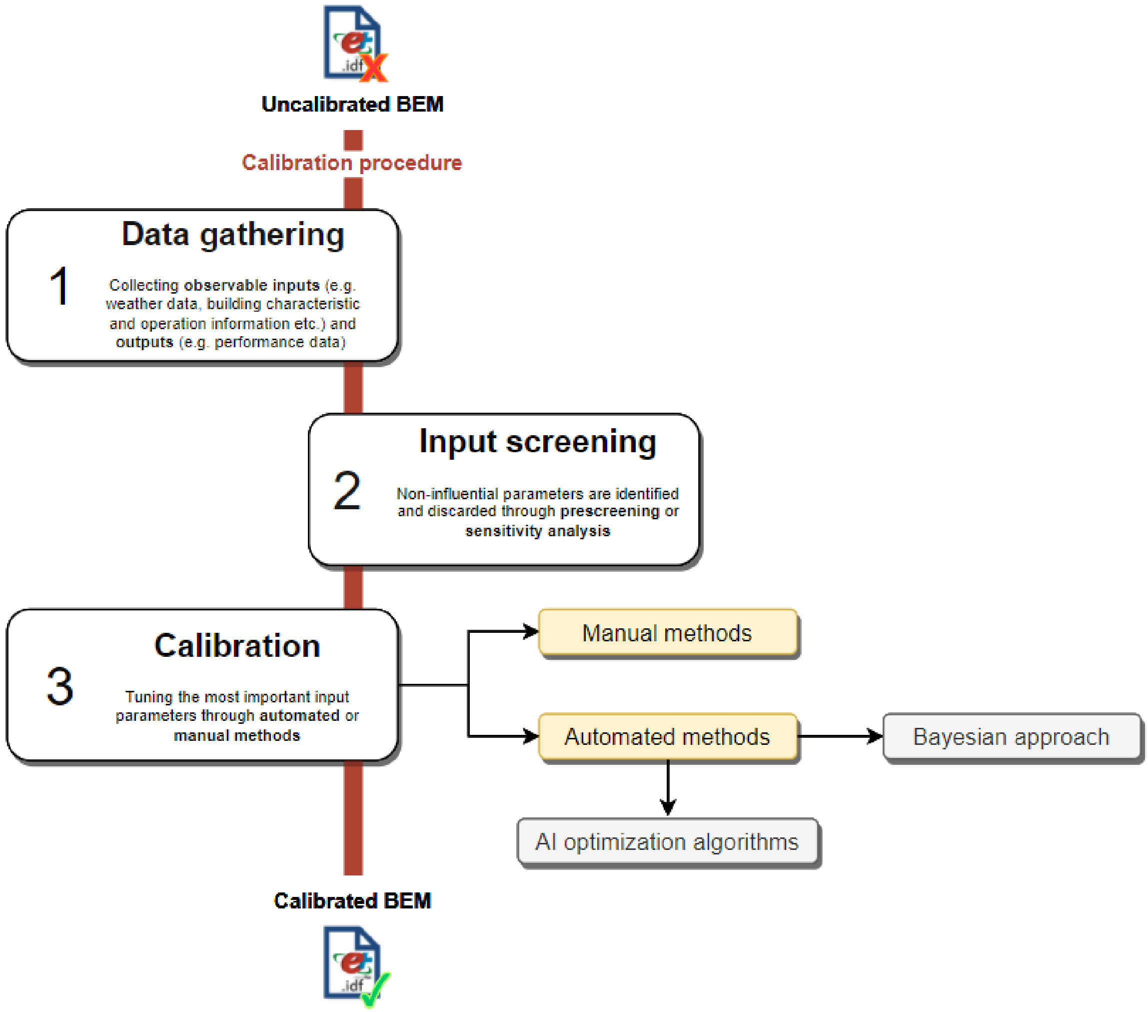


Figure 2: General workflow of state-of-the-art BC procedures.

2.2 Analytical techniques and approaches supporting BC

Several analytical techniques are combined in the literature to both manual and automated calibration approaches to assist and improve calibration efficiency. According to [7] and [9], the existing techniques can be subdivided into three main categories, i.e.:

a) **Model Characterization Techniques**, useful to characterize the physical and operational characteristics of the building being modeled;
b) **Model Simplification Techniques**, which aim to reduce the complexity of simulation models by reducing or aggregating the number of simulation variables to have a more efficient calibration process;
c) **Procedural Extensions**, regarding the use of standard processes or techniques to improve the calibration process.

The most relevant and important techniques are reported in Table 3 along with a brief description.

Table 3: Analytical approaches, tools, and techniques used to support the calibration process in the literature. Adapted from [7] and [9].

Acronym	Name	Description
Model characterization techniques		
AUDIT	Detailed audit	A detailed audit can help to gain a better knowledge of the building systems and characteristics (geometry, HVAC systems, lighting, equipment, and occupancy schedules).
EXPERT	Expert knowledge	Using expert knowledge or judgment is a key element of the calibration process to reduce user input requirements, involving the use of typical building templates and databases for typical building parameters and components.
INT	Intrusive testing	Intervention in the operation of the actual building, such as 'Blink Tests', which consists in turning on and off groups of end-use loads such as plugs loads, lighting, etc., in a controlled sequence to determine their overall impact on building load.
HIGH	High-res data	The adoption of high-resolution measurement data (generally hourly or sub-hourly) has better accuracy than that obtained with daily or monthly temporal resolution data.
STEM	Short-term energy monitoring	Metering equipment for a short period (>2 weeks) for identifying typical energy end-use profiles and/or baseloads.
Model simplification techniques		
BASE	Base-case modeling	Using measured baseloads to calibrate the building model. For example, when (a) heating and cooling loads are minimal to better characterize weather-independent variables such as internal loads, or (b) internal loads are minimal and the HVAC system is not operating to better characterize weather-dependent variables such as the building envelope.
PARRED	Parameter reduction	Reducing the requirement for detailed input for variable schedules (e.g., plug loads, lighting, occupancy, equipment, etc.). Day-typing and Zone-typing, for example, allow reducing detailed schedules into typical day-type schedules and aggregating spaces with similar thermal zones, respectively.
Procedural extension		

SA	Sensitivity analysis	Provide insights on how much the input uncertainty affects the outputs and then determine non-influential parameters that can be ignored during calibration or set priorities for future efforts (measurements or detailed investigation).
UQ	Uncertainty quantification	Assessment of parameter uncertainties as part of the calibration process. It can be used to directly assist in model calibration or for risk quantification within the results (e.g. uncertainty related to the risk quantification in energy conservation measure analysis).
EVIDENCE	Evidence-based model development	A procedural approach to model development and calibration, making changes according to source evidence rather than ad hoc intervention. Often requires model development version controls to keep track of the changes.

2.3 BEM calibration approaches

In BC, the discrepancy between observed and predicted data is generally evaluated through two error functions, i.e. the Coefficient of Variation of Root Mean Square Error, CV(RMSE), which indicates how close the numerical prediction is to the measured data, and the Normalized Mean Bias Error (NMBE), which is an indicator of the overall bias in the simulation predictions. In the literature, also the RMSE is used as an error function when measured and simulated air temperatures are considered [16–19].

In particular, the RMSE and CV(RMSE) are computed as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(m_i - s_i)^2}{n}} \times 100 \qquad (1)$$

$$CV(RMSE) = \frac{1}{\overline{m}}\sqrt{\frac{\sum_{i=1}^{n}(m_i - s_i)^2}{n}} \times 100 \qquad (2)$$

where m_i and s_i are the measured and simulated values respectively, $\overline{m}$ is the average of the measured values, n is the number of data points and n is the sample size.

The NMBE is computed as:

$$NMBE = \frac{\sum_{i=1}^{n}(m_i - s_i)}{m_i} \qquad (3)$$

According to the relevant Standards on BEM calibration, a BEM can be considered “calibrated” if the CVRMSE and the NMBE computed on hourly or monthly energy consumption fall below specific thresholds, synthetically reported in Table 4 [12–15]. The same Standards do not specify any thresholds or error functions for BEMs to be calibrated towards indoor air temperatures, despite this being an increasingly common approach for naturally ventilated buildings (NVBs) [16–19]. In this case, reference can be made to the case studies published in the literature, where calibrated BEMs generally achieve an NMBE ranging from –5.9% to 1.9%, a CV(RMSE) ranging from 1.7% to 20.3%, and an RMSE between

0.9 and 2.9 °C (with a maximum absolute error between 1.0 and 21.9 °C) [16–19]. These values, and in particular the RMSE ones, can be taken as a reference for BEM calibrated toward indoor air temperatures.

Table 4: Requirements for models calibrated on hourly or monthly energy consumption according to different guidelines and literature.

Reference	Monthly criteria (%)		Hourly criteria (%)	
	NMBE	CV(RMSE)	NMBE	CV(RMSE)
ASHRAE Guideline 14 [13]	±5	15	±10	30
IPMVP [14]	-	-	±5	20
FEMP [12]	±5	15	±10	30
Martinez et al. [15]	-	N/A	-	3

To achieve the required level of accuracy, different BC approaches can be adopted, which can be classified into two main categories, i.e. manual and automated methods [7].

Manual methods involve manual trial-and-error processes, whose success relies on the iterative pragmatic intervention of the energy modeler. The main drawbacks of these methods are that they are generally time-consuming, and their efficacy and efficiency are highly dependent on the experience and expertise of the modeler [20].

Automated methods have been developed to overcome the main limitations of manual methods. They tune model parameters to maximize the model's fit to observations by using computerized processes, which allow reducing time, increasing prediction accuracy, and minimizing the dependence of the calibration results on the experience and expertise of the energy analyst.

Due to their advantages, automated calibration approaches have raised increasing interest in the research field in the last decades [9]. The most adopted automated approaches are based on optimization algorithms, followed by the Bayesian-based approaches [9].

The latter has the main advantage of incorporating uncertainty in the calibration process, combining prior information with measured data to derive posterior estimates of the model parameters [21]. However, they have the main drawback of requiring very high computational costs and being highly data demanding [22].

Optimization-based calibration approaches are usually considered more efficient than Bayesian approaches, with much fewer requirements. They implement one or more penalty functions to determine the optimal set of variables that minimize the error between simulated and measured data and are generally suitable for optimization frameworks that minimize one (single-objective optimization) or more (multi-objective optimization) error functions.

In optimization-aided BC studies, metaheuristic Evolutionary Computation (EC) algorithms that mimic biological evolution, such as the Genetic Algorithm (GA) and Particle Swarm Optimization (PSO), are the most used [9,20,23]. These algorithms have proven to be very efficient to search for solutions in extremely large search spaces, also reducing the risk of converging to local minima [24]. Indeed, these techniques initialize the optimization process by considering a set of randomly (or quasi-randomly

[25]) distributed points, maintaining a set of possible solutions during the process, rather than performing a strictly gradient-based approach, allowing reducing the risk of converging to a local minimum.

Concerning error functions to be minimized by the optimization algorithm, CV(RMSE) and the NMBE are the most adopted ones. This is quite expected due to the CV(RMSE) and NMBE thresholds specified by BEM Calibration Standards to consider a BEM as calibrated (see) [12–14]. However, the NMBE suffers from cancellation effects between positive and negative bias, leading to erroneous interpretations of predictive performance. As a result, it cannot be efficiently adopted as a fitness function in optimization processes [15,26]. Conversely, according to the literature, the CV(RMSE) can be considered the most robust error function for candidates' fitness evaluation in an optimization-aided process [15].

It should be noted that the efficacy of automatic calibration approaches is not completely independent of energy modelers' expertise. Indeed, the calibration of a BEM is a highly undetermined problem (i.e., the parameters to be tuned are more than the data points to match), which can easily lead to equifinality issues (different solutions are obtained with the same prediction accuracy). Thence, if the search problem is formulated and solved blindly, i.e., without the guidance of an expert energy modeler, it could easily conduct to a situation in which, for example, the utility data are matched, but the model does not match reality. This may lead to the design of erroneous ECM and unsafe energy-saving predictions. For example, if a good fit with energy consumption has been obtained by inferring a poor coefficient of performance (COP) in the model when, in reality, infiltration rates are higher than those assumed, an HVAC retrofit ECM might be recommended but will likely provide lackluster results. Such a situation can be mitigated by applying, for example, a Bayesian or other regularization methodology, which biases the resulting model toward initial expectations (prior distributions). However, in this case, the data needed to build the prior distributions in the Bayesian framework for each to-be-calibrated input (data on the confidence we have in each piece of our model) must be gathered, validated, and subsequently used, then still heavily relying on experts' judgments, also being highly data-intensive for the most common applications.

2.4 Limitations of existing automatic calibration software tools

In the literature, most authors developed in-house software codes to automatize the state-of-the-art procedure depicted in Figure 2. However, to the authors' knowledge, only a few of them make their code available and understandable to the practitioner and then largely applicable in the engineering practice. In Table 5, a list of software tools specifically developed for BC and published in the literature is provided, along with their main features. As can be seen, most tools do not integrate the SA and EXPERT techniques within their codes. These techniques are fundamental to speed up the calibration process and minimize the dependence of the procedure efficiency on the experience of energy modelers, then making them suitable for use even by non-expert users. As a result, the state-of-the-art procedure is today scarcely applicable in most common applications and manual calibration approaches are generally preferred by the practitioner, slowing down the building renovation design process. Thus, a software tool able to perform BC by including both SA and EXPERT techniques is strongly needed to spread the use of automated calibration approaches in engineering practice.

Table 5: Software tools specifically developed for BC in the literature and main characteristics, i.e.: implemented optimization algorithms; simulation engine; observed outputs and related resolution that can be used for calibration; parallel computing; local simulation (or if a cloud-based provider is needed); SA and EXPERT techniques integration.

Reference	Name	Type	Opt. Algorithm	Simulation Engine	Observed Output(s)	Resolution	Parallel Comp.	Local Comp.	Cloud Comp.	SA	EXPERT
[20,27]	Autotune	Python code	EC	E+	All	Any	Yes	Yes	Yes	No	No
[28]	CBES Toolkit	Web based	Pattern-based	E+	Building Electricity and Gas	Monthly	No	No	Yes	No	Yes
[29]	OpenStudio PAT	App	NSGA-II, PSO, SPEA	E+	Utility bills	All	Yes	No	Yes	No	No
[30]	ExCalibBEM	App	PSO, GPS-HJ	E+, DOE-2, TRYNSIS, Dymola, IDA-ICE	All	All	Yes	Yes	No	No	No

EC: Evolutionary Computation, NSGA-II: Non-dominated Sorting Genetic Algorithm-II, PSO: Particle Swarm Optimization; SPEA: strength Pareto Evolutionary Algorithm; GPS-HJ: Generalized Pattern Search Hooke-Jeeves

3. BIMSPEED BEM calibration procedure

3.1 General description and workflow

The developed automated BC procedure and tool aim to spread and encourage the application of state-of-the-art BC in the engineering practice by simplifying and speeding up the entire calibration process through its automatization, then allowing the fast delivery of the calibrated BEM with the level of accuracy required for achieving reliable performance predictions to support renovation design, advanced control of renovated building operation as well as energy/environmental/economic assessment. The procedure, compliant with relevant BC and BIM-BEM Standards [12,14,31], assists the energy modelers over the entire calibration process, from the data gathering to the BEM optimization process, passing through model enrichment and SA technique.

