Building energy model calibration: a detailed case study using sub-hourly measured data

Dimitri Guyot, Florine Giraud, Florian Simon, David Corgier, Christophe Marvillet, Brice Tremeac

▶ **To cite this version:**

Dimitri Guyot, Florine Giraud, Florian Simon, David Corgier, Christophe Marvillet, et al.. Building energy model calibration: a detailed case study using sub-hourly measured data. Energy and Buildings, 2020, pp.110189. ⟨10.1016/j.enbuild.2020.110189⟩. ⟨hal-02798684⟩

HAL Id: hal-02798684

https://hal.science/hal-02798684v1

Submitted on 22 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons CC BY-NC 4.0 - Attribution - Non-commercial use - International License

Version of Record: https://www.sciencedirect.com/science/article/pii/S0378778820300505
Manuscript_9eecc09833748248517b19a4e7cc0484

Building energy model calibration: a detailed case study using sub-hourly measured data

Dimitri Guyot[1, 2], Florine Giraud[1], Florian Simon[2], David Corgier[2], Christophe Marvillet[1], Brice Tremeac[1]
Corresponding author: brice.tremeac@lecnam.net

[1]Laboratoire du froid, des systèmes énergétiques et thermiques (Lafset), CNAM, 292 rue Saint Martin – 75003 Paris, France
[2]MANASLU Ing., Savoie Technolac, BP80209 – 73374 Le Bourget du Lac

Abstract

This paper details a manual calibration of a 132 zones, highly monitored 5 434 m² office building located in France. The calibration is conducted over one year of simulation at a five minute time-step. The model is mostly fed with input files containing measured data for each simulation time-step. First simulation results show that the model cannot be considered as being calibrated considering usual guidelines based on statistical indicators calculated on monthly energy consumption. Instead of going through a classical tuning process to match measured and simulated monthly energy consumption data, the choice was made to focus on explaining those discrepancies using available building information and monitoring data. A first dynamical heating and cooling powers analysis has been made, resulting in good correlation with measured values and explaining most of the discrepancies previously noticed, followed by a dynamical temperature analysis performed at the zone scale, and completed by a global statistical analysis at the building scale for the 132 zones of the building. Results showed very good agreement between measured and simulated temperature values with 90 % of data points ranging between -1.72 °C and 2.02 °C of error. This demonstrate that a model that at first glance does not perform well considering particular statistical indicators can show an accurate dynamical behavior and a good performance when other variables and performance criteria are considered. This is followed by a discussion of the possible model improvements along with the limitations of statistical analysis compared to dynamic physical analysis for calibration performance assessment. The paper concludes with several proposals to improve the guidelines for calibration procedures involving the use of detailed monitoring data.

Keywords: Simulation, calibration, building, performance gap, monitoring, sub-hourly, EnergyPlus.

1. Introduction

1.1. BES models and characteristics

With more than 40 % of the energy use and 36 % of the CO_2 emissions [1], the European building sector is one of the main contributors to global warming. It also represents a tremendous potential for improvement through the retrofit of the existing building stock and the construction of new efficient buildings. Building energy simulation (BES) became an unmissable and powerful tool to overcome this challenge and now plays a significant role in the building industry. Many BES tools

© 2020 published by Elsevier. This manuscript is made available under the CC BY NC user license
https://creativecommons.org/licenses/by-nc/4.0/

exists, each of them having different capabilities and limits [2]. As law-driven models (also known as physical models or white-box models), BES models rely on the thermodynamics rules that govern a system to predict its behavior given its properties and external conditions. Most of them are based on the nodal approach [3], which consists in considering each zone of the building as an homogeneous volume characterized by discrete values of state variables such as temperature, pressure, or humidity. Each zone represents one node of the mesh, in which the thermal transfer equations are solved over each time-step. This method can be considered as a one-dimensional approach. Those models require large computation time and a thorough knowledge of the building to be modeled. They are able to model a system given a set of previously unobserved initial conditions. However, detailed data may not be available during the design stage of a building, which forces the user to make different assumptions. This could result in an inaccurate simulation leading to a well-known problem called the "performance gap" (i.e. the difference between the simulated and measured values). Several studies pointed out significant discrepancies between the BES program outputs and the actual measured value [4,5], discrediting those tools and their users [6] and questioning their usefulness in the building sector.

1.2. The performance gap

The performance gap results from the existence of a variety of uncertainties in the process of building modeling. Their study is the subject of specific researches, and two main fields can be distinguished, uncertainty and sensitivity analysis. While uncertainty analysis aims at calculating the likely variation in the outputs due to the variations in the inputs, the sensitivity analysis, as stated by Saltelli *et al* [7]*,* aims at studying “how the uncertainty in the output of a model can be apportioned to different sources of uncertainty in the model input”. In other words, the aim of sensitivity analysis is to identify the input parameters to which the output is sensitive. As a result, the model output may be sensitive to an input but, if this parameter is well known, it will not be a critical parameter in an uncertainty analysis [8]. A complete inventory of sensitivity analysis methods applied to the study of BES models is made in [9] and [10], while a review of uncertainty analysis results and techniques for building energy assessment can be found in [11].

According to Imam *et al.*[6], the performance gap find its origin in the three phases of a building's life: design, construction and operation. Many studies already identified uncertainty sources occurring at different stages [12–14], classified as follow:

- ***Specification uncertainties:*** Incomplete, improper or inaccurate parameters specification such as building geometry or HVAC (Heating, Ventilation, and Air Conditioning) systems properties. This can be due to a user's lack of experience or negligence, or to an incomplete knowledge of inputs parameters.

- ***Modelling uncertainties:*** Simplifications and assumptions made in the models used by the BES program to represent the physical phenomenon (nodal model, space and time discretization, etc.). This does not directly rely on the user, but he can choose to reduce this uncertainty by for example reducing the time-step or choosing the most appropriate model to represent a given phenomenon.

Discrepancies can also occur during the construction process, identified in [6] as the "construction gap". Even if a very particular attention is given during this phase to produce a high quality building,

it will not be as perfect as it can be in a BES program, which will lead in unavoidable differences. Finally, during the operational stage occurs the last source of discrepancies with the predictions, identified as the scenario uncertainties source.

- ***Scenario uncertainties:*** Errors resulting from the assumptions made on external parameters such as weather data or occupants' behavior. The stochastic nature of those parameters inevitably lead to a form of uncertainty. However, it can be drastically reduced during the design phase, yet it entirely relies on the user's knowledge. Even if representative weather files are now easily accessible, poor occupants' scenario assumptions (occupancy, internal loads, etc.) can still be made and will inevitably impact the simulation performance.

Many researchers have worked to identify the most influential input parameters along with their associated uncertainty to quantify their effect on the model outputs. As an example Menberg *et al.* [15] compared several methods for sensitivity analysis of BES models. They studied the influence of eleven parameters that are typically selected for sensitivity analysis, and identified set point temperature, infiltration, building thermal properties, and ventilation related parameters as influential for building energy models. They stated that these findings are in good agreement with the conclusions of previous studies. However, they did not consider weather data in their analysis, even though this was identified as one of the sources of uncertainty in BES models [11].

Specification uncertainties have been widely studied in the literature. Based on the work of Clarke *et al.* [16], Macdonald [8] established detailed uncertainty data, especially for conductivity, density, specific heat capacity, or emissivity and absorptivity of construction materials. Other key parameters of building envelope such as infiltration rate [17] or thermal bridges [18] have also been studied. As previously mentioned, specification uncertainty can also be related to HVAC systems properties. Wang *et al.*[19] studied the influence of different building operation strategies related to HVAC systems on the energy consumption of an office building. They showed that the combined uncertainties of those parameters could result in an uncertainty on the annual energy consumption ranging from -15.8 % to 70.3 %.

Although less frequently studied, it is possible to find studies dealing with modeling uncertainties. As an example, the effect of simulation time-step has been studied in [20]. The author simulated the same model at various time-steps (1, 5, 10, 15, 20, 30, and 60 min) and found differences up to 3.6 % in daily cooling load when simulating at a 60 min time-step, compared to 1 min time-step simulation. When considering daily peak load (which is an often used sizing factor for building design), discrepancies goes up to 75 %. Dos Santos and Mendes [21] found other results with zone temperature differences up to 4 °C and 10 % variations in humidity ratio between 1 s time-step and 3600 s time-step simulations. They conclude that a 60 min time-step can lead to strong discrepancies in indoor air temperature and consequently in building load calculation.

Scenario uncertainty, such as weather data, have also been studied in the literature. Thus, Wang *et al.*[19] investigated the influence of weather data among other parameters on the energy use in an office building. They considered four different cities for which they gathered 10 to 15 years of weather data. They found that the energy could vary from - 4.0 % to 6.1 % depending on the city and year compared to the typical meteorological year that is usually considered.

Together, these studies provide insight into the various sources of uncertainty and their effects on the model outputs. However, these results are highly dependent on the building under

consideration. Therefore, they remain trends that can be expected, but are by no means directly transferable to any given building.

The uncertainty on the metered data used for the model performance assessment due to the sensors' accuracy can also explain a part of the performance gap. Although this may be negligible regarding the discrepancies usually observed, it should be considered or at least mentioned.

The complexity of BES programs and the complete freedom the user benefit in their utilization make the modeler the main reason of a potential "performance gap". This statement can be illustrated in the work of Imam *et al.*[6] where the authors asked 108 modellers to comment the importance of 21 common modelling input variables such as U-values, ventilation rate or glazing ratio. They found very heterogeneous results with no correlation between modellers' answers since many of them identified parameters to be important when others thought they were irrelevant. In addition, the authors found little correlation between the variables which were thought to be important by the modellers and which proved to be objectively important. Those results corroborate Guyon findings [22] where 12 users where asked to model the same residential house using the same software. The author found energy consumption predictions ranging from -41 % to +39 % around the average value. As mentioned by Imam *et al.*[6], modellers rarely compare energy consumption results with measured values of the building during its operational stage. Consequently, there is no feedback about good or wrong practices, causing the knowledge to drift over time. To overcome this issue, calibration process should be conducted regularly by modellers. However, this implies to have highly monitored buildings available and to give a real value to a calibrated model since the process is very time consuming.

1.3. Calibration methods and model performance assessment

The calibration process aims at closing, or at least reducing the performance gap. It consists in tuning the different unknown input parameters within a defined range in order to match the simulated and measured values. Reddy [12] proposed four classes to classify the different calibration approaches:

- Manual, iterative and pragmatic intervention
- A suite of informative graphical comparative displays
- Special tests and analytical procedures
- Analytical/Mathematical methods of calibration

The details of each category as well as examples of applications can be found in his review. Those different approaches for BES models calibration have been broadly classified in two main categories [13]:

- ***Manual:*** This gathers every method that does not rely on automated procedures, these are the most commonly used. They are conducted by the user, in an iterative and "trial and error" manner, and entirely rely on the user's experience, knowledge and judgment. Those methods can be considered as subjective. The majority of these methods involve the use of graphical representations [23] and comparative displays including time-series plots, 3D comparative plots [24,25], signature analysis methods [26], or statistical displays [24].