The developed procedure can be subdivided into two main phases (Figure 3):

- **A data-gathering phase** (Section 3.3), aimed at obtaining the model inputs (weather conditions, schedule information, etc.) and outputs (e.g., energy consumption) required for the calibration process. The level of accuracy of the calibration process depends on the availability of these data, which can vary greatly from case to case. Then, different accuracy levels can be considered in this phase, from the lowest accuracy (minimum requirements) to the highest one;
- **An automated calibration phase** (Section 3.4), that can be carried out through the developed BEM-Calibration Tool (described in Section 4), aimed at speeding up and simplifying the calibration process for the end-user, and at increasing the predictive accuracy of the calibrated model with respect to manual approaches. In particular, for the first time in the literature, the tool integrates within the same workflow EXPERT techniques, SA, and Artificial Intelligence (AI) optimization algorithms, minimizing the number of inputs required from the practitioners in BC processes while maintaining an easy-to-use interface to facilitate its use in the engineering practice. The tool has been integrated into the BIMSPEED platform as a third-party service associated with *.idf* files.

It should be noted that, although automated, the BC procedure does not guarantee alone the calibration of the BEM. The success of the calibration process depends on the data availability and the accuracy and correctness of the input inserted by the user in the model (e.g. the geometry). For this reason, the developed BC procedure should be intended as an iterative process (Figure 3), with iterative model refinements and deepening of input and output until a calibrated model is obtained.

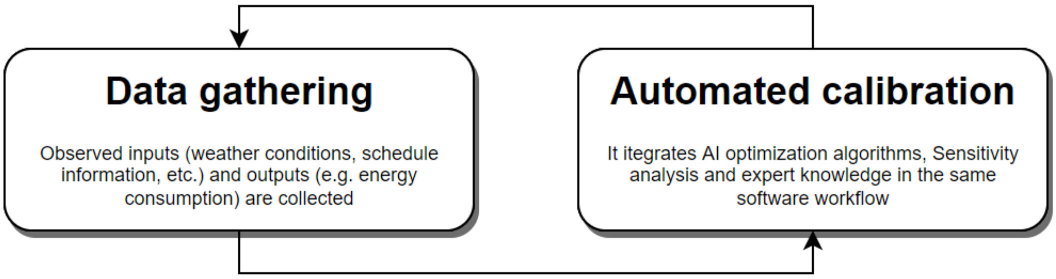


Figure 3: Iterative BC workflow.

3.2 Preliminary considerations on Building Energy Modeling and BIM-to-BEM

Before starting BC, one should be sure that the geometry and architectural features of the building (and of the surroundings such as near buildings and vegetation) are correctly inserted in the BEM and are consistent with the actual building configuration. This step is especially useful for BEMs created from BIMs. Indeed, these BEMs are often characterized by missing parts due to interoperability issues in the BIM-to-BEM process (see Deliverable 3.1 and 3.2). The user should be also aware that over-specifying architectural details may not have an impact on building energy use but strongly increase the computational cost. Then, the BEM should be maintained as simple as possible to speed up the calibration process. For example, if the target output of the BC is the energy consumption of a single apartment in a multifamily building, only the thermal zone of the target apartment (or the thermal zone plus the thermal zone of near apartments) could be modeled.

3.3 Data gathering

In this section, the data requirements for BEM calibration are described, also considering the different levels of accuracy that can be obtained in the different cases based on data availability while still being compliant with the BC Standards. Then, methods and guidelines to collect the most important data are briefly indicated.

3.3.1 Data requirements and accuracy levels

In the *data-gathering phase*, the data required for BEM calibration are collected. A correct data collection is required for the success of whatever model calibration process. Indeed, even when a reasonable match is obtained between the simulated and measured outputs, it is not certain that the model is a good representation of the building. Data collection must be then carried out to ensure that the calibrated BEM reflects as much as possible the actual building operation, performance, and characteristics. In general, since BC is a highly undetermined problem, the higher the number of information gathered, the higher the accuracy of the calibration procedure.

Depending on data availability, and according to the literature [11,32], five different levels of calibration can be identified in the most common applications (see Table 6):

- **Level 1** is the minimum requirement for BEM calibration and can be considered the most common in usual applications while being still compliant with BC standards. At this level, utility bills (generally monthly) and weather data (dry bulb temperature, solar radiation, relative humidity, wind speed, and direction) are collected to compare the predicted energy consumption with the observed ones. In NVBs, indoor air temperatures can be considered instead of energy consumption. The information about the building characteristics and operation is not very detailed and is not cross-checked with on-site visits;

- At **Level 2**, detailed site investigation or inspections allow verifying as-built data and collecting more information about building and system characteristics and operations;
- **Level 3** is based on a detailed audit of the case study, where on-spot measurements of the building operation and energy consumption are also collected;
- Finally, **Levels 4** and **5** are the most detailed and accurate levels of calibration, where data loggers are installed in the building to collect all the required missing information. In this case, BEMs can be calibrated against both utility energy consumption for space heating/cooling and/or indoor air temperatures (e.g., for calibrating BEM during free-floating periods).

This is not intended to be an exhaustive list. Other combinations of data can be considered if adequately justified. It should be noted that SA, carried out with the developed BC tool described in Section 3.4, can help to determine which input parameters are the most important and thus require higher accuracy in the collection phase, as well as those that do not affect model prediction and thus that can be excluded from the data collection plan.

Table 6: Calibration levels are based on building information availability [11,32].

Calibration Levels	Building Input/Output data availability						
	Weather data	As-Built Data	Utility Bills (or Indoor air temperatures in NVB)	Detailed Site investigation	Detailed Audit	Short-Term Monitoring	Long-Term Monitoring
1	X	X	X				
2	X	X	X	X			
3	X	X	X	X	X		
4	X	X	X	X	X	X	
5	X	X	X	X	X	X	X

3.3.2 Methods and guidelines for data collection

Methods and guidelines for data collection are here briefly described. Other methods can be followed if justified [14,31].

3.3.2.1 Monitoring period

Generally, the monitoring period must include the full range of expected operating conditions, modes, and independent variables. Typically, the baseline period is the period immediately before the retrofit and should represent one or more complete operating cycles to minimize bias. For weather-dependent loads, the baseline period for data collection should be a full year. If data cannot be obtained for a full cycle of operation, shorter periods that are representative of each operating mode (e.g., one month in each season) may be acceptable, especially if the data collection interval is reduced (e.g., from monthly to hourly) or if they can be considered as representative of another not-considered period [33]. In all cases, care must be taken to ensure that the baseline period is representative of typical conditions and does not over or underemphasize specific operating conditions.

3.3.2.2 Energy consumption

Utility bills can be used to obtain energy consumption for space heating/cooling to be used in the BC process. In some cases, however, the energy used for space heating and domestic hot water (DHW) is reported in the same bill, as can happen in residential buildings. In this case, it is not possible to understand which part of the energy is used for space heating and which for DHW. In these cases, it is suggested to estimate DHW energy consumption from summer bills, assuming a constant DHW energy consumption during the whole year, then obtaining the energy consumption for space heating by subtraction in the other months.

Demand energy data for space heating could be also useful when energy consumption data for space heating are not available from meters (e.g., in the case of district heating). In the latter case, equipment to collect such data should be installed.

When available, also hourly energy consumption and spot measurements can be collected and used to increase the calibration accuracy. Monthly utility bills should be used only when hourly data are not available or cannot be collected.

3.3.2.3 Weather data

The most common weather data affecting model output are outdoor air temperature and humidity (outdoor air conditions). Solar radiation (or cloud cover), and wind speed and direction can also affect building energy use. Accurate and consistent measurements and observations of weather conditions are critical.

At a minimum, the modeler collects (or gathers) hourly weather data corresponding to the same period as the energy use data to which the model will be calibrated. Data obtained from government weather stations can be considered the most reliable source of data for sites near the station. However, these data are limited to the location where weather stations are placed. Moreover, variations in microclimates can produce significant variations in weather data even over short distances due to changes in terrain, altitude, and building density. This may justify the use of on-site instrumentation.

When using government weather stations, the station that most closely represents the microclimatic conditions at the project site should be used, even if there are other stations closer to the project site. Where a nearby weather station is unavailable, a more distant station may be used if its weather pattern is well correlated to the pattern at the facility, even if the total heating or cooling conditions are somewhat different. In this case, short-term weather data from the site could be compared with the weather observations recorded at several weather stations to determine which station most closely corresponds to the site’s local weather conditions.

Although some modelers have reported using average or typical year weather data for model calibration, this approach is not recommended, as the comparison utility data are probably related to actual weather from the time in question. Several studies have shown that using an average year weather file in simulation can introduce errors into the simulation that is large as some of the differences that are being sought in the analysis.

3.3.2.4 Detailed site investigation

The following data *can* be collected from on-site surveys or occupants' interviews to be used as model input data to increase the calibration accuracy:

- *HVAC systems data* such as quantities, capacities, operating characteristics, and part-load performance curves of primary equipment (e.g. chillers and boilers). Since in most cases it may not be possible to exactly simulate a building's HVAC system or to retrieve all this information, especially in the case of existing residential buildings, an estimate of the global seasonal efficiency of the system and operating schedules can be sufficient for the aim of the calibration.
- *Lighting systems* data such as fixture counts, types, nameplate data from lamp and ballasts, 24-hour weekday, weekend, and holiday schedule of lighting use, characteristics of fixtures for estimating radiative and convective heat flows thermal zone assignments and diversity of operation.
- *Plug loads* data such as counts of and nameplate data from plug-in devices, 24-hour weekday and weekend schedules, and diversity of operation. Although measurements are preferable, plug loads may be estimated by taking inventory and summing connected loads. When doing so, the nameplate should not be entered into the simulation software. On average, most plug-load devices operate at an average power much lower than that of the nameplate rating. The actual operating power is obtained by multiplying the nameplate power by a use factor (generally 0.3 is used as a common rule of thumb).
- *Building occupants' activity*. Population counts; weekday, weekend, and holiday schedules; activity levels; assignment to thermal zones. Schedules of occupancy and operation can be investigated by data collection forms provided to the occupants. Information should include building use, thermostat settings, occupancy, operational data, windows opening, lighting, and equipment use.

3.3.2.5 Spot, short-term, and long-term measurements

Spot measurements are generally taken for a moment using handled instruments, while short- and long-term ones entail the use of instruments having data logging capabilities that are set up and left in place. Spot and short-term measurements can be used to better define model inputs. For example, short-term measurements provide valuable information regarding the schedule of use. Long-term measurements can be used to monitor both model input and output (e.g., energy consumption at a higher resolution scale through smart metering or indoor air temperatures). The most important in-situ measurements are the following (most important first):

- *Energy use and related operating schedules for space heating/cooling* through smart metering to be used as a model output and input, respectively;
- *The indoor air temperature* that can be used as both model input (e.g. as heating setpoint) and output;

- *Windows opening schedules, building ventilation, and infiltration* to be used as model input. If resources permit, building infiltration should be measured through blower door tests because these values often vary from expectations, also having a high impact on energy consumption [19].
- *Lighting systems and plug loads*. Operating schedules through smart metering and electric power to be used as model input;
- *Building occupancy* through motion detection, CO2 sensor, and lighting levels.

3.3.2.6 Missing data

Missing data may be estimated or interpolated from measured data using statistically valid engineering techniques. The data used to interpolate or estimate the missing data shall represent the full range of operating conditions experienced during the missing data interval (if the dependent variable data are missing) or similar adjacent intervals (if data for the independent variables are missing). The data set used for interpolation or estimation of missing data should be an order of magnitude greater than the missing data interval (e.g., for monthly data, 12 months; for daily data, 7 to 14 days; for hourly data, 12 hours; etc.).

3.4 Automated calibration

In this section, the automatic calibration process and its implementation in the BEM-Calibration Tool are described.

3.4.1 General workflow

The general workflow of the automatic procedure consists of three main phases:

- **A parameter prescreening phase** (see Section 3.4.2), which, based on EXPERT techniques:
    - defines which are the most important model inputs that should be considered for calibration, allowing for reduction of the parameters' search dimensionality, and speeding up the automatic process;
    - assigns to each of them a range of variation based on building typological characteristics (e.g. residential buildings) for the parameterization.
- **A SA phase** (see Section 3.4.3), based on the Morris method, to determine:
    - which parameters, among those identified in the previous phase, are the most influencing on the model prediction and for which a more accurate estimation could be required;
    - which parameters are non-influential and can be discarded from the next phase to reduce the computational burden of the optimization process;
- **An optimization phase** (see Section 3.4.4), that implements AI optimization algorithms to efficiently determine the model input values that provide the best fit between experimental and predicted measurements.

At the end of the automated process, the building energy modeler compares the obtained goodness of fit of the optimized model (e.g. the CV(RMSE) value) with the thresholds defined by the Standards to determine if the model can be considered calibrated or not (see Section 2.3). If not, one or more observed inputs that are not varied during optimization (e.g. the schedule of the HVAC activation) or the range of variation of unobserved inputs, are further investigated in the data-gathering phase and the automated process restarted until the calibration thresholds are met, as depicted in Figure 3.

3.4.2 Parameters pre-screening

A typical BEM in EnergyPlus has approximately 3000 input parameters that must be specified for a given building [20]. As a result, the search space in the calibration problem is extremely large. If each parameter were simply binary, the number of possible models is 2^{3000}, i.e. higher than the number of subatomic particles in the observable universe (10^8). Since many of the BEM parameters are continuous values, the actual size of the search space is almost infinite. Thus, it is strongly needed to reduce the number of parameters to be considered in BC to limit the time and computational effort of building calibration processes. Based on these considerations, the parameter pre-screening phase entails the automatic identification of the most important unobserved parameters to reduce the parameter space search dimensionality and computation time of the calibration problem [7].

In a BC process, model input can be subdivided into an input that can be observed (or estimated with a good approximation) and those that will remain unobserved (or uncertain).

We can also subdivide model input into variables and parameters, where variables refer to input that varies over time (e.g., windows opening schedules), while parameters do not relate to time-varying values. In some cases, an input can be a variable or a parameter depending on how it is modeled [7].

The unobserved parameters are generally assumed to be the main responsible for the discrepancy between simulation and measured output(s) in automatic procedures [9]. Indeed, unobserved inputs to be calibrated in automatic procedures generally include building envelope characteristics (material properties and infiltration rate), internal gains characteristics such as occupant, lighting, and equipment power density, and zone heating or cooling setpoints. Conversely, schedules (i.e. unobserved variables) are typically not adjusted. This is due to the sharp increase in computational cost if every schedule parameter is considered in the calibration [7].

Considering that the most common observed outputs in BEM calibration are the energy consumption for space heating and indoor air temperatures, the most influential parameters that are automatically identified and parametrized are the following:

- heating setpoint (for each thermal zone);
- ventilation flow rates (for each thermal zone);
- infiltration flow rates (for each thermal zone);
- internal loads (lighting, people, and equipment for each thermal zone);
- the thermal transmittances of the opaque elements (internal and external);
- the thermal transmittances of the windows (U-Factor);
- the Solar Heat Gain Coefficients of the windows (SHGC);
- Heating/Cooling system efficiency or Coefficient of Performance (CoP).

The density of the opaque elements is also considered given the impact that it may have on indoor air temperatures. A range of variation is then automatically assigned to each identified parameter based on building typology (residential buildings), defined base on the information from the TABULA/EPISCOPE projects database and relevant standards [34,35] (EXPERT technique). The range of variation can be then modified by the user at the end of the assignment process to consider specific knowledge of building characteristics.

3.4.3 Sensitivity analysis: Parameter Screening with Morris method

A SA is carried out to define the parameters that have the highest impact on the model output (for which a further investigation can be required in the data-gathering phase), and those that are not influential and thus negligible in the calibration process.

Among SA methods, global SA methods are the most used in BC procedures, thanks to their ability to provide an overall view of the importance of different inputs also considering their possible interactions [9,36]. The Sobol' SA is considered the most accurate global SA method for quantifying the impact of input parameters on output ones [3,37]. However, it requires a high amount of model evaluations and thus high computational resources to obtain reliable results. The Morris method, instead, is a global sensitivity screening method generally considered more suitable for models that are typically non-linear and with a high dimensional parameter space (as BEMs) requiring a quite low computational cost without losing accuracy if compared to Sobol' sensitivity. For this reason, in this work, the Morris method is used as a screening method to simply identify and drop noninfluential parameters.