- ***Automated:*** This include all automated approaches based on mathematical or statistical tools that cannot be considered as user driven. Most of these techniques require to identify a set of

influential parameters to be calibrated as well as their maximum range of variation using sensitivity analysis. The objective is then to find, among the multitude of possible configurations, the set of input parameters that minimizes a cost function, usually representing the error between measurement and simulation [27,28]. Another approach regularly used is the Bayesian calibration, which has the advantage of naturally taking into account the uncertainties in the calibration process [29].

Detailed explanations as well as examples of the application of each of the above methods can be found in [13] and [30]. Regardless of the calibration technique used, the process remains time-consuming and costly. Indeed, manual techniques require the time of an experienced human to analyze the results of the multiple simulations that will be conducted. Similarly, automatic methods also require the time of a human expert to be developed, as well as the computing power and time needed to perform a large number of simulations. According to Fabrizio *et al.*[30], there are five different calibration complexity levels, depending on the data available for comparison, from simple utility bills comparison to long-term monitoring:

- **Level 1:** The first level of calibration is based on energy bills only. This type of calibration will therefore only be based on monthly or annual energy consumption data. As the dynamic aspect is not taken into account, the calibration can be qualified as static.

- **Level 2:** The data available for a level 1 calibration may be supplemented by visual inspections on site to verify some of the design data (equipment installed, internal geometry of the building, zones occupancy, etc.). The building manager may also be questioned. Calibration will therefore always be based on monthly energy consumption data, and the dynamic aspect will not be considered.

- **Level 3:** In addition to the information available for the previous levels, information can be obtained from detailed on-site audits through the use of measuring instruments (oscilloscope, luxmeter, balometer, temperature probe, etc.). These measurements may be carried out at different times of the day but remain punctual, as no instrument is left on site. The dynamic aspect may be considered, essentially for systems, over short periods of time only, around the hour.

- **Level 4:** At this level, measuring instruments are installed on site to measure different variables (temperatures, thermal or electrical energy, humidity, etc.) over a relatively short period of time (typically several weeks to a few months), at short time intervals (from minutes to hours). Here, calibration can be done on other variables than the simple monthly energy consumption, such as zone temperatures or the thermal or electrical powers. Because the measurement period has become sufficiently long, the building dynamics can be considered over specific periods of time only.

- **Level 5:** This last level of calibration involves an almost exhaustive instrumentation, which is permanently installed on the building. The instrumentation was therefore designed beforehand, during the design phase of the building. There are no limits to the duration of the monitoring, meaning that several years of data are usually available. The recording interval is small enough to capture the dynamics of rapid transient phenomena (of the magnitude of a minute). The data is usually accessible remotely and the building is controlled through a BEMS (Building Energy Management System) which also allows its monitoring along with the

modification of its operation in real time. Calibration can be carried out in a totally dynamic way, in order to reproduce the dynamic behavior of the building over the entire period under consideration by working on variables such as thermal and electrical power or the temperature and humidity of the zones.

Evaluating a model's accuracy asks the question of the indicators that have to be used. For the sake of uniformity and standardization, statistical indices summarized in Table 1 have been selected by three international organizations [31–33] to assess whether a model can be considered as being calibrated or not [13,30]. The maximum value of each criterion varies depending on whether the model is calibrated on an hourly or monthly basis. Note that the hourly NMBE criterion of the IPMVP guideline is more restrictive than the monthly criterion, which goes against the trend observed for the other two guidelines. According to every guideline, the performance calculation is mostly made on energy consumption data by comparing the simulated and the measured values.

Organization/Guideline	Monthly calibration		Hourly calibration	
	CvRMSE [%]	NMBE [%]	CvRMSE [%]	NMBE [%]
IPMVP [31]	-	20	20	5
ASHRAE - Guideline 14 [32]	15	5	30	10
FEMP - M&V Guidelines [33]	15	5	30	10

Table 1: Validation criteria for BES models

A confusion between the Mean Bias Error (MBE) and the Normalized Mean Bias Error (NMBE) has been identified in the guidelines. The origins of this error, the potential consequences on the results, and the articles concerned are reported in [34]. According to the authors, the most common mistake is the use of the abbreviation “MBE” when referring to the “NMBE”.

- **NMBE:** It corresponds to the normalization of the mean bias error (MBE) using the mean of the measured values to make the results comparable regardless of the unit used. It indicates the overall bias of the model. However, it suffers from cancellation effect, when negative and positive error values offset each other. For this reason, NMBE cannot be considered alone, another indicator has to be introduced.

$$NMBE\ [\%] = \frac{1}{\overline{m}} \frac{\sum_{i=1}^{n}(M_i - S_i)}{n-p} \times 100 = \frac{n}{n-p} \frac{\sum_{i=1}^{n}(M_i - S_i)}{\sum_{i=1}^{n} M_i} \times 100 \quad (1)$$

Where:

M is the hourly or monthly measured value of energy consumption
S is the hourly or monthly simulated value of energy consumption
n is the number of data points ($n_{monthly} = 12$; $n_{hourly} = 8760$)
p is the number of adjustable model parameters

The number of adjustable model parameters p is suggested to be set to zero for calibration procedures [34], simplifying NMBE formula to:

$$NMBE\ [\%] = \frac{\sum_{i=1}^{n}(M_i - S_i)}{\sum_{i=1}^{n} M_i} \times 100 \quad (2)$$

- **CvRMSE (Coefficient of variation of the Root Mean Squared Error):** This corresponds to the RMSE (Root Mean Squared Error) divided by the averaged value of the measured data. Because of squaring, it tends to maximize the impact of important errors in comparison to

smaller errors. It is always positive and allows to determine how well a model fits the reality. The lower the CvRMSE, the more accurate the model.

$$CvRMSE\ [\%] = \frac{1}{\overline{m}}\sqrt{\frac{\sum_{i=1}^{n}(M_i - S_i)^2}{n}} \times 100 = \frac{\sqrt{\frac{\sum_{i=1}^{n}(M_i - S_i)^2}{n}}}{\frac{\sum_{i=1}^{n} M_i}{n}} \times 100 \quad (3)$$

In spite of the equifinality of BES models (i.e. multiple model configurations may produce the same results [35]), calibration performance assessment is rarely made on other variables like zones temperatures or humidity. This means for example that a model can be considered as being calibrated with no regards to the temperature evolution in the different zones, even if BES is also meant to evaluate the building's thermal comfort. Moreover, to the authors' knowledge, there is no guideline regarding calibrations conducted on a sub-hourly time-step.

2. Building description

The case study is the Hikari project located in Lyon, France (45°44'31.1"N 4°49'09.5"E, altitude 240 m), represented in Figure 1. It has been designed to be the first net zero energy city block in Europe and is operating since July 2015. The project of 12 310 m² is composed of three distinct mixed-use buildings (from left to right in Figure 1: Nishi, Minami and Higashi) built on a common basement gathering a parking lot and the mutual technical facilities providing the energy needs. Hikari is fully monitored with more than 10 000 measurement points and controlled through the BEMS (Building Energy Management System).

Figure 1: South facade of the Hikari project

The ground floors of each of the three buildings are occupied by different shops. Higashi is composed of seven floors separated in 14 offices areas (two on each floor) subdivided in offices according to the user's needs. Minami building is separated in 32 flats (four on each floor). The first five floors of Nishi are divided in 10 offices areas (two on each floor) that are also subdivided in offices. The two Nishi's upper floors are composed of four duplex apartments. The different uses and floor areas are summarized in Table 2. The offices zones are occupied during daytime, Monday to Friday, from 8 am to 6 pm, approximately.

Areas (m²)	Offices	Housing	Shops	TOTAL
Higashi	5 434	-	567	**6 001**
Minami	-	2 959	289	**3 248**
Nishi	2 338	570	153	**3 061**
TOTAL	**7 772**	**3 529**	**1 009**	**12 310**

Table 2: Floor areas summary

2.1. Envelope characteristics

Since the calibration work focus on the office zones of Higashi building, the characteristics of the other buildings will not be mentioned in this part. A special attention has been given on the envelope quality in order to meet the net zero energy target. The main envelope characteristics are gathered in Table 3.

Parameter	Higashi
Wall's U value [W/(m².K)]	0.16
Ground floor U value [W/(m².K)]	0.18
Roof U value [W/(m².K)]	0.15
Windows U value [W/(m².K)]	1.4
Solar heat gain coefficient	0.58
Visible light transmission	0.78

Table 3: Envelope main characteristics

Several measures have been taken, including exterior wall insulation, double glazing low emissivity, and thermal bridges treatment. The airtightness has been meticulously treated with an objective of air leakage rate under 50 Pa of pressure difference between the interior and the exterior of the building (n_{50} criterion) below 0.6 vol/h, which correspond to the Passivhaus standard. A blower door test has been performed on the whole office volume corresponding to the seven floors of Higashi. The airtightness has been measured at 0.595 vol/h under 50 Pa.

2.2. Energy production and distribution systems

The three buildings pool their energy through different and innovative energy production systems represented in Figure 2. The heating needs are provided by a rapeseed oil cogeneration unit (98 kW of thermal power) and a gas boiler (300 kW) coupled to 7 hot water storage tanks of 4 m^3 each. This corresponds to an available energy of 1.3 MWh when considering a use involving a temperature difference of 40 °C between the inlet and outlet of the tanks. The cooling needs are provided by an absorption chiller of 46 kW and a magnetic levitation chiller of 300 kW coupled to a chilled water storage tank using phase change materials (50 m^3, 0.61 MWh, melting occurs à 10 °C). The electricity produced by the cogeneration unit (75 kW of electric power) and photovoltaic panels (188 kWp) can be consumed on site, injected on the grid, or stored in a battery pack (100 kWh).

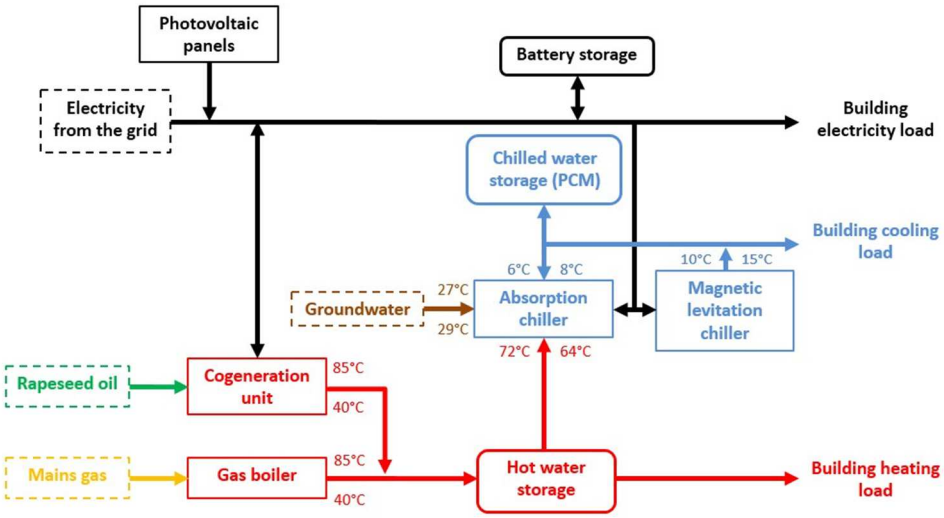

Figure 2: Energy production and distribution systems with their operating temperature range

2.3. Zones equipment and control strategies

Each office is independently controlled in temperature, ventilation, lighting level and is equipped with motorized external venetian blinds. The offices areas are heated and cooled by radiant ceilings and the ventilation is of single-flow type. The radiant ceilings' water supply temperature varies depending on the measured outdoor temperature to benefit from increased heating power when energy needs are high. Radiant ceilings' main characteristics based on measurements made by the manufacturer are presented in Table 4.