In particular, the Morris method combines an efficient parameter screening method with a factorial sampling strategy to identify uninfluential parameters, i.e. those parameters that can be fixed at any value within their range of variation without affecting the variance of the simulation results [38].

Concerning the sampling, the parameters are firstly transformed into dimensionless variables in the interval (0;1). Then, the parameter space is discretized by dividing each parameter interval into a certain number (*p*) of levels, forming a regular grid in the unit-length hypercube H^k.

The sampling starting point is then randomly chosen, while each sample differs from the previous one in one coordinate only. Then, a trajectory is generated, i.e. a sequence of *k*+1 points (*k* is the number of input parameters) where each parameter changes only once by a pre-defined value Δi. Each point of this trajectory corresponds to an evaluation of the model. The variation magnitude in the model output due to the variation of one parameter is called the elementary effect (EE) and can be computed as follow:

$$EE_i = \frac{[Y(X + e_i \Delta_i) - Y(X)]}{\Delta_i} \quad (4)$$

where:

- X is an N*k matrix of model inputs with N samples (or trajectories) and k input parameters defined within a uniform range of variation;
- Y(X) is a mathematical function representing the model (and then the model output);
- e_i is the vector of zeros except for the *i-th* element that will be equal to ±1 representing an incremental change in the *i-th* parameter.

One trajectory allows evaluating an elementary effect for each parameter. Then, a set of N trajectories enables statistical evaluation of the elementary effects. The most used quantitative sensitivity measures in the Morris methods are the absolute mean (μ^*) and the standard deviation (σ), computed as follows:

$$\mu_i^* = \frac{1}{N}\sum_{N=1}^{r} |EE_{it}| \tag{5}$$

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{N=1}^{r} (EE_{iN} - \mu_i)} \tag{6}$$

In particular, μ* (mu_star) measures the overall influence of the input parameter on the model output and is often used to rank the parameters according to their importance. σ (sigma), instead, assesses the effect of the input due to the interaction with other parameters as well as its non-linearity.

The computational burden of the Morris method, i.e. the number of simulations required to rank parameters with sufficient accuracy, is N *(k + 1), with N generally ranging between 5 and 15 [39]. In this work, to reduce the computational burden without losing accuracy, an N value equal to 5 is assumed, as made in [40].

The ranking of parameters according to the μ^* *values* obtained through the Morris method can be biased by the presence of outliers in the EE, especially in the case of low N values combined with a large parameter space (Table 1), which leads to a low number of EEs for each parameter [36]. To overcome this issue, one possible strategy, which does not recur to the simple increase of N, consists of using the absolute median χ^* (median_star) to characterize the EE distribution. This statistical parameter, indeed, is a robust measure to characterize skewed distributions, reducing the impact of the outliers in the result [36]. For this reason, in this work, we use the χ^* value for the ranking of the parameters.

Once parameters are ranked from the most important to the less important, only the group of most important parameters is considered in the calibration step. This group is composed of the first parameters whose sum of their medians of elementary effects – i.e. median star - is higher than 90% of the total effects. Other parameters are considered non-influential or not significant and will be discarded in the next optimization/calibration phase.

3.4.4 Model optimization: The NSGA-II algorithm

In the *model optimization* phase, Artificial Intelligence (AI) optimization algorithms are adopted to define the best set of values for the inputs parameters that minimize one (single-objective optimization) or more error functions (multi-objectives optimization). It should be noted that only the parameters selected in the previous sensitivity/screening phase are varied in this phase.

In particular, the non-dominated sorting genetic algorithm II (NSGA-II) is used in this work since able to obtain a better spread of solutions and convergence than other evolutionary algorithms, and for this reason, the most adopted in the literature for BC problems [41].

The algorithm starts creating a random or quasi-random population (a set of "individuals", or candidate solutions, i.e. models with different input values) based on the ranges given to the model inputs (defined in Section 3.4.2). In the context of BC, a candidate solution is a BEM (.idf) characterized by a list of values (chromosome), one for each unobserved parameter to be tuned. In our approach, the first population is sampled through the Latin-Hypercube Sampling (LHS) technique, which is a stratified random sampling technique that allows ensuring that the entire search space is homogeneously sampled. A sample size equal to 4 times the considered number of parameters is also adopted [42,43].

Then, each individual or BEM is evaluated through a fitness function. The obtained "fitness" is a problem-dependent measure of how well each candidate solves the specific problem (in evolutionary terms, fitness is a measure of survivability of the specific individual). For calibration problems, different fitness functions can be used. In our approach, the CV(RMSE) is considered (see Eq. (1)), since proven to be the most efficient in driving the automatic optimization process [15]. In the case of multi-output optimization, the average of the CV(RMSE) of the different outputs is adopted as a fitness function [15]. The candidates are then ranked based on nondominated sorting according to their fitness values.

When the solutions are ranked, genetic operators are applied, which manipulate the selected chromosomes to generate new offspring. These operators drive the evolutionary process. The most used ones are selection, crossover, and mutation.

The selection operator copies the individual strings from the parent chromosomes into the new population. The most adopted selection method is tournament selection, where individuals are randomly chosen from the population and compared with each other in terms of fitness values. Then, the best is chosen as a parent of the offspring.

The crossover operator, which is the most important genetic-mimicking probabilistic operator, then combines two parent solutions with high fitness to create a new generation.

Finally, the diversity within a population is guaranteed by the mutation operator, which randomly acts after the crossover operator to avoid the loss of genetic material that may occur due to the previous operators. It acts by randomly modifying the chromosome values of one or more offspring by introducing new genetic material.

The elitism operator can also be adopted which guarantees that the best solutions (maximum fitness value) directly pass to the next generation, improving the speed of convergence without losing any best solutions. In this work, a crossover rate, mutation rate and tournament size equal to 1, 0.2 and 2, respectively, are considered [43].

The stopping criterion of the optimization process is hard to be defined. For this reason, in this procedure, it is set as a maximum number of generations or imposed by the user during the calculation (e.g. when no more improvement is obtained between two or more consecutive generations).

After the optimization procedure is completed, a model evaluation is performed to define if the optimized BEM can be considered calibrated or not according to the thresholds reported in Section 2.3. Expert judgment should also be used in this phase to discard unreasonable calibrated solutions (e.g. with unreasonable input values).

4. Software implementation: The BEM-Calibration Tool

The BEM-Calibration Tool (BIM-Calibration service library, from DoA) implements the automatic optimization-aided calibration process described in the previous sections. The tool has been specifically developed for its use in combination with EnergyPlus models (.idf, IDF in the following). EnergyPlus is a free, open-source, and cross-platform building energy simulation program, that can be considered the most used for dynamic and detailed energy simulations all around the world. In particular, the BEM-Calibration Tool is EnergyPlus version agnostic, thus it accepts IDF models built with whatever EnergyPlus version, provided that the correct version of EnergyPlus matching the model version is correctly installed in the local machine to perform the simulations.

The software architecture of the BEM-Calibration Tool is shown in Figure 4 while its graphical interface is shown in Figure 5. The software functioning is split into three main modules, that exchange information with each other. These are the IDF analysis module, the Sensitivity Analysis module, and the Calibration module. Each module corresponds to a different button in the Graphical User Interface (GUI, see Figure 5) that can be pressed only if the previous button/analysis is already carried out and its results are saved in the working folder (IDF folder).

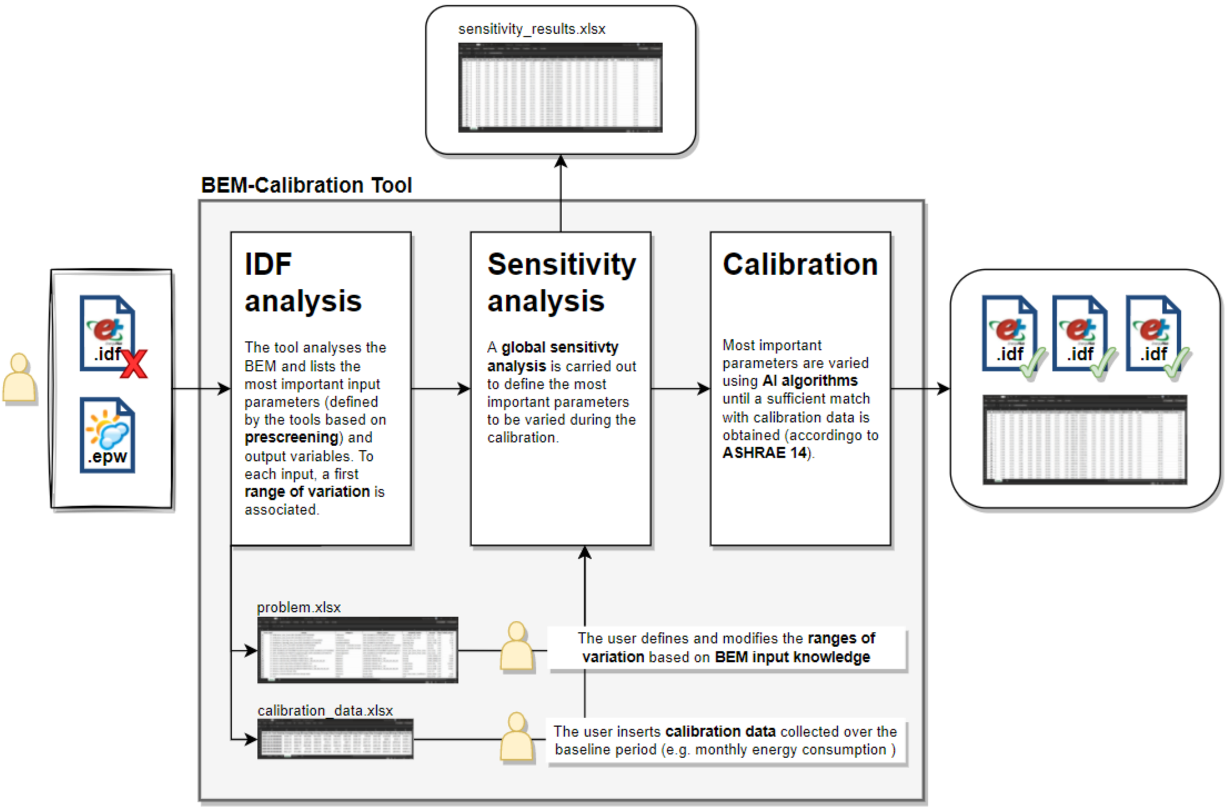


Figure 4: BEM-Calibration Tool workflow with a short description of the three different main modules: IDF analysis, Sensitivity analysis and Calibration. The workflow initial inputs (on the left) are the uncalibrated model (IDF) and the weather file (EPW). After the run of the first module (IDF analysis), the tool asks for modifying the range of variation of the selected parameters defined in the "problem.xlsx" to proceed with the sensitivity (if needed) and to insert the calibration data into the "calibration_data.xlsx" file needed for the calibration. The results of the entire process are different calibrated IDFs and a summary table with calibrated input values for each calibrated model.

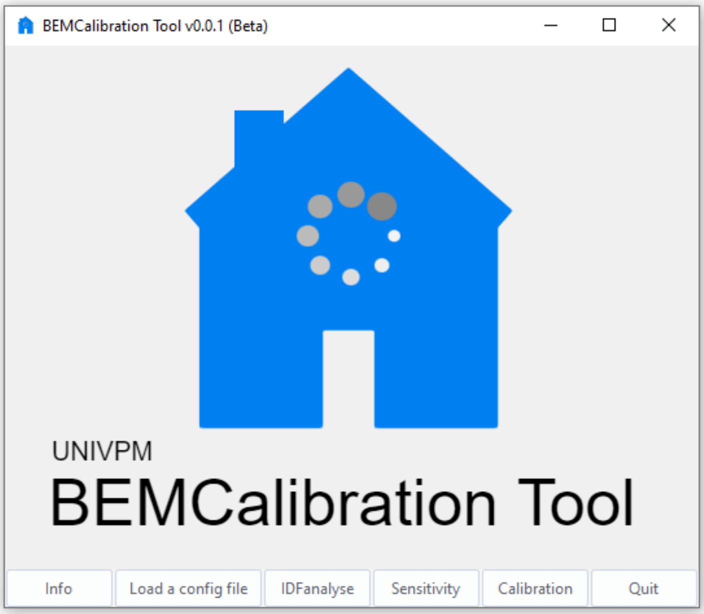


Figure 5: BEM-Calibration Tool GUI. Once the configuration file is loaded (json), the procedure starts by pressing, once at a time and with the provided order, the three buttons related to the three modules of the tool.

4.1 Load a config file

First, the user presses the “Load a config file” button that allows the user to load its specific configuration file (.json). An example of a configuration file is reported in Figure 6. In particular, the configuration file specifies:

- the location of the IDF in the local computer. The IDF folder automatically becomes the working folder of the software, where all the results will be saved. The IDF also defines which version of EnergyPlus must be used (this should be installed in the local computer in the folder suggested by EnergyPlus during installation);
- the location of the weather file (.epw, EPW in the following) in the local computer;
- the initial CoP value and its range of variation that is applied to obtain energy consumption from energy demand values obtained from the Ideal Load system;
- the output variable name, frequency, and key (e.g. thermal zone name) for which the sensitivity elementary effects of the Morris method should be computed;
- the maximum number of generations to be created in the optimization process.

```json
{
    "epw_file": "C:/Tool/2019-2020.epw",
    "idf_file": "C:/Tool/test/test.idf",
    "initial_CoP": 0.8,
    "CoPbounds": [0.7, 0.9],
    "variable_name": "Electricity:Facility",
    "variable_key": null,
    "variable_frequency": "Timestep",
    "max_generations": 30
}
```

Figure 6. An example of a configuration file (.json).

4.2 IDFanalyse

Once the configuration file is loaded, the *IDFanalyse* button can be pressed to start the IDF analysis, which automatizes the parameter pre-screening phase described in Section 3.4.2. The new command opens a new window terminal where the results of the analysis are plotted and instructions for the user are given (see Figure 7).

First, all the new folders needed for storing the results are created in the IDF folder. Previous calibration results are also eliminated. Then, the tool analyses the provided BEM. During this phase, the following information is plotted in the window terminal for user verification and model checking (see Figure 7):

- the number of heated and not heated thermal zones;
- the number of non-adiabatic surfaces, specifying the wall typology (wall, exterior floor, ceiling, roof, and window) and boundary conditions where applicable (exterior, interior, or ground floors);
- the number of stratigraphies characterizing the non-adiabatic surfaces, referred to as “constructions”, also specifying the related surface typology and boundary conditions where applicable;

- for each non-adiabatic construction, the most important layers, referred to as "materials", are listed, distinguishing between opaque materials (only those thicker than 0.02 m are considered to reduce the number of layers thermal properties to be varied), no mass materials (air gap), and windows materials. For opaque thick materials, a distinction is also made between insulating and massive materials, based on thermal conductivity λ values, using $\lambda = 0.1$ W/m^{2}K as a limit threshold between the two material classes.

For each opaque element, only the thickest layers among insulation materials and massive materials are considered for the parametrization in the sensitivity and calibration process. For massive materials, only the density values are parametrized, while the conductivity is parametrized in absence of an insulation layer. For insulation layers, only the conductivity value is parametrized. This approach aims to use a single value to modify the entire U-value of a single construction element, minimizing as much as possible the possible compensation errors and loss of efficiency that can be obtained in the sensitivity and calibration phases if all the layers are parametrized at the same time. This procedure is valid regardless of the number of layers that compose a stratigraphy, thus applying to both detailed and simplified modeling approaches. When a material is shared among two different constructions, it is automatically duplicated to have a different material for each construction. This is carried out to avoid that a change in a material parameter value may affect the U-value of two different construction elements (e.g. internal and external walls that can share the same masonry layer).

```
shared folder exists
/results folder created.
/finalpop_IDFs folder created.
/data folder created.
/ParametricIDFs folder created.
--- IDF ANALYSIS ---
--- Thermal zones ---
There are:
- 5 thermal zones, 4 heated

--- Surfaces ---
There are:
- 61 non-adiabatic wall surfaces: 52 exterior and 9 interior;
- 0 non-adiabatic exterior floor surfaces, 2 interior and 1 ground floors;
- 0 non-adiabatic ceiling surfaces;
- 3 non-adiabatic roof surfaces;
- 33 windows surfaces.

--- Constructions ---
From the above printed non-adiabatic surfaces,there are:
- 2 exterior wall constructions;
- 2 interior wall constructions;
- 0 exterior floor constructions;
- 1 interior floor constructions;
- 1 ground floor constructions;
- 0 ceiling constructions;
- 1 roof constructions;
- 3 windows constructions.

--- Materials ---
From the above printed constructions, there are:
- 9 opaque thick materials: 0 insulation thick materials and 9 massive thick materials;
- 1 no mass materials (airgap);
- 3 windows materials.

Preparing idf for sensitivity analysis and calibration..
The IDF for sensitivity analysis is now stored in ParametricIDFs folder as C:/Tool_4/test\results/data/parametric_IDFs/parametric_idf.idf
Duplicating shared materials...
M12_Mattoni_forati (80mm) material has been duplicated.
Now, there are:
- 10 opaque massive thick materials;
- 0 opaque insulation thick materials.


Defining thickest insulation and massive materials for each construction and typology...
Problem setting...
END
Please check the ranges of variation of parameters in 'problem.xlsx' before proceeding
```

Figure 7: Example of IDF analysis results plotted in the window terminal for user verification. The module plots the number of heated and not heated thermal zones, the number and type of non-adiabatic surfaces; the number of stratigraphies (constructions) of the non-adiabatic surfaces; the most important layers, referred to as "materials", for each non-adiabatic construction.

After this process, the main result of the pre-screening phase is plotted in an editable file (*problem.xlsx*), where the selected, most important parameters are listed along with their suggested range of variation. An example is reported in Figure 8. Each parameter is characterized in terms of variables name used by the software, input category, object name, property name, suggested range of variation according to the adopted EXPERT technique (bounds), distribution type (uniform, this is the only admissible in the version) and initial value set by the user in the original IDF. Before proceeding with the sensitivity analysis, the user can modify the range of variation to infer specific knowledge about building characteristics acquired during the data collection phase (for example, reducing the range of variation of thermal conductivity values of a specific element) or keep the suggested ranges.

	num_vars	names	category	object_name	property_name	bounds	dists	initial_values
0	33	Infiltration_rate:Z01_Flat_2R_Var.0	Infiltration	Z01_Flat_2R_IG1 Infil	Air_Changes_per_Hour	[0.3, 1.2]	unif	0.6
1	33	Infiltration_rate:Z02_Flat_3R_Var.1	Infiltration	Z02_Flat_3R_IG4 Infil	Air_Changes_per_Hour	[0.25, 1.0]	unif	0.5
2	33	Infiltration_rate:Z03_Flat_3L_Var.2	Infiltration	Z03_Flat_3L_IG1 Infil	Air_Changes_per_Hour	[0.25, 1.0]	unif	0.5
3	33	Infiltration_rate:Z04_Flat_4R_Var.3	Infiltration	Z04_Flat_4R_IG1 Infil	Air_Changes_per_Hour	[0.25, 1.0]	unif	0.5
4	33	Ventilation_rate:Z02_Flat_3R_Var.4	VentilationDFR	Z02_Flat_3R_IG5 Ventil	Air_Changes_per_Hour	[0.2, 0.8]	unif	0.4
5	33	Heating_set_point:Z01_Flat_2R_Var.5	Thermostat - Schedule:Constant	Heating_set_pointZ01_Flat_2R	Hourly_Value	[18, 22]	unif	20
6	33	Heating_set_point:Z02_Flat_3R_Var.6	Thermostat - Schedule:Constant	Heating_set_pointZ02_Flat_3R	Hourly_Value	[18, 22]	unif	20
7	33	Heating_set_point:Z03_Flat_3L_Var.7	Thermostat - Schedule:Constant	Heating_set_pointZ03_Flat_3L	Hourly_Value	[18, 22]	unif	20
8	33	Heating_set_point:Z04_Flat_4R_Var.8	Thermostat - Schedule:Constant	Heating_set_pointZ04_Flat_4R	Hourly_Value	[18, 22]	unif	20
9	33	Metabolic rate:Z02_Flat_3R_Var.9	People - Schedule:Compact	Z02_Flat_3R_IG1 ActivityLevel	Field_4	[80, 200]	unif	120
10	33	Lights:Z02_Flat_3R_Var.10	Lights	Z02_Flat_3R_IG2 Lights	Watts_per_Zone_Floor_Area	[3.3, 13.2]	unif	6.6
11	33	OtherEquipment:Z02_Flat_3R_Var.11	OtherEquipment	Z02_Flat_3R_IG3 ElecEq	Power_per_Zone_Floor_Area	[2.7, 10.8]	unif	5.4
12	33	Exterior wall Density:M12_Mattoni_forati (80mm)_Var.12	Material	M12_Mattoni_forati (80mm)	Density	[400, 2500]	unif	800
13	33	Exterior wall Density:Duplicated M12_Mattoni_forati (80mm) for C02_CV2_Tamponatura_con_fini_Var.13	Material	Duplicated M12_Mattoni_forati (	Density	[400, 2500]	unif	800
14	33	Ground_floor Density:M03_Ground_floor_generic (300mm 0.6)_Var.14	Material	M03_Ground_floor_generic (300	Density	[400, 2500]	unif	1000
15	33	Interior floor Density:M16_Solaio_in_laterocemento (250mm)_Var.15	Material	M16_Solaio_in_laterocemento (	Density	[400, 2500]	unif	1800
16	33	Interior walls Density:M11_Mattone_forato (110mm)_Var.16	Material	M11_Mattone_forato (110mm)	Density	[400, 2500]	unif	800
17	33	Interior walls Density:M11_Mattone_forato (70mm)_Var.17	Material	M11_Mattone_forato (70mm)	Density	[400, 2500]	unif	800
18	33	Roof Density:M09_Ladrillo_hueco_12 (120mm 0.6)_Var.18	Material	M09_Ladrillo_hueco_12 (120mm	Density	[400, 2500]	unif	100
19	33	Exterior wall Conductivity:M12_Mattoni_forati (80mm)_Var.19	Material	M12_Mattoni_forati (80mm)	Conductivity	[0.1, 2.4]	unif	0.414
20	33	Exterior wall Conductivity:Duplicated M12_Mattoni_forati (80mm) for C02_CV2_Tamponatura_con_fini_Var.20	Material	Duplicated M12_Mattoni_forati (	Conductivity	[0.1, 2.4]	unif	0.414
21	33	Ground_floor Conductivity:M03_Ground_floor_generic (300mm 0.6)_Var.21	Material	M03_Ground_floor_generic (300	Conductivity	[0.1, 2.4]	unif	1.42857
22	33	Interior floor Conductivity:M16_Solaio_in_laterocemento (250mm)_Var.22	Material	M16_Solaio_in_laterocemento (	Conductivity	[0.1, 2.4]	unif	0.757576
23	33	Interior walls Conductivity:M11_Mattone_forato (110mm)_Var.23	Material	M11_Mattone_forato (110mm)	Conductivity	[0.1, 2.4]	unif	0.3
24	33	Interior walls Conductivity:M11_Mattone_forato (70mm)_Var.24	Material	M11_Mattone_forato (70mm)	Conductivity	[0.1, 2.4]	unif	0.3
25	33	Roof Conductivity:M09_Ladrillo_hueco_12 (120mm 0.6)_Var.25	Material	M09_Ladrillo_hueco_12 (120mm	Conductivity	[0.1, 2.4]	unif	1.2
26	33	Window_UFactor:V04_PVC_Doppio_vetro (10mm)_Var.26	Windows	V04_PVC_Doppio_vetro (10mm)	UFactor	[0.6, 7.0]	unif	3.1
27	33	Window_SolarHeatGainCoefficient:V04_PVC_Doppio_vetro (10mm)_Var.27	Windows	V04_PVC_Doppio_vetro (10mm)	Solar_Heat_Gain_Coefficient	[0.2, 0.99]	unif	0.7
28	33	Window_UFactor:V01_AL2_Vetro_doppio (10mm)_Var.28	Windows	V01_AL2_Vetro_doppio (10mm)	UFactor	[0.6, 7.0]	unif	3.1
29	33	Window_SolarHeatGainCoefficient:V01_AL2_Vetro_doppio (10mm)_Var.29	Windows	V01_AL2_Vetro_doppio (10mm)	Solar_Heat_Gain_Coefficient	[0.2, 0.99]	unif	0.7
30	33	Window_UFactor:V02_AL_Vetro_singolo (10mm)_Var.30	Windows	V02_AL_Vetro_singolo (10mm)	UFactor	[0.6, 7.0]	unif	5.7
31	33	Window_SolarHeatGainCoefficient:V02_AL_Vetro_singolo (10mm)_Var.31	Windows	V02_AL_Vetro_singolo (10mm)	Solar_Heat_Gain_Coefficient	[0.2, 0.99]	unif	0.7
32	33	CoP_Var.32	HVAC	None (HVAC)	CoP	[0.6, 1]	unif	0.8

Figure 8: Example of problem.xlsx file where each screened parameter is characterized in terms of variables name, input category, object name, property name, and suggested range of variation.

4.3 Sensitivity screening

Once the IDF analysis is carried out, the *Sensitivity* button can be pressed. The tool runs a first test simulation to verify the correctness of the IDF file. Then, a sensitivity analysis is carried out considering the impact of the parameters on the output variable indicated by the user in the configuration file. Figure 9 shows the sensitivity analysis information plotted by the software in the window terminal.

```
--- Running test E+ simulation ---
EnergyPlus Completed Successfully.
...saving
    50.71 seconds for running the test E+ simulations
The number of variables are 33
The number of Morris simulations are 170

--- Running 170 E+ simulations ---
  0%|                                                                    | 0/170 [00:00<?, ?it/s]
nergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.

EnergyPlus Completed Successfully.
 91%|██████████████████████████████████████████               | 154/170 [03:13<00:19,  1.20s/it]
nergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
175it [03:41,  1.26s/it]

--- 3.7 minutes for running 170 E+ simulations ---
Computing sentivity indexes ...
END
Please insert calibration data in 'calibration_data.xlsx' before proceeding
```

Figure 9: Sensitivity analysis information plotted in the window terminal. The number of Morris simulations to be run is indicated before running the simulation.

When all the simulations are completed, the results of the sensitivity analysis are saved in the *Sensitivity_results.xlsx* file within the IDF folder. The file reports the main results that can be obtained from the Morris method. An example is reported in Figure 10 where the input values are sorted in order of importance according to the median of the elementary effects (median_star, see Section 3.4.3).

It should be noted that the sensitivity analysis can be performed more times without carrying out the IDF analysis again by varying the ranges of variation in the problem.xlsx file.

	names	mu_star	mu_star_conf	sigma	mu	mu_test	median_star	median_star_conf
6	Heating_set_point:Z02_Flat_3R_Var.6	9836.533916	2093.318032	3172.337581	9836.533916	9836.533916	8887.75	2616.64754
32	CoP_Var.32	5785.132308	3579.926875	4714.931501	-5785.132308	-5785.132308	3145.2	4474.908594
1	Infiltration_rate:Z02_Flat_3R_Var.1	2868.45	801.8436621	968.7840865	2868.45	2868.45	2968.25	1002.304578
19	Exterior wall Conductivity:M12_Mattoni_forati (80mm)_Var.19	5869.039161	3730.465064	4792.206891	5869.039161	5869.039161	2623.75	4663.081331
5	Heating_set_point:Z01_Flat_2R_Var.5	1811.82	580.9385889	887.7203034	-1811.82	-1811.82	2079.25	726.1732361
8	Heating_set_point:Z04_Flat_4R_Var.8	1674.35035	393.516374	622.2233718	-1674.35035	-1674.35035	1884.115385	491.8954675
10	Lights:Z02_Flat_3R_Var.10	1462.761189	335.3914606	422.3440344	-1462.761189	-1462.761189	1613.5	419.2393257
4	Ventilation_rate:Z02_Flat_3R_Var.4	1625.45	399.3973596	547.2386259	1625.45	1625.45	1605.75	499.2466995
30	Window_UFactor:V02_AL_Vetro_singolo (10mm)_Var.30	1038.240769	365.8544538	436.2570524	1038.240769	1038.240769	1124.75	457.3180672
28	Window_UFactor:V01_AL2_Vetro_doppio (10mm)_Var.28	1281.96	384.7615836	495.6818869	1281.96	1281.96	1098	480.9519796
31	Window_SolarHeatGainCoefficient:V02_AL_Vetro_singolo (10mm)_Var.31	1011.415035	173.2712809	194.8171761	-1011.415035	-1011.415035	968.1136364	216.5891011
11	OtherEquipment:Z02_Flat_3R_Var.11	946.4213287	138.8638524	184.6799784	-946.4213287	-946.4213287	946.0384615	173.5798155
29	Window_SolarHeatGainCoefficient:V01_AL2_Vetro_doppio (10mm)_Var.29	923.2384615	262.4444237	325.5571603	-923.2384615	-923.2384615	801	328.0555297
22	Interior floor Conductivity:M16_Solaio_in_laterocemento (250mm)_Var.22	794.3748252	585.0630425	1096.911948	377.5748252	377.5748252	655.4423077	731.3288031
20	Exterior wall Conductivity:Duplicated M12_Mattoni_forati (80mm) for C02_CV2_Tamponatura_con_fini_Var.20	391.31	138.7579934	168.5711252	391.31	391.31	483.75	173.4474917
26	Window_UFactor:V04_PVC_Doppio_vetro (10mm)_Var.26	299.6130769	75.43901994	107.325901	299.6130769	299.6130769	337.5	94.29877492
7	Heating_set_point:Z03_Flat_3L_Var.7	342.82	54.12410958	74.55186617	-342.82	-342.82	330.75	67.65513698
27	Window_SolarHeatGainCoefficient:V04_PVC_Doppio_vetro (10mm)_Var.27	230.741958	68.95055548	88.61415842	-230.741958	-230.741958	222.75	86.18819435
9	Metabolic rate:Z02_Flat_3R_Var.9	214.9746154	62.2575414	70.16009749	-214.9746154	-214.9746154	200.75	77.82192675
23	Interior walls Conductivity:M11_Mattone_forato (110mm)_Var.23	102.1107692	55.19038858	70.17512642	102.1107692	102.1107692	88	68.98798573
21	Ground_floor Conductivity:M03_Ground_floor_generic (300mm 0.6)_Var.21	41.7	21.11700197	27.61645796	-41.7	-41.7	46.75	26.39625247
24	Interior walls Conductivity:M11_Mattone_forato (70mm)_Var.24	39.59545455	20.22117013	26.49260383	-39.59545455	-39.59545455	38.25	25.27646267
12	Exterior wall Density:M12_Mattoni_forati (80mm)_Var.12	75.45384615	78.46496872	100.198666	-75.45384615	-75.45384615	29.5	98.08121091
17	Interior walls Density:M11_Mattone_forato (70mm)_Var.17	34.31363636	36.5749972	52.76123123	-34.31363636	-34.31363636	21.06818182	45.7187465
15	Interior floor Density:M16_Solaio_in_laterocemento (250mm)_Var.15	17.91958042	3.798938038	20.35168249	3.01958042	3.01958042	17.48076923	4.748672548
25	Roof Conductivity:M09_Ladrillo_hueco_12 (120mm 0.6)_Var.25	18.91328671	9.798299728	13.29448956	18.91328671	18.91328671	12.63461538	12.24787466
0	Infiltration_rate:Z01_Flat_2R_Var.0	6.736013986	8.141705127	10.89731229	6.736013986	6.736013986	2.045454545	10.17713141
13	Exterior wall Density:Duplicated M12_Mattoni_forati (80mm) for C02_CV2_Tamponatura_con_fini_Var.13	2.33951049	1.660368811	2.957756097	-1.63951049	-1.63951049	1.75	2.075461014
14	Ground_floor Density:M03_Ground_floor_generic (300mm 0.6)_Var.14	0.945454545	0.462850218	1.135727094	0.345454545	0.345454545	1	0.578562772
16	Interior walls Density:M11_Mattone_forato (110mm)_Var.16	1.348251748	0.960224731	1.810108066	0.520979021	0.520979021	1	1.200280914
3	Infiltration_rate:Z04_Flat_4R_Var.3	6.25	7.660913358	9.979666828	6.25	6.25	0.75	9.576141697
18	Roof Density:M09_Ladrillo_hueco_12 (120mm 0.6)_Var.18	0.51	0.384606201	0.702139587	0.19	0.19	0.5	0.480757751
2	Infiltration_rate:Z03_Flat_3L_Var.2	0.184615385	0.267158462	0.324823575	0.184615385	0.184615385	0	0.333948077

Figure 10: Sensitivity analysis results in the *Sensitivity_results.xlsx* file. The file reports the main results of the Morris method in terms of mu_star, sigma, and median_star of the Elementary effect. The confidence interval is also reported for each quantity. The ranking of the parameters is carried out in terms of median_star values, according to Section 3.4.3.