	Heating mode	Cooling mode
Inlet water temperature [°C]	45	20
Outlet water temperature [°C]	40	23
Flow [l/h]	128	76
Power density [W/m²]	215	77
Surface [m²]	1.28	
Heating capacity density [J/K.m²]	5 307	

Table 4: Radiant ceilings characteristics, power density is given in heating and cooling mode for ambient temperatures values of 20 °C and 27 °C, respectively. Heating capacity density does not take into account the mass of water contained in the panel.

Occupants can choose the temperature set point in occupied mode within the range from – 2 °C to + 2 °C around a fixed temperature set point. If nobody is detected in the zone during the predefined occupied hours, the temperature set-point switches to standby mode. It corresponds to a slightly lower (or higher in cooling mode) value allowing to keep the zone at an acceptable temperature so the occupied mode set point can be reached faster if someone enter the zone again. If nobody is detected outside the predefined occupied hours, the temperature set point switches to unoccupied mode, corresponding to a lower (or higher in cooling mode) value. The radiant ceiling's 2-way valve are controlled to reach the office's temperature set point. Each window is equipped with opening sensor so the heating or cooling systems can be turned off in case of a window opening. The on/off ventilation control is based on presence detection. If someone is detected, the damper opens to provide the nominal air flow in the office. The air is pre-heated or pre-cooled depending on the measured outdoor temperature by AHU (Air Handling Units) for each office area. This allows for an increased heating power when energy needs are high. The lighting level is controlled according to the presence detection and the measured lighting level. If someone is detected and if the measured

lighting level is below the set point, the light is activated. Each blind is independently controlled in angle and opening according to the horizontal direct solar radiation measured on the building's roof and the information about the shadow presence on the window (numerically calculated depending on the window position and orientation on the building, and on the solar height and angle along the year). If there is no shadow on the window and if the measured direct solar radiation exceeds 150 W/m², the blind is closed. The occupants can choose to manually control the blinds, but they are automatically switch back to automatic control twice a day.

3. Model and calibration procedure description

3.1. Calibration methodology

Due to the large amount of monitoring data available and the unrestricted access to the design data (plans, diagrams, data sheets, calculation notes, etc.), we chose to conduct a manual calibration [13] of level 5 [30]. The calibration process was conducted in three main steps:

- **Update of the geometry of the building:** the model created during the design phase has been revised by drawing more precisely some facade elements that could results in solar masks. Other close masks such as non-existing buildings in the design phase have been added. The dimensions of some of the windows have also been modified. Finally, the partitioning of each level, which was not known at the design stage, has been entirely designed in accordance with the existing layout chosen by the occupant.

- **Model parameter setting:** the conformity of the model parameters with the available data (plans, calculation notes, visits and audits, etc.) has been checked, and then modified in case of non-conformity. In the case of parameters that remained unknown, they are simply be set to their most likely value.

- **Input data creation:** the monitoring data were extracted from the BEMS, analyzed (checking of their consistency, of the units, etc.), then formatted (removal of abnormal values, resampling, etc.) to be given as input to the model. This step accounts for most of the time needed to calibrate the model due to the large amount of data to be processed. Furthermore, generic processing is sometimes not possible, resulting in a particularly time-consuming specific processing. The "annual schedules" generated in this way are in the form of .csv input files containing for each variable a value for each simulation time-step (8 760 hours at a five minutes time-step, i.e. 105 120 lines).

A conservative approach was chosen, allowing us to maintain a full understanding and control over the calibration process. No parameters were tuned, the approach simply consists in feeding the model with as much known data as possible in order to constrain it to the maximum to limit the unknown domain. By this approach, unknown parameters are not blindly tuned to match measured data. Since a given variable can be involved in several physical phenomena, it will not disrupt other phenomena in a wrong manner. Our approach is similar approach to the one described in [36] and applied in [37], and can be considered as the first part of a classic manual calibration process, where all the known variables are set before starting the tuning process. Each decision taken along the process is based on a factual observation of the results. This “evidence-based methodology” [36],

implies a complete understanding of the physical phenomena involved and a thorough knowledge of the building. Decision-making along the process was based on the use of statistical analysis tools such as the calculation of NMBE and CvRMSE indicators on monthly energy consumption, or the calculation of error on zone temperatures. Statistical visualization tools such as probability density plots, scatter plots, or box plots were also used. These statistical analyses were systematically completed by the analysis of the associated physical parameters, especially through the use of time-series plots. As recommended in [36], a version control file was created to store each revision of the model along with the corresponding modifications made and the obtained results, allowing to keep a complete history of the calibration process.

Due to several measurement errors and dysfunctions, data were not available to calibrate a few zones in the buildings. In these cases, the configuration used in design phase is not modified. Since the measurement uncertainty is not known for all the variables used for the calibration, it has not been considered in the following analysis.

3.2. Model input parameters

The model input parameters can be directly obtained from monitoring or from the information available on the building. Some may also result from a calculation mixing measured variables, design data, and assumptions. A summary of the calibrated parameters can be found in Table 5 where simulation inputs are written in bold. An indication of the confidence given to each parameter is given in the last column. Monitoring data along with data from on-site audits are considered to be particularly reliable. In addition, a commissioning process has been carried out, allowing greater confidence to be placed in the design data.