Only the group of most important parameters are considered in the next calibration step. This group is composed of the first parameters (listed from the highest to the lowest median_star) whose sum of their median star is equal to 90% of the total effects. Other parameters are considered non-influential or not significant and will be discarded for the optimization problem of the next step.

4.4 Calibration

Before starting the optimization process, the calibration data (measured outputs) should be inserted in the *calibration_data.xlsx* file saved by the code in the IDF folder during the sensitivity analysis. This file contains a void editable column for each output specified in the IDF. To facilitate the search of the output variable, the excel file is subdivided into four sheets, one for each reporting frequency (Timestep, Hourly, Daily, Monthly, and RunPeriod). In particular, the RunPeriod frequency can be useful to consider the actual bills reporting period (to be specified in the IDF), which can be different from monthly periods. Moreover, if the CoP is specified in the configuration file, the simulation output related to the energy demand is converted into energy consumption by using the CoP as heating system efficiency. Then, energy consumption data should be inserted in the related column (such as energy consumption from bills).

In Figure 11, an example of the *calibration_data.xlsx* file is reported. It should be noted that not all the void columns must be filled by the user, but only those for which the measured data are available. For each filled column, the software computes a CV(RMSE) from the comparison between predicted and simulated output, allowing to both perform single-objective (one filled column) and multi-objective (more than one filled column) optimization problems. The multi-objective optimization is reduced to a single-objective problem, considering the average of all the CV(RMSE) as error functions, to increase the efficacy and accuracy of the optimization algorithm.

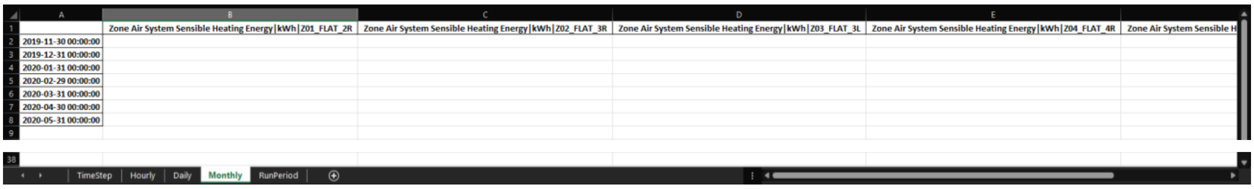
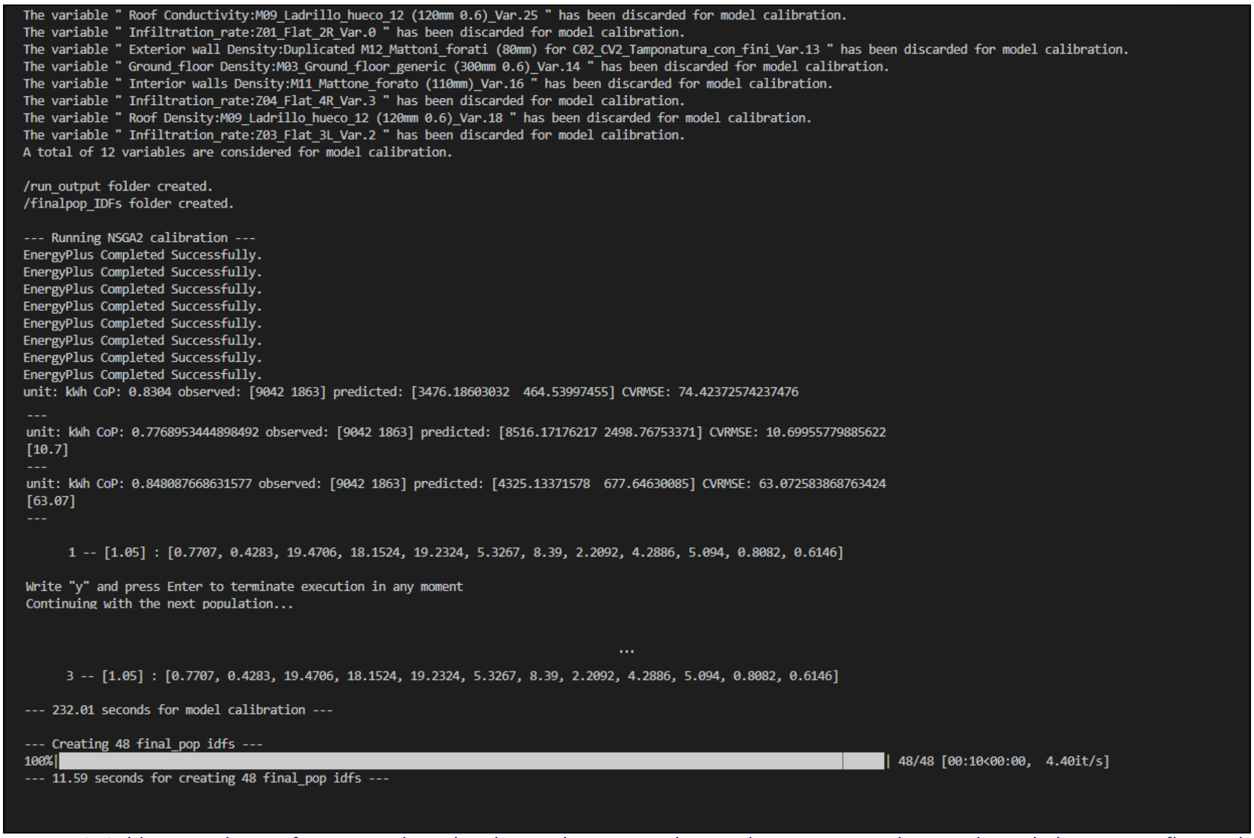


Figure 11: An example of *Calibration_data.xlsx*. It is composed of different sheets, one for each frequency (Timestep, Hourly, Daily, Monthly, Yearly, and RunPeriod). Within each sheet, the outputs set in the IDF file are reported with void fields (void time series) to be filled in case of real data availability.

Once the calibration_data.xlsx data is filled, the user can start the calibration process by clicking on the *Calibration* button. In Figure 12, the information plotted by the software tool in the window terminal during the calibration phase is reported. First, the names of the input parameters that are discarded from sensitivity are plotted as well as the number of variables considered in the optimization process. Then, for each simulation, the predicted and observed output are plotted along with the related CV(RMSE). When all the candidate solutions of a specific population are evaluated, the best solution with the lowest CV(RMSE) is plotted along with the related set of input parameter values (chromosome). Then, a new generation is created, and the evaluation process restarts. The optimization process ends when the maximum number of generations specified in the configuration file is reached or when the user chooses to terminate the execution.

Then, the IDFs belonging to the last population (optimized BEMs) are created and saved in the finalpop_IDFs folder created within the IDF folder. For each optimized BEM, the list of optimized property values is reported in the *final_pop.xlsx* file along with the obtained CV(RMSE). Then, the user can identify a calibrated BEM according to the obtained CV(RMSE).



```
The variable " Roof Conductivity:M09_Ladrillo_hueco_12 (120mm 0.6)_Var.25 " has been discarded for model calibration.
The variable " Infiltration_rate:Z01_Flat_2R_Var.0 " has been discarded for model calibration.
The variable " Exterior wall Density:Duplicated M12_Mattoni_forati (80mm) for C02_CV2_Tamponatura_con_fini_Var.13 " has been discarded for model calibration.
The variable " Ground_floor Density:M03_Ground_floor_generic (300mm 0.6)_Var.14 " has been discarded for model calibration.
The variable " Interior walls Density:M11_Mattone_forato (110mm)_Var.16 " has been discarded for model calibration.
The variable " Infiltration_rate:Z04_Flat_4R_Var.3 " has been discarded for model calibration.
The variable " Roof Density:M09_Ladrillo_hueco_12 (120mm 0.6)_Var.18 " has been discarded for model calibration.
The variable " Infiltration_rate:Z03_Flat_3L_Var.2 " has been discarded for model calibration.
A total of 12 variables are considered for model calibration.

/run_output folder created.
/finalpop_IDFs folder created.

--- Running NSGA2 calibration ---
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
EnergyPlus Completed Successfully.
unit: kWh CoP: 0.8304 observed: [9042 1863] predicted: [3476.18603032  464.53997455] CVRMSE: 74.42372574237476
 ---
 unit: kWh CoP: 0.7768953444898492 observed: [9042 1863] predicted: [8516.17176217 2498.76753371] CVRMSE: 10.69955779885622
 [10.7]
 ---
 unit: kWh CoP: 0.848087668631577 observed: [9042 1863] predicted: [4325.13371578  677.64630085] CVRMSE: 63.072583868763424
 [63.07]
 ---

      1 -- [1.05] : [0.7707, 0.4283, 19.4706, 18.1524, 19.2324, 5.3267, 8.39, 2.2092, 4.2886, 5.094, 0.8082, 0.6146]

 Write "y" and press Enter to terminate execution in any moment
 Continuing with the next population...

                                                   ...

      3 -- [1.05] : [0.7707, 0.4283, 19.4706, 18.1524, 19.2324, 5.3267, 8.39, 2.2092, 4.2886, 5.094, 0.8082, 0.6146]

 --- 232.01 seconds for model calibration ---

 --- Creating 48 final_pop idfs ---
 100%|████████████████████████████████████████| 48/48 [00:10<00:00,  4.40it/s]
 --- 11.59 seconds for creating 48 final_pop idfs ---
```

Figure 12: Calibration phase information plotted in the window terminal. First, the parameters that are discarded, since uninfluential, are listed in the terminal. Then, for each simulation and each output to be calibrated, the unit of measure, the related CoP (if applicable), along with observed and predicted values, and the obtained CV(RMSE) are reported. At the end of each generation, the best IDF is plotted in terms of CV(RMSE) (the average if there is more than one output considered for calibration) and input values. When the maximum number of generations specified in the config_files is reached, or when the user decides to stop the calibration, the optimized IDFs of the last generation and the summary excel table (final_pop.xlsx) with input values and average CV(RMSE) are plotted.

	Infiltration_rate:Z(	Ventilation_rate:Z	Heating_set_point	Heating_set_point	Heating_set_point	Lights:Z02_Flat_3R	OtherEquipment:Z	Exterior wall Cond	Window_UFactor:	Window_UFactor:	Window_SolarHea	CoP_Var.32	CVRMSE (%) | Index n.0
0	0.770650969	0.428339847	19.47057702	18.15238707	19.23241199	5.326730533	8.389993929	2.209177888	4.288595687	5.093984927	0.808244292	0.614599364	1.049432771
1	0.589042242	0.391359463	19.02482204	19.27662328	21.82904886	8.771339453	8.954538286	2.330987583	3.259658317	5.680211328	0.738480389	0.632321502	2.45943561
2	0.600204136	0.328375271	21.53875338	19.60580843	19.70634526	5.143076309	4.457364207	2.310371549	5.121982483	1.892169244	0.926396984	0.703572456	2.933028563
3	0.319650735	0.440941672	21.37918288	20.47239589	21.89191356	12.25342936	6.559755196	1.87421597	2.138190352	4.891641338	0.949413387	0.644802661	3.093199501
4	0.313913898	0.715322173	19.88495075	19.08827319	18.7431726	4.268978396	4.187723456	1.163292589	3.781189209	4.642087908	0.797525109	0.695491458	3.135033726
5	0.280079733	0.212421225	19.80286347	20.11708107	19.53943383	12.8460307	8.674031926	1.853180768	4.533253862	4.843811964	0.948328253	0.648103759	4.414757334
6	0.4701	0.7413	18.3913	19.9565	19.087	11.5859	10.5359	0.575	6.9304	3.1739	0.6379	0.6826	4.839256938
7	0.671066702	0.463542221	18.58154111	18.17154798	19.54517436	6.314326232	5.230838739	1.586676086	1.198645507	3.585100223	0.311324645	0.610291856	5.180699875
8	0.5842	0.5717	19.2609	19.3478	21.7826	8.788	8.9511	2.325	0.6696	5.6783	0.5349	0.6304	5.209415712
9	0.634028206	0.442034539	19.1659189	19.98091424	20.02106735	10.47950218	2.895376836	0.402983425	5.582531938	5.455762674	0.654554665	0.612986097	5.408749388
10	0.982163	0.599038525	20.86113477	19.92513086	19.87438368	6.247193574	8.562455924	0.980062597	3.310045494	3.609846811	0.658889082	0.785710005	5.664442819
11	0.454504327	0.606830166	19.45887062	18.11147411	19.53810937	7.454830947	2.715032288	1.441350568	4.219657583	3.568540833	0.305400139	0.608455556	5.711045461
12	0.310221437	0.712734376	19.7251888	19.1144331	21.24736533	8.776031799	3.938792168	1.212320259	6.901583671	4.647380324	0.244996452	0.695219014	6.007399951
13	0.6332	0.7804	21.4348	19.6957	19.4348	9.8641	4.3728	1.275	5.1217	1.9217	0.6895	0.7087	6.068808744
14	0.307949814	0.601112233	19.94111541	19.97446572	19.94876973	10.86462854	4.293159005	1.1489337	3.399495044	5.444588517	0.652490742	0.726675228	6.244396673
15	0.26824779	0.560485568	19.74205801	19.00788882	21.26448538	7.348314603	3.93749229	1.665260219	6.895879202	5.021661184	0.247093677	0.691221314	6.622388615
16	0.274959857	0.245521327	19.11801315	20.14042722	20.12192978	12.71570051	2.873430499	1.015200962	6.085757977	4.143784537	0.899490493	0.665331146	7.169202777
17	0.683047755	0.390353451	19.23456921	21.0555587	19.99598516	10.59528791	6.178931684	0.449695009	3.588797895	5.511615281	0.720553203	0.71914927	7.984120387
18	0.513432892	0.418089592	20.21441861	20.4400424	19.72215222	7.495599486	7.165591335	2.317674834	2.032618175	3.84022586	0.947878032	0.745750197	7.988353376
19	0.983110754	0.600291976	19.09570398	19.98068956	20.00076884	10.63708148	6.318587608	0.404757606	3.307543389	5.530162092	0.655642286	0.726559275	8.075581899
20	0.776300117	0.432827678	20.40534316	20.43608058	19.18046976	5.148869416	9.054442122	2.147749987	1.453963617	5.109533761	0.831776267	0.890463275	8.188784304
21	0.719831781	0.756456928	19.77928082	20.57457487	19.51227437	10.5299917	7.380857852	1.866251538	6.2803823	4.855952403	0.946687105	0.952495595	8.231260643
22	0.2908	0.2326	19.7826	20.1304	19.5217	13.0924	3.3163	1.875	5.8174	4.9826	0.9471	0.6478	8.701639613
23	0.6000153	0.779792074	18.85920255	20.48735928	19.68567076	10.2344816	7.174048994	2.31502837	5.316269005	3.892969281	0.899948719	0.959958289	8.717779494
24	0.718431225	0.284989929	18.82666069	19.69097309	19.3582621	3.942848565	5.931070526	1.728548964	6.230805483	3.021941052	0.933819166	0.856122244	8.754492319
25	0.717946524	0.766982003	18.85976324	20.56963237	19.34823286	10.55453017	7.14751952	1.668577333	6.286174487	3.880110333	0.903867885	0.950880533	8.774847894
26	0.645123341	0.784972353	21.38176019	19.08535263	19.38577849	9.883225412	4.194735595	1.269782448	5.49983134	1.878209406	0.789450281	0.709566584	9.060765447
27	0.6168	0.4283	21.7826	19.8696	19.2609	5.5598	8.0707	0.675	4.287	4.8435	0.3116	0.613	9.423057887
28	0.686424826	0.390360098	21.10613614	21.08170487	19.77704295	6.661797024	7.189259455	1.088249597	5.272518267	3.890771599	0.717172561	0.885424197	9.463932966
29	0.731	0.4413	21.5217	20.4783	21.9565	12.2315	6.4859	1.825	2.4783	6.513	0.5005	0.7174	9.549204089
30	0.585553054	0.391013446	19.41680388	18.15471186	19.51122001	5.303002218	8.909331537	2.329356612	3.299461546	5.702356081	0.738779581	0.615950914	9.612057356
31	0.682126632	0.38950531	19.0214595	21.03073954	21.83026019	6.140110196	8.421346044	1.02200165	3.205836879	4.024557498	0.741392275	0.795747895	9.628743238
32	0.633495864	0.749824729	18.72214662	19.69859157	19.42415474	8.343823691	5.873035793	1.339628451	5.260617517	3.064330756	0.693003016	0.870327544	9.773579174
33	0.598606873	0.785432957	18.7559584	20.47206995	20.4540286	9.997214017	8.445311387	2.298744678	4.501383555	3.663863623	0.904063762	0.958538747	10.27952831
34	0.7769836	0.303999018	20.91010801	21.80318446	19.88939922	6.201570564	9.617678152	0.328034095	3.559767183	4.35573406	0.623312397	0.776895344	10.6995578
35	0.6821	0.3891	21	21	19.8696	6.2054	8.4228	1.025	3.5913	3.5913	0.7238	0.7783	10.72408482
36	0.6850164	0.396000982	21.00289199	21.06641554	20.54540078	6.424529436	8.460521848	1.071965905	4.457632817	3.66166594	0.721287603	0.884004656	10.85626924
37	0.284469175	0.718913258	19.22674291	19.1044701	18.70783776	9.858720556	6.091199572	1.023985535	3.583692672	4.716000267	0.719274367	0.694992585	10.97265441
38	0.67728439	0.569845846	19.25753746	21.10191625	21.78381132	6.156770744	8.417907757	1.016014067	0.615778562	4.02264617	0.537811886	0.793826393	10.99112459
39	0.288914743	0.243125075	18.86308242	20.12545751	19.35765849	13.11693847	3.082961668	1.677325796	5.823192187	4.00675793	0.90428078	0.646184938	11.1752708
40	0.977151498	0.709929563	19.09257913	19.70484442	18.84419322	4.022305928	6.391216465	0.424334048	3.311106214	4.771152077	0.700726372	0.694508922	11.25216879
41	0.976602816	0.606736212	21.6921216	19.88269425	19.22286767	5.963748602	6.136178846	0.698593414	3.296157588	5.056687454	0.311459368	0.726269419	11.41231217
42	0.724082411	0.765039052	19.18862582	20.57950494	19.47378984	10.48278465	7.423604225	1.992168098	5.56003133	5.389638805	0.657569619	0.943146854	11.58569681
43	0.519	0.3761	20.2174	19.4348	18.3043	7.0663	7.8946	1.475	2.0609	5.9565	0.449	0.7522	11.61353293
44	0.6984	0.5587	19.9565	20.0435	18.0435	5.3446	4.0207	0.475	6.2348	6.3739	0.2945	0.7609	11.71611131
45	0.307579076	0.717520093	18.22740386	20.73004024	21.55271309	8.738323776	3.866168562	1.21802895	6.545755618	4.65849687	0.203361614	0.953198144	12.0078512
46	0.463957189	0.224789521	20.40165518	21.64771526	19.45067639	7.213694174	7.199814934	0.922604992	6.050153696	4.191202869	0.506669828	0.888228237	12.05148933
47	0.265605429	0.565271284	18.24427308	20.62349596	21.56983314	7.31060658	3.864868685	1.67096891	6.54005115	5.03277773	0.205458839	0.949200445	12.16814796

Figure 13: Example of *final_pop.xlsx* file. For each IDF of the last generation (rows), the input values, and the related “fitness” are reported.

5. An exemplary application of BEM calibration

5.1 Description of the case study and aim of the calibration

The considered demo is a six-story U-shaped building built in 1958 and located in a densely urbanized context in the city of Vitoria-Gasteiz, Spain. Figure 14 reports the aerial photo of the site with an indicative view of the urban context, while Figure 15 shows the main and rear façades of the demo building. The building has a footprint, consisting of 2 apartments on each floor (for a total of 8 apartments), except for the ground floor, where a parking lot access and a bar are present, and for the last floor, which is a not-heated attic with a sloped roof. Figure 16 reports the plan of a representative floor.

The architectural and constructive characteristics of the building are consistent with typical buildings constructed in the same period, characterized by hollow brick cavity walls with no insulation [44]. The building is then characterized by poor energy performance, also showing condensation and humidity damp issues. Each apartment is equipped with an autonomous heating system (gas boiler) used for both space heating and DHW production having a rated capacity of 24kW. No cooling systems or mechanical ventilation systems are installed. A summary of the demo general data is reported in Table 7.

The objective of the BC is to calibrate the BEM of the selected demo case with respect to both the heating system energy consumption and the indoor air temperatures of the left-side third-floor apartment (3R apartment, see Figure 16).

a b

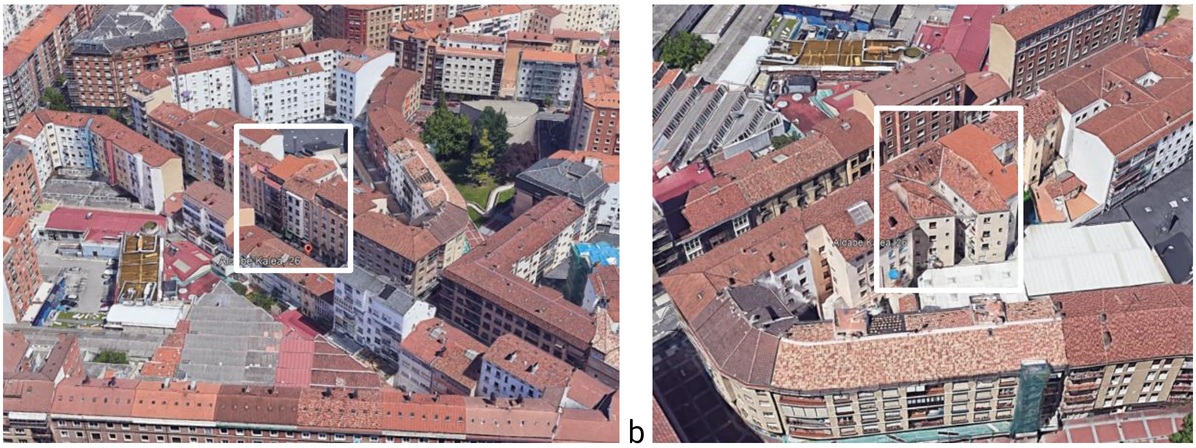


Figure 14: Aerial views of the urban context and building location. a) main facade; b) rear façade.

Figure 15 a) main and b) rear façade of the building case study. The main façade is southwest-oriented.

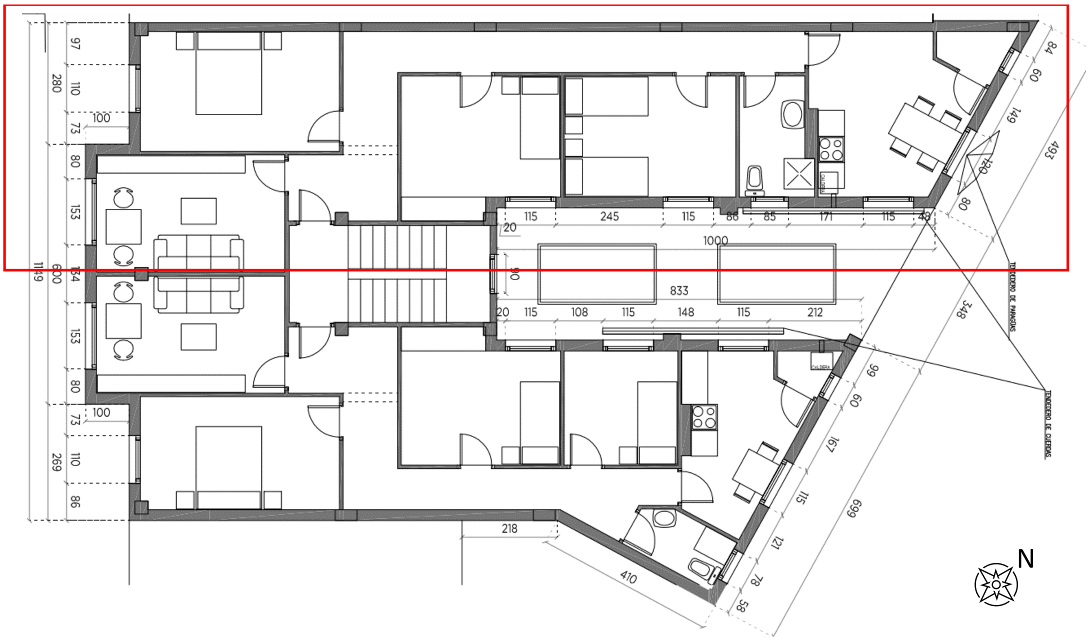


Figure 16. Third-floor plan. The apartment object of calibration (3R) is highlighted.

Table 7: General information of the Vitoria/Gasteiz democase.

General information	
Location	Vitoria-Gasteiz (Spain)
Use category	Residential
Building type	Multi-family house
Construction year	1958
Renovation year	2020
Number of floors	6
Number of apartments/units	8

5.2 Data gathering

The energy consumption of the heating system and the hourly indoor air temperature of the living room have been collected from 21 November 2019 to 30 June 2020, thus considering all the relevant seasons. Concerning energy consumption, the monitoring period is covered by three different bills, whose energy consumption and reporting periods are shown in Table 8. From May, no space heating was required in the apartment due to the high outdoor air temperatures. Then, the energy consumption of the last bill (from 21/05/2020 to 30/06/2020) has been used to estimate a constant average daily energy consumption for DHW (i.e. about 5.7 kWh per day). This value has been used to compute by subtraction from the total heating system energy consumption the net energy consumption for space heating only to be used as a calibration target (see Table 8).

Table 8: Energy consumption for space heating and DHW.

Reporting period	Energy consumption for space heating and DHW (kWh)	Energy consumption for space heating (kWh)
from 21/11/2020 to 13/03/2020	6907	6604
from 14/03/2020 to 20/05/2020	1863	1475
from 21/05/2020 to 30/06/2020	263	-

To identify the internal gains schedule related to electric equipment, hourly electricity consumption has been also monitored as short-term measurements from 14 to 18 November. Finally, the weather data related to the same monitoring period has been obtained from the MEREEN tool and then merged for creating a single EPW (Energy Plus Weather File) to be used in the numerical simulations.

5.3 Building energy modeling

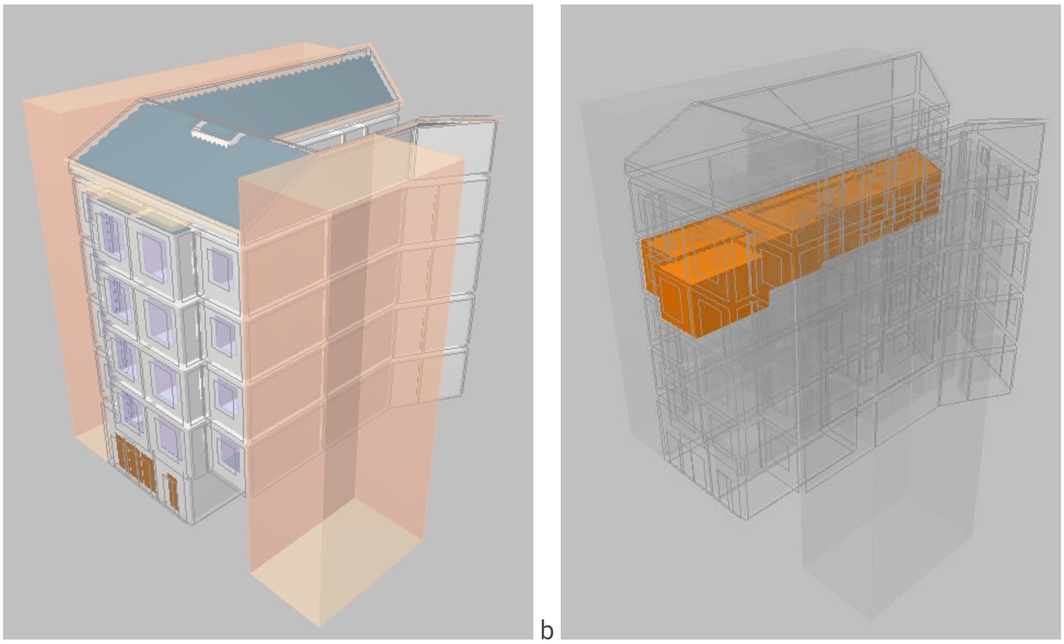
Figure 17a provides the 3D graphical representation of the original BEM as completed in CypeTherm Eplus. Each apartment is modeled as a single thermal zone, while near buildings in front of the main and rear façades are modeled as additional external surfaces casting shadows on the building. Since the calibration process will be focused on a single apartment only (see Figure 17b and Figure 16), the BEM has been simplified to speed up the calibration process by considering in the model only the target thermal zone and the surrounding ones.

The elements, as well as the single materials, have been created and stored in structured libraries. Table 9 summarises all the materials implemented within the BEM, while Table 10 and

Table 11 report the construction systems and windows properties, respectively.

Internal gains in terms of maximum power densities are summarized in

Table 12. Occupational patterns have been assumed based on standard residential uses and some information collected from the users, while internal gains schedules are defined based on the electricity consumption collected data. The heating setpoint has been assumed to vary according to the monitored indoor air temperatures, which are then used as an input of the model for the heating season [9].



a b

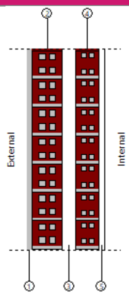
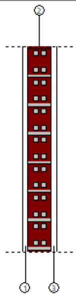
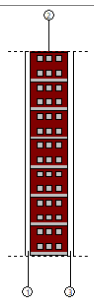
Figure 17: 3D graphical representation of the Frigento BEM.