	Parameter	Origin	Confidence
Weather	**Temperature**	**Measured via monitoring**	**High**
	Humidity		
	Wind speed		
	Wind direction		
	Direct solar radiation		
	Diffuse solar radiation		
Heating and cooling	**Heating temperature setpoint**	**Measured via monitoring**	**High**
	Nominal heating power	**Design data**	
	Cooling temperature setpoint	**Measured via monitoring**	
	Nominal cooling power	**Design data**	
Mechanical ventilation	Damper position	Measured via monitoring	High
	Nominal air flow	Design data	High
	Actual air flow in each zone	**Calculated**	**High**
Occupancy	Nominal number of people	Design data	Medium
	Occupancy density schedule	Hypothesis	Medium
	Actual occupancy in each zone	**Calculated**	**Low**
Blinds	Direct solar radiation	Measured via monitoring	High
	Shadow presence	Design data	High
	Manual takeovers	Unknown	-
	Blind position for each window	**Calculated**	**Medium**
Natural ventilation	Window opening status	Measured via monitoring	High
	Number of windows in each zone	Design data	High
	Nominal air flow	Hypothesis	Medium

	Actual air flow in each zone	**Calculated**	**Medium**
Electrical internal loads	Lighting power	Measured via monitoring	High
	Plugs power	Measured via monitoring	High
	Miscellaneous power	Measured via monitoring	High
	Floor area	Design data	High
	Electrical internal loads in each zone	**Calculated**	**Medium**
Infiltration	Building infiltration air flow	Measured via on-site audit	High
	Building leakage distribution	Observed via on-site audit	Medium
	Walls area	Design data	High
	Leakage distribution in each zones	**Calculated**	**Medium**

Table 5: Model input parameters

The parameters presented in this table are among the most influential and are those requiring some clarification regarding on how they were obtained. Therefore, the list in this table is not exhaustive, as a multitude of other input parameters are also known, such as the geometry of the building, the composition of the walls, or their thermo-physical properties. Details regarding the generation of each input parameter mentioned in Table 5 are given below:

- **Weather:** Thanks to the weather station installed on the roof, a weather file is created containing the on-site measured values of temperature, humidity, wind speed and direction, direct solar radiation rate and diffuse solar radiation rate at a five minutes time-step over the whole year.

- **Heating and cooling:** The measured heating and cooling set points of each zone at a five minutes time-step are gathered in an input file fed in the model. The radiant ceiling sizing heating and cooling power of each zone are replaced by the real installed powers.

- **Mechanical ventilation:** The ventilation sizing air flow of each zone is replaced by the real nominal air flow and the measured ventilation damper position of each zone at a five minutes time-step are gathered in an input file fed in the model. The supply air temperature set point (depending on the outside temperature) measured at a five minutes time-step is also fed in the model.

- **Occupancy:** The nominal number of people of each zone is fed in the model, based on the information given by the building tenant. Since the occupancy sensor (image-based sensor, moving object detection by image processing) only tells if a zone is occupied or unoccupied, it is not possible to know exactly the number of people in a zone at each time-step. Consequently, the nominal number of people is simply multiplied with a normalized occupancy density weekly schedule, visible in Figure 3, determined according to the measured arrival and departure mean times of the occupants. This seems to be a reasonable approximation given the open-space organization of the offices and the uniformity of use inducing common work schedules.

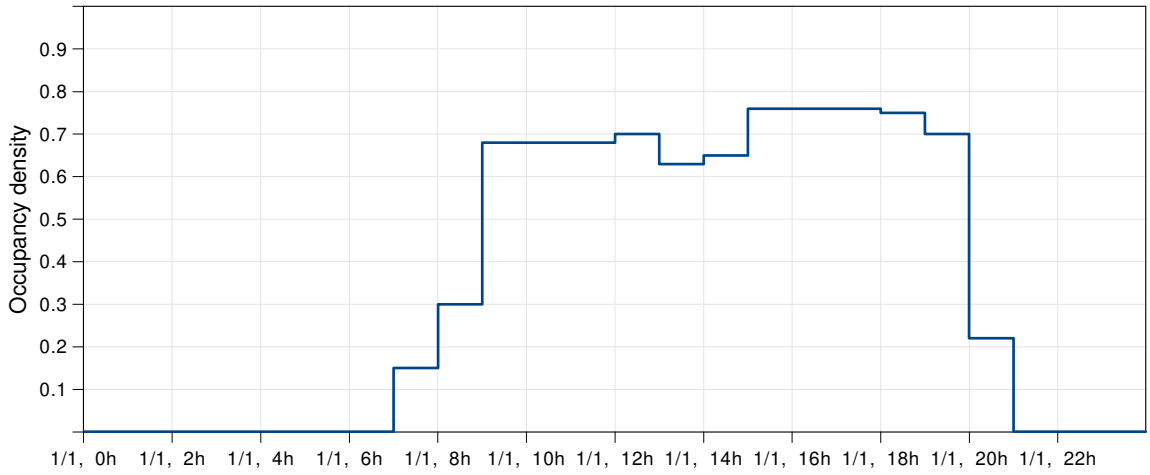

Figure 3: Normalized occupancy density weekly schedule

- **Blinds:** The shadow presence on the window is calculated at every time-step by the model during the simulation. This information is crossed with the measured value of direct solar radiation rate at a five minutes time-step to produce a blind opening schedule for each window at the simulation time-step for the whole year. Even if the blind opening is measured, creating a specific schedule for each window would have been time-consuming. The adopted approach in the simulation is very close from the real functioning, the only error is coming from the periods of breakdowns or when the blinds have been manually controlled.