Table 9: Materials

Layers					
Material	e	ρ	λ	RT	Cp
Cementitiousplaster	1.50	1000.00	1.400	0.01	1000.00
Hollow bricks 9cm	9.00	1000.00	0.500	0.18	1000.00
Holow bricks 7cm	7.00	1000.00	0.500	0.14	1000.00
Gypsum plaster	1.50	1000.00	0.300	0.05	1000.00
Hollow bricks 11cm	11.00	1000.00	0.770	0.14	1000.00
Hollow bricks 7cm	7.00	1000.00	0.500	0.14	1000.00
Teja Ceramica	1.00	35.00	1.000	0.01	1000.00
Mort. cemento 5	5.00	2000.00	1.250	0.04	1000.00
Ladrillo hueco 12	12.00	100.00	0.500	0.24	1000.00
Tiles	1.50	1000.00	0.210	0.07	1000.00
Screed	5.00	1000.00	1.400	0.04	1000.00
Brick concrete slab	25.00	1000.00	1.316	0.19	1000.00
Ground floor - generic	30.00	1000.00	1.429	0.21	1000.00

Used abbreviations			
e	*Thickness cm*	RT	*Thermal resistance (m²·K)/W*
ρ	*Density kg/m³*	Cp	*Specific heat J/(kg·K)*
λ	*Thermal conductivity W/(m·K)*		

Table 10: Main construction systems for the target apartment.

1.1 Façades		
CV1_External wall_23cm		Layer list: 1 - Cementitious plaster 1.50 cm 2 - Hollow bricks 9cm 9.00 cm 3 - Air gap 4.00 cm 4 - Holow bricks 7cm 7.00 cm 5 - Gypsum plaster 1.50 cm Properties Thermal transmittance, U: 0.87 W/(m²·K) Total thickness 23.00 cm
2.1 Internal vertical partitioning		
PV1_Internal partition_10cm		Layer list: 1 - Gypsum plaster 1.50 cm 2 - Hollow bricks 7cm 7.00 cm 3 - Gypsum plaster 1.50 cm Properties Thermal transmittance, U: 2.00 W/(m²·K) Total thickness 10.00 cm
PV2_Internal partition_14cm		Layer list: 1 - Gypsum plaster 1.50 cm 2 - Hollow bricks 11cm 11.00 cm 3 - Gypsum plaster 1.50 cm Properties Thermal transmittance, U: 1.99 W/(m²·K) Total thickness 14.00 cm
2.2 Internal horizontal partitioning		

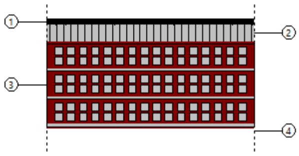
PO1_Interstorey floor_33cm	 Layer list: 1 - Tiles 1.50 cm 2 - Screed 5.00 cm 3 - Brick concrete slab 25.00 cm 4 - Gypsum plaster 1.50 cm Properties Thermal transmittance, U: 1.83 W/(m²·K) Total thickness 33.00 cm

Table 11: Windows properties.

AV3 PVC double glass	Thermal transmittance, U: 2.98 W/(m²·K) Solar factor, g: 0.700 Opaque fraction, Ff: 0.200
AL2_Double glass aluminum	Thermal transmittance, U: 3.64 W/(m²·K) Solar factor, g: 0.700 Opaque fraction, Ff: 0.200
AL1_Single glass aluminum	Thermal transmittance, U: 5.72 W/(m²·K) Solar factor, g: 0.700 Opaque fraction, Ff: 0.200

Table 12: Initial internal gains features

OCCUPIED Space	Infiltration rates	INTERNAL GAINS (maximum value)	PEOPLE	ACTIVITY level (maximum value)
All apartments	0.5 ACH	8 W/m^2	30 m^2/person	120 W/person

5.4 Calibration procedure

5.4.1 Phases

The calibration process is subdivided into three main phases (two-staged calibration [9]):

- first, a sensitivity analysis is carried out to identify the most important parameters and discard uninfluential ones from the calibration process;
- then, the BEM model has been calibrated in terms of indoor air temperatures (Phase 1 calibration) considering the first week of May, i.e. when the target flat (3R) operated in free-floating conditions;
- finally, starting from the calibrated model obtained in the previous phase, the calibration is repeated on energy consumption (Phase 2 calibration) to find the CoP value that provides the best fit for the energy consumption for space heating obtained from bills.

5.4.2 Sensitivity/screening analysis

The sensitivity/screening analysis is carried out considering the indoor air temperatures as model output and is related to the first week of May. The considered ranges of variation for each parameter selected by the tool during the IDF analysis phase are reported in Table 13. The results of the Morris method, carried out with respect to indoor air temperatures, are reported in Figure 18, where the most and less important parameters in the calibration process can be identified.

Table 13: Screened variables and related range of variation.

ID	Zone/Layer	Parameter	Range of variation	Initial value	Ref.
0	Flat_2R	Infiltration_rate (ACH)	[0.3, 1.2]	0.5	[45]
1	Flat_3R	Infiltration_rate (ACH)	[0.3, 1.2]	0.5	[45]
2	Flat_3L	Infiltration_rate (ACH)	[0.3, 1.2]	0.5	[45]
3	Flat_4R	Infiltration_rate (ACH)	[0.3, 1.2]	0.5	[45]
4	Flat_3R	Metabolic rate (W/m^2)	[80, 140]	120	[46]
5	Flat_3R	Electrical internal gains (W/m^2)	[3, 12]	8	Exp.
6	Ext. wall Hollow_bricks (90mm)	Density (kg/m^3)	[800, 1600]	1000	[47]
7	Int. floor Brick-concrete slab (250mm)	Density (kg/m^3)	[1800, 2000]	1000	[47]
8	Int. walls Hollow_bricks (110mm)	Density (kg/m^3)	[800, 1600]	1000	[47]
9	Int. walls Hollow_bricks (70mm)	Density (kg/m^3)	[800, 1600]	1000	[47]
10	Ext. walls Hollow_bricks (90mm)	Conductivity (W/mK)	[0.18, 0.5]	0.5	[35]
11	Int. floor Brick concrete slab (250mm)	Conductivity (W/mK)	[0.475, 0.76]	0.48	[35]
12	Int. walls Hollow_bricks (110mm)	Conductivity (W/mK)	[0.18, 0.5]	0.3	[35]
13	Int. walls Hollow_bricks (70mm)	Conductivity (W/mK)	[0.18, 0.5]	0.3	[35]
14	PVC_Double_glass (10mm)	Window_Ufactor (W/m^{2}K)	[2.9, 3.3]	3.1	[35,48]
15	PVC_Double_glass (10mm)	Window_SHGC (-)	[0.55, 0.99]	0.7	Assumed
16	AL2_ Double_glass (10mm)	Window_Ufactor (W/m^{2}K)	[3.4, 4.5]	4.12	[35,48]
17	AL2_ Double_glass (10mm)	Window_SHGC (-)	[0.55, 0.99]	0.7	Assumed
18	AL1_Single_glass (10mm)	Window_Ufactor (W/m^{2}K)	[4.6, 6.10]	5	[35,48]
19	AL1_ Single_glass (10mm)	Window_SHGC (-)	[0.55, 0.99]	0.7	Assumed

20	Z02_Flat_3R	CoP (-)	[0.6, 0.9]	0.7	Assumed

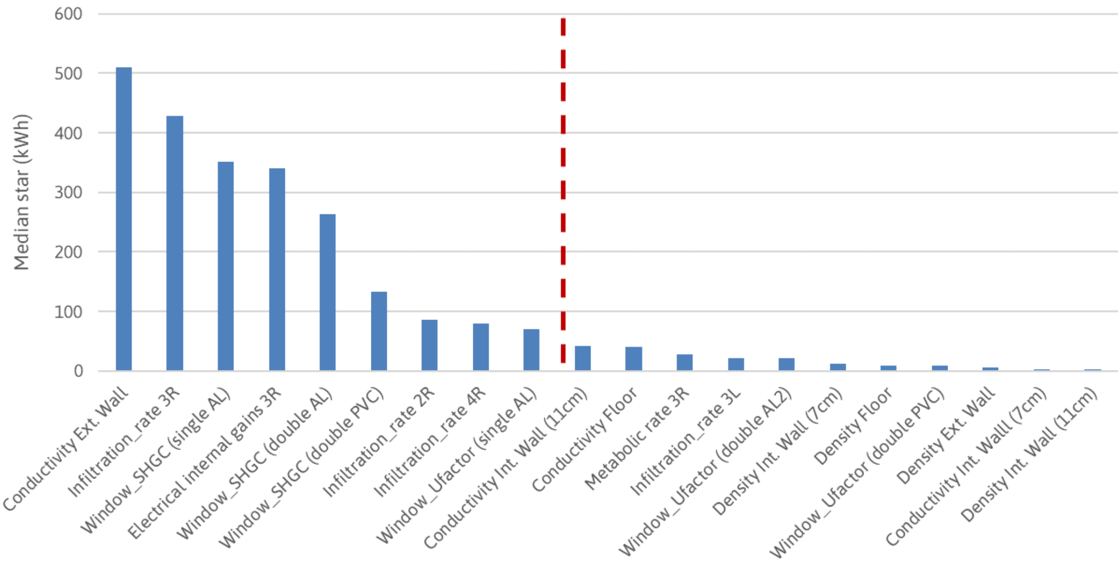


Figure 18: Results of the Morris method in terms of the median_star of the elementary effects on indoor air temperatures. The parameters on the right of the red dotted lines can be considered as not influential in the calibration process.

5.4.3 Calibration of indoor air temperatures

The results of the automated calibration process carried out with respect to indoor air temperatures are reported in Figure 19 in terms of indoor air temperature, and in **Error! Reference source not found.** in terms of the calibrated values (from most to less important according to Morris screening).

The most important calibrated input values, listed in **Error! Reference source not found.**, are here briefly commented. In particular, the conductivity value of the hollow bricks layer of the vertical wall passes from 0.4 W/mK of the not calibrated model to 0.18 W/mK of the calibrated one. This leads to a U-value of the vertical walls (0.68 W/m^{2}K) lower than that initially assumed (0.87 W/m^{2}K), which means a lower heat exchange for conductivity between indoor and outdoor environments. Conversely, the infiltration rate passes from 0.5 to 0.63 ACH, showing a slight increase in the heat exchanged with the outdoor air for convection. Similarly, the window SHGCs denote higher values than that initially assumed, which increase the heat gain during the day to have similar temperature maximum values with experimental results. The obtained values are also in line with those obtained for similar windows in the literature [49]. Finally, the electric internal gains are reduced from 8.0 to 3.66 W/m^2.

Table 14: Calibrated values of unobserved parameters, ordered from most to less important according to the sensitivity results.

ID	Zone/Layer	Parameter	Initial values	Calibrated values

10	Ext. walls Hollow_bricks (90mm)	Conductivity (W/mK)	0.4	0.18
1	Flat_3R	Infiltration_rate (ACH)	0.5	0.63
19	AL1_ Single_glass (10mm)	Window_SHGC (-)	0.7	0.95
5	Flat_3R	Electrical internal gains (W/m^2)	8.0	3.66
17	AL2_ Double_glass (10mm)	Window_SHGC (-)	0.7	0.97
15	PVC_Double_glass (10mm)	Window_SHGC (-)	0.7	0.70
0	Flat_2R	Infiltration_rate (ACH)	0.5	1.14
3	Flat_4R	Infiltration_rate (ACH)	0.5	1.06
18	AL1_Single_glass (10mm)	Window_Ufactor (W/m^{2}K)	5.0	5.72

Concerning the error metrics, the CV(RMSE) obtained from the comparison of the numerical and experimental air temperature datasets passes from 6.7% for the original BEM to 3.4% for the calibrated one, corresponding to an RMSE passing from 1.52 to 0.76°C. This value denotes a good level of accuracy of the calibrated BEM, especially if compared to the values obtained in literature for similar buildings calibrated toward indoor air temperatures where calibrated BEMs achieve a minimum RMSE of 0.9 °C [16–19] (see Section 2.3). Some differences can still be noted between the two datasets even after calibration, especially in terms of peak maximum indoor air temperatures. This difference can be mainly attributable to different occupants' behaviors during the different days, with respect to the window opening and indoor activities, that could be better investigated to obtain a more accurate result.

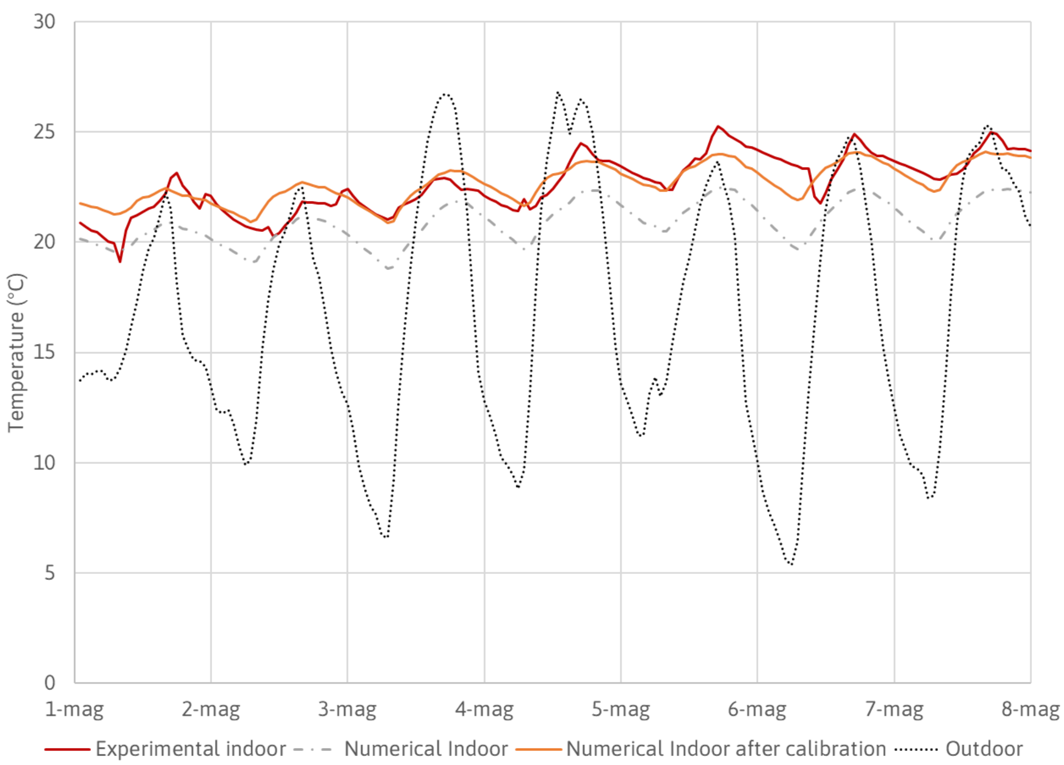


Figure 19: Results of the automated calibration process in terms of indoor air temperatures.

5.4.4 Calibration of energy consumption

Finally, the results of the automated calibration process carried out with respect to energy consumption (Phase 2 calibration) are reported in Table 15, obtained by simply modifying the remaining, not calibrated CoP value. In particular, the CV(RMSE) obtained from the comparison of the numerical and experimental datasets passes from 48.1% for the original BEM to 10.3% for the calibrated one (assuming a CoP equal to 0.83). This result denotes the higher level of BEM accuracy achieved through the calibration process also in terms of energy consumption.

Table 15: Energy consumption for space heating and DHW.

Reporting period and CV(RMSE)	Experimental	Original BEM	Phase 1 Calibration	Phase 2 Calibration
from 21/11/2019 to 13/03/2020	6604 kWh	9190 kWh	7630 kWh	6427 kWh
from 14/03/2020 to 20/05/2020	1475 kWh	2399 kWh	2416 kWh	2034 kWh
CV(RMSE)	-	48.1%	24.4%	10.3%

6. Conclusion

In energy renovation of existing buildings, BEMs are generally used to estimate ECMs energy savings, verify the compliance of ECM with the requirements set by the National Standards, and then select the best ECM among different available options. However, a significant discrepancy called the "energy performance gap", is often found between simulated and measured energy use, which can be traced back to the difficulty in obtaining the exact values of all the thousands of inputs needed for characterizing a BEM.

A BEM with inaccurate input data and/or inaccurate energy predictions may lead to the design of erroneous Energy Conservation Measures (ECM) in building renovation projects. For example, if a coefficient of performance (COP) of the heating system lower than the actual one is set in the model, an ECM concerning the retrofit of the HVAC might be recommended. When put in place, however, this ECM will likely provide lackluster results in terms of energy-saving. To minimize this risk and then to have a better design of ECMs, it is of paramount importance to increase the accuracy and reliability of BEM. At this aim, a BEM calibration is generally undertaken, consisting of fine-tuning model input parameters to minimize the discrepancy between simulated and measured data.

However, to date, there is still no universal consensus on which is the best calibration procedure to be used. Indeed, while there are standard criteria for validating a calibrated model, there is still a lack of formal and recognized methodology or guidelines for BEM calibration, which makes the BEM calibration processes highly dependent on the user's skills and judgments.

In this work, an automated BEM calibration procedure is developed and applied to a BIMSPEEED demo case. The procedure, compliant with relevant BC and BIM-BEM Standards, is aimed at facilitating its application in engineering practice by simplifying and speeding up the entire calibration process, assisting the energy modelers from the data gathering to the BEM optimization process, passing through model enrichment and sensitivity analysis techniques. The procedure consists of two main phases, i.e. a data-gathering phase and an automated calibration phase. The data-gathering phase is aimed at obtaining the model inputs (weather conditions, schedule information, etc.) and outputs (e.g., energy consumption) required for the calibration process. The calibration phase is aimed, on one hand, at speeding up and simplifying the calibration process for the end-user, and, on the other hand, at increasing the predictive accuracy of the calibrated model with respect to manual approaches.

To facilitate the application of the automated calibration phase, an automatic BEM-Calibration Tool has been developed and described in the deliverable. This tool integrates, for the first time in the literature, expert knowledge, sensitivity analysis, and Artificial Intelligence (AI) optimization algorithms within the same workflow, minimizing the number of inputs required from the practitioners in BEM calibration processes while maintaining an easy-to-use interface.

An exemplary application of the developed procedure to a BIMSPEED demo case located in Spain has been also reported in this report, used to calibrate indoor air temperature and energy consumption. To do so, the energy consumption of the heating system and the hourly indoor air temperature of the living room of a target apartment have been monitored from 21 November 2019 to 30 June 2020, along with weather data, thus considering all the relevant seasons. Hourly electricity consumption from 14 to 18 November has been also collected to obtain the internal gains to be considered in the model.

The calibration process was subdivided into three main phases, i.e. a sensitivity analysis to identify the most important parameters and discard uninfluential ones from the calibration process; the BEM calibration in terms of indoor air temperatures (Phase 1 calibration) considering the first week of May, i.e. when the target flat (3R) operated in free-floating conditions; and the calibration on energy consumption (Phase 2 calibration) to find the CoP value that provides the best fit for the energy consumption for space heating obtained from bills.

Concerning the indoor air temperatures, the model error, estimated through the CV(RMSE) obtained from the comparison of the numerical and experimental datasets, passes from 6.7% to 3.4%, corresponding to an RMSE of 1.52 and 0.76°C, respectively, that can be considered acceptable according to the literature, where calibrated BEMs achieve a minimum RMSE of 0.9 °C [16–19]. Concerning the energy consumption, the CV(RMSE) passes from 48.1% for the original BEM to 10.3% for the calibrated one, denoting a good level of accuracy achieved through the calibration process. Thence, the adopted procedure and tool have been proven to be a very strong ally to practitioners to reduce model inaccuracies and then increase the effectiveness of ECM in energy renovation projects.

The entire described calibration process represents the lessons learned from BEM calibration as input for standardization in Task 5.1 “Cooperation with standardization bodies”.

7. References

[1] EU Commission, Energy performance of buildings, (n.d.). https://ec.europa.eu/energy/en/topics/energy-efficiency/energy-performance-of-buildings (accessed July 19, 2019).
[2] E. Di Giuseppe, G. Maracchini, A. Gianangeli, G. Bernardini, M. D'Orazio, Internal Insulation of Historic Buildings: A Stochastic Approach to Life Cycle Costing Within RIBuild EU Project, in: Sustain. Energy Build., 2020: pp. 349–359. https://doi.org/10.1007/978-981-32-9868-2_30.
[3] E. Di Giuseppe, M. D'Orazio, G. Du, C. Favi, S. Lasvaux, G. Maracchini, P. Padey, A Stochastic Approach to LCA of Internal Insulation Solutions for Historic Buildings, Sustainability. 12 (2020) 1535. https://doi.org/10.3390/su12041535.
[4] EUROSTAT, Energy consumption and use by households - Product - Eurostat, (2019). https://ec.europa.eu/eurostat/web/products-eurostat-news/-/DDN-20190620-1 (accessed November 20, 2020).
[5] International Energy Agency (IEA), Outlook for energy demand – World Energy Outlook 2020 – Analysis, (2020).
[6] P. De Wilde, The gap between predicted and measured energy performance of buildings: A framework for investigation, Autom. Constr. 41 (2014) 40–49. https://doi.org/10.1016/j.autcon.2014.02.009.
[7] D. Coakley, P. Raftery, M. Keane, A review of methods to match building energy simulation models to measured data, Renew. Sustain. Energy Rev. 37 (2014) 123–141. https://doi.org/10.1016/j.rser.2014.05.007.
[8] H. Yoshino, T. Hong, N. Nord, IEA EBC annex 53: Total energy use in buildings—Analysis and evaluation methods, Energy Build. 152 (2017) 124–136. https://doi.org/10.1016/j.enbuild.2017.07.038.
[9] A. Chong, Y. Gu, H. Jia, Calibrating building energy simulation models: A review of the basics to guide future work, Energy Build. 253 (2021) 111533. https://doi.org/10.1016/j.enbuild.2021.111533.
[10] T.A. Reddy, Literature Review on Calibration of Building Energy Simulation Programs: Uses, Problems, Procedures, Uncertainty, and Tools, ASHRAE Trans. 112 (2006) 226–240.
[11] E. Fabrizio, V. Monetti, Methodologies and advancements in the calibration of building energy models, Energies. 8 (2015) 2548–2574. https://doi.org/10.3390/en8042548.
[12] FEMP, M & V Guidelines: Measurement and Verification for Federal Energy Projects. Version 3.0, 2008.
[13] ASHRAE Guideline 14, Measurement of Energy, Demand, and Water Savings, (2014).
[14] EVO, IPMVP - International Performance Measurement and Verification Protocol - Concepts and Options for Determining Energy and Water Savings Volume 1, (2012).
[15] S. Martínez, P. Eguía, E. Granada, A. Moazami, M. Hamdy, A performance comparison of multi-objective optimization-based approaches for calibrating white-box building energy models, Energy Build. 216 (2020) 109942. https://doi.org/10.1016/j.enbuild.2020.109942.
[16] Y. Wang, E. Long, S. Deng, Applying passive cooling measures to a temporary disaster-relief prefabricated house to improve its indoor thermal environment in summer in the subtropics, Energy Build. 139 (2017) 456–464. https://doi.org/10.1016/j.enbuild.2016.12.081.
[17] C. Wang, S. Deng, J. Niu, E. Long, A numerical study on optimizing the designs of applying PCMs to a disaster-relief prefabricated temporary-house (PTH) to improve its summer daytime indoor thermal environment, Energy. 181 (2019) 239–249. https://doi.org/10.1016/j.energy.2019.05.165.
[18] A. O' Donovan, P.D. O' Sullivan, M.D. Murphy, Predicting air temperatures in a naturally ventilated nearly zero energy building: Calibration, validation, analysis and approaches, Appl. Energy. 250 (2019) 991–1010. https://doi.org/10.1016/j.apenergy.2019.04.082.
[19] G. Maracchini, M. D'Orazio, Improving the livability of lightweight emergency architectures: A numerical investigation on a novel reinforced-EPS based construction system, Build. Environ. 208 (2022) 108601. https://doi.org/10.1016/j.buildenv.2021.108601.
[20] G. Chaudhary, J. New, J. Sanyal, P. Im, Z. O'Neill, V. Garg, Evaluation of "Autotune" calibration against manual calibration of building energy models, Appl. Energy. 182 (2016) 115–134. https://doi.org/10.1016/j.apenergy.2016.08.073.
[21] A. Chong, K. Menberg, Guidelines for the Bayesian calibration of building energy models,

Energy Build. 174 (2018) 527–547. https://doi.org/10.1016/j.enbuild.2018.06.028.
[22] M. Fernández, B. Conde, P. Eguía, E. Granada, Parameter identification of a round-robin test box model using a deterministic and probabilistic methodology, J. Build. Perform. Simul. 11 (2018) 623–638. https://doi.org/10.1080/19401493.2017.1420824.
[23] T. Yang, Y. Pan, J. Mao, Y. Wang, Z. Huang, An automated optimization method for calibrating building energy simulation models with measured data: Orientation and a case study, Appl. Energy. 179 (2016) 1220–1231. https://doi.org/10.1016/j.apenergy.2016.07.084.
[24] Z. Michalewicz, D.B. Fogel, How to Solve It: Modern Heuristics, Springer Sci. Bus. Media. (2013).
[25] L. Rivalin, P. Stabat, D. Marchio, M. Caciolo, F. Hopquin, A comparison of methods for uncertainty and sensitivity analysis applied to the energy performance of new commercial buildings, Energy Build. 166 (2018) 489–504. https://doi.org/10.1016/j.enbuild.2018.02.021.
[26] S. Martínez, P. Eguía, E. Granada, A. Moazami, M. Hamdy, A performance comparison of multi-objective optimization-based approaches for calibrating white-box building energy models., Energy Build. 216 (2020). https://doi.org/10.1016/j.enbuild.2020.109942.
[27] A. Garrett, J. New, Scalable tuning of building models to hourly data, Energy. 84 (2015) 493–502. https://doi.org/10.1016/j.energy.2015.03.014.
[28] K. Sun, T. Hong, S.C. Taylor-Lange, M.A. Piette, A pattern-based automated approach to building energy model calibration, Appl. Energy. 165 (2016) 214–224. https://doi.org/10.1016/j.apenergy.2015.12.026.
[29] E. Hale, L. Lisell, D. Goldwasser, D. Macumber, J. Dean, I. Metzger, A. Parker, N. Long, B. Ball, M. Schott, Others, Cloud-Based Model Calibration Using OpenStudio, in: ESim, Ottawa, Canada, 2014.
[30] S. Sansregret, K. Lavigne, A. Daoud, L.-A. Leclaire, ExCalibBEM Tool Development to Calibrate Building Energy Models, Proc. ESim 2014 8th Conf. IBPSA-Canada, May 8-9. (2014) 7A.5.1-7A.5.11.
[31] ASHRAE, ASHRAE Guideline 14 - Measurement of Energy, Demand, and Water Savings, (2014).
[32] J. Sun, T.A. Reddy, Calibration of building energy simulation programs using the analytic optimization approach (RP-1051), HVAC R Res. 12 (2006) 177–196. https://doi.org/10.1080/10789669.2006.10391173.
[33] H. Lim, Z. (John) Zhai, Influences of energy data on Bayesian calibration of building energy model, Appl. Energy. 231 (2018) 686–698. https://doi.org/10.1016/j.apenergy.2018.09.156.
[34] I. Ballarini, S.P. Corgnati, V. Corrado, Use of reference buildings to assess the energy saving potentials of the residential building stock: The experience of TABULA project, Energy Policy. 68 (2014) 273–284. https://doi.org/10.1016/J.ENPOL.2014.01.027.
[35] T. Loga, B. Stein, N. Diefenbach, TABULA building typologies in 20 European countries—Making energy-related features of residential building stocks comparable, Energy Build. 132 (2016) 4–12. https://doi.org/10.1016/j.enbuild.2016.06.094.
[36] K. Menberg, Y. Heo, R. Choudhary, Sensitivity analysis methods for building energy models: Comparing computational costs and extractable information, Energy Build. 133 (2016) 433–445. https://doi.org/10.1016/j.enbuild.2016.10.005.
[37] E. Baldoni, S. Coderoni, E. Di Giuseppe, M. D'Orazio, R. Esposti, G. Maracchini, A Software Tool for a Stochastic Life Cycle Assessment and Costing of Buildings' Energy Efficiency Measures, Sustainability. 13 (2021) 7975. https://doi.org/10.3390/su13147975.
[38] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola, Global Sensitivity Analysis. The Primer, Wiley, John, Chichester, 2008. https://doi.org/10.1002/9780470725184.
[39] M.H. Kristensen, S. Petersen, Choosing the appropriate sensitivity analysis method for building energy model-based investigations, Energy Build. 130 (2016) 166–176. https://doi.org/10.1016/j.enbuild.2016.08.038.
[40] Z. Yang, B. Becerik-Gerber, A model calibration framework for simultaneous multi-level building energy simulation, Appl. Energy. 149 (2015) 415–431. https://doi.org/10.1016/j.apenergy.2015.03.048.
[41] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Trans. Evol. Comput. 6 (2002) 182–197. https://doi.org/10.1109/4235.996017.
[42] S. Martínez, E. Pérez, P. Eguía, A. Erkoreka, E. Granada, Model calibration and exergoeconomic optimization with NSGA-II applied to a residential cogeneration, Appl. Therm. Eng. 169 (2020) 114916. https://doi.org/10.1016/j.applthermaleng.2020.114916.
[43] I. Costa-Carrapiço, R. Raslan, J.N. González, A systematic review of genetic algorithm-based

multi-objective optimisation for building retrofitting strategies towards energy efficiency, Energy Build. 210 (2020). https://doi.org/10.1016/j.enbuild.2019.109690.
[44] T. Loga, B. Stein, N. Diefenbach, TABULA building typologies in 20 European countries—Making energy-related features of residential building stocks comparable, Energy Build. 132 (2016) 4–12. https://doi.org/10.1016/J.ENBUILD.2016.06.094.
[45] L. Carnieletto, M. Ferrando, L. Teso, K. Sun, W. Zhang, F. Causone, P. Romagnoni, A. Zarrella, T. Hong, Italian prototype building models for urban scale building performance simulation, Build. Environ. 192 (2021) 107590. https://doi.org/10.1016/j.buildenv.2021.107590.
[46] ASHRAE 55, Ashrae Handbook Fundamentals, 2017. https://doi.org/10.1039/c1cs15219j.
[47] UNI, UNI 10351:2015 - Materiali e prodotti per edilizia - Proprietà termoigrometriche - Procedura per la scelta dei valori di progetto, (2015).
[48] ISO, EN ISO 10077-1:2018. Thermal performance of windows, door and shutters - Calculation of thermal transmittance - Part 1: General, (2018).
[49] K.A.R. Ismail, J.R. Henríquez, Modeling and simulation of a simple glass window, Sol. Energy Mater. Sol. Cells. 80 (2003) 355–374. https://doi.org/10.1016/j.solmat.2003.08.010.