- **Natural ventilation:** The window opening status of each window is known from the monitoring. Based on the measured number of open windows in the zone and a nominal air flow rate of 0.5 volume per hour per opening, the natural ventilation air flow is fed in the simulation at a five minutes time-step.

- **Electrical internal loads:** The lighting, plugs, and miscellaneous power (gathering the electrical energy consumption of all the regulation and control elements such as radiant ceilings 2-ways valves, ventilation dampers, light sensors, occupancy sensors, light dimmers, blinds motors) are not known for each zone since it's only measured for each of the 14 office areas. Thus, the power density flux is calculated for each office area based on the power measurement and the electrical power is then assigned to each zone on a floor-area-weighted basis at a five minutes time-step.

- **Infiltration:** A blower door test has been performed on the whole heated volume, that is, the seven floors of the Higashi building. This allows to calculate the total infiltration air flow rate of the two heated volume at each time-step of the simulation depending on the measured wind speed and indoor/outdoor temperature difference. According to where the infiltrations are observed in the zones (floor, ceiling, walls, windows), the obtained total infiltration air flow rate is distributed in each zone according to the area of each walls' type.

3.3. Model overview

It has been chosen to focus on the calibration of the energy needs only, to have a better control of the simulation gaps and to exclude the error induced by the energy production and distribution systems modeling. Consequently, the simulation account for energy needs only, energy production and distribution systems have been specifically modeled using Matlab/Simulink and have undergone

a specific calibration process, which is beyond the scope of this paper. The three buildings have been designed with Design Builder 4.6 and simulated using EnergyPlus 8.3.2. An overview of the model is shown in Figure 4. The simulation is conducted over one year at a five minutes time-step, from the 3rd of July 2017 to the 30th of June 2018. The CTF (Conduction Transfer Function) algorithm is used and the TARP and DOE-2 algorithms are used for interior and exterior convection respectively. The temperature is controlled on the operative temperature, defined according to equation (4).

$$T_{op} = \frac{h_{rad}T_{rad} + h_{conv}T_{air}}{h_{rad} + h_{conv}} \tag{4}$$

With T_{rad} the mean radiant temperature defined according to equation (5), T_{air} the air temperature, h_{rad} the radiation heat transfer coefficient, and h_{conv} the convective heat transfer coefficient.

$$T_{rad} = \sqrt[4]{\sum_j (F_{jh}T_j{}^4)} \tag{5}$$

With F_{jh} the configuration factor from a surface j towards human body and T_j the temperature of surface j. If we define the radiant fraction γ as:

$$\gamma = \frac{h_{rad}}{h_{rad} + h_{conv}} \tag{6}$$

The operative temperature T_{op} can be written as:

$$T_{op} = \gamma T_{rad} + (1 - \gamma)T_{air} \tag{7}$$

In the simulation, the radiant fraction used to calculate the operative temperature has been set to 0.2. This assumption has been made to take into consideration the fact that control thermostats do not measure a perfect air temperature since the radiant temperature due to their wall-mounted position influences them.

The power to be injected in the zone to reach the set point is calculated by EnergyPlus, and is limited by the previously set installed power. Part of this power is emitted into the zone by convection, the other part by radiation. The convective and radiant fraction are considered to be constant over the operating range of the emitters and were determined on the basis of measurements made on a test bench by the manufacturer of the radiant ceilings. Thus, the power emitted by a radiant panel is considered to be 40 % convective and 60 % radiant under standard conditions of use.

To keep a full understanding of the model and in order not to introduce additional errors, mechanical ventilation, natural ventilation and infiltration have been modeled separately. Indeed, even if the *AirflowNetwork Model* of EnergyPlus allows for a coupled consideration of these phenomena, it also requires to know detailed information such as the opening states of internal openings or pressure losses and air flow rates in the ventilation pipes. In our case, such requirements would have induced strong hypothesis, resulting in important, uncontrolled error.

Thermal bridges have been identified and calculated during the design phase using dedicated tools (COMSOL Multiphysics, Mold Simulator). They have been considered in the simulation using the following method:

For different thermal bridges identified on a wall and characterized by calculation, it is possible to calculate the total losses generated by these thermal bridges, and thus the total heat transfer coefficient H_{tot} through the wall according to equation (8):

$$H_{tot}[W/K] = U_{wall}A_{wall} + \sum_{k=1}^{m} \Psi_k l_k + \sum_{i=1}^{n} \chi_l \quad (8)$$

With:

U_{wall} the wall heat transfer coefficient $[W/(m^2.K)]$

A_{wall} the surface of the wall $[m^2]$

Ψ the linear thermal bridge heat transfer coefficient $[W/(m.K)]$

χ the punctual thermal bridge heat transfer coefficient $[W/K]$

Then, the thermal conductivity of the insulation material is increased to obtain an adjusted wall heat transfer coefficient $U_{wall\ adjusted}$ allowing to take the thermal bridges impact into consideration:

$$H_{tot}[W/K] = U_{wall\ adjsuted}A_{wall} \quad (9)$$

Figure 5 shows a floor view of the second floor of Higashi, only the office zones (in yellow) are calibrated. The core zones (stairs, corridors, toilets, meeting rooms, technical premises) are simulated but are not considered in the calibration process. Even if the present paper focuses on Higashi's office zones, the other buildings have also been simulated for potential further studies.

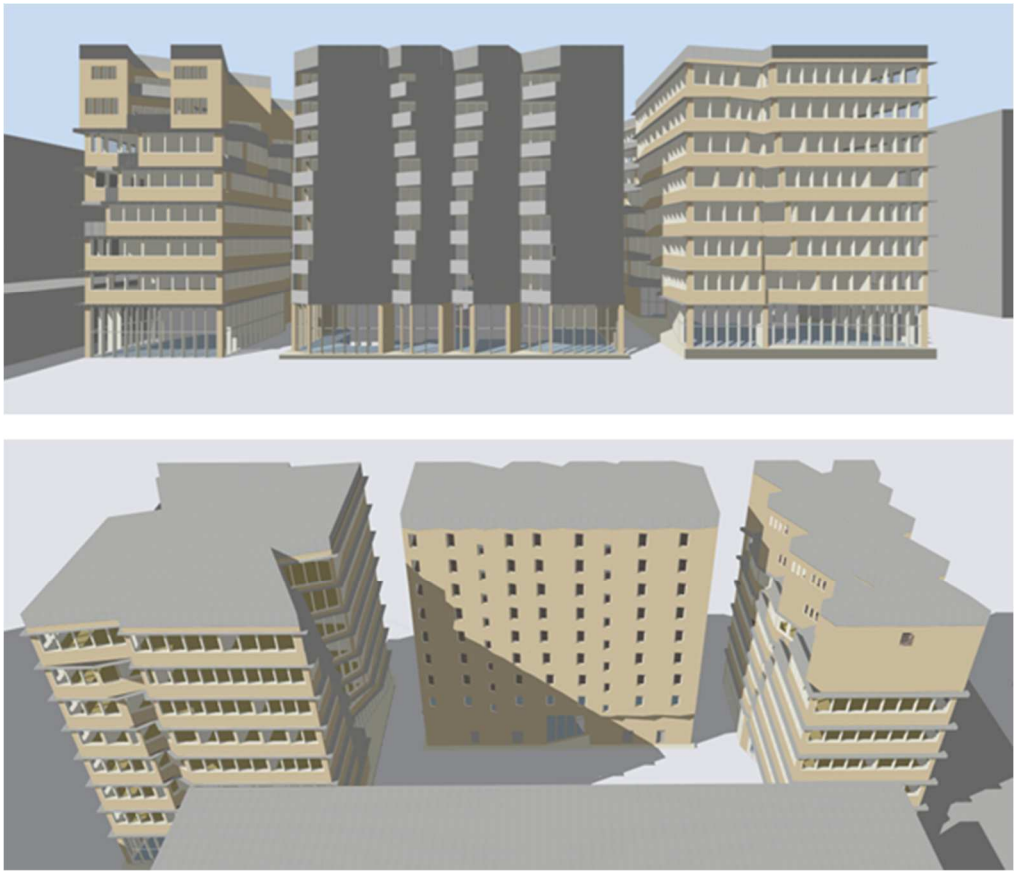

Figure 4: South facade (top) and north facade (bottom) – DesignBuilder model

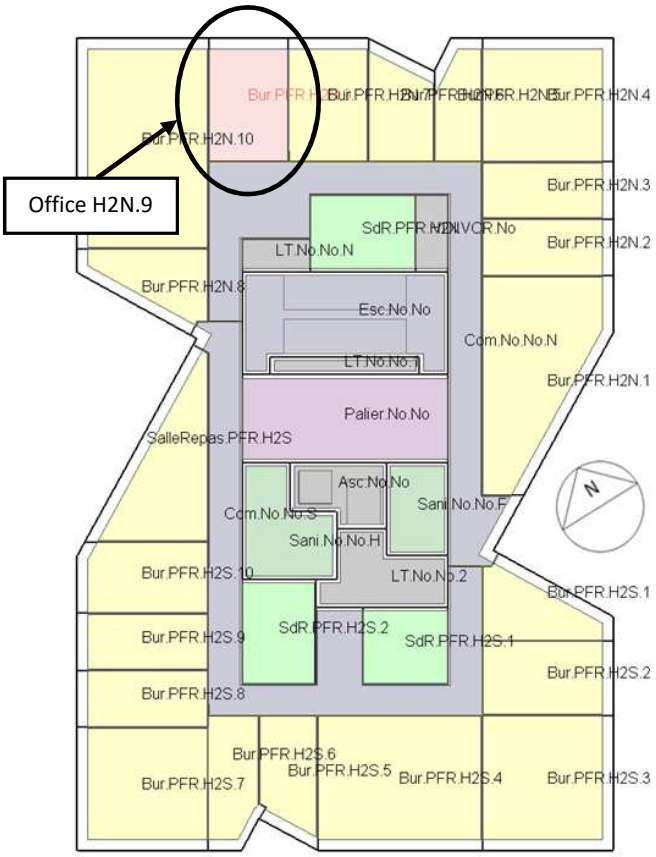

Figure 5: Second floor – DesignBuilder model

4. Results analysis

Results will first be shown at the building scale considering annually and monthly energy needs for both heating and cooling. Based on those results, a first calibration performance assessment is made on statistical indicators recommended in the guidelines previously mentioned. Then, a second performance evaluation is done through a dynamic physical analysis of the heating and cooling powers. In a third part, the model performance is also assessed through a statistical and a dynamical analysis of the zone's temperatures. Due to several measurement issues, temperature data are sometimes unavailable for comparison with simulated data, those periods have been identified and quantified. In addition, the temperature probes used for temperature measurements typically have an uncertainty of 0.1 °C.

4.1. Measured and simulated energy needs analysis

4.1.1. Annual and monthly results

Performance indicators have been calculated on monthly energy needs, corresponding to the energy consumed by the AHU and the radiant ceilings, in both heating and cooling mode according to the guidelines previously mentioned [31–33] and are compared to the recommended values in Figure 6.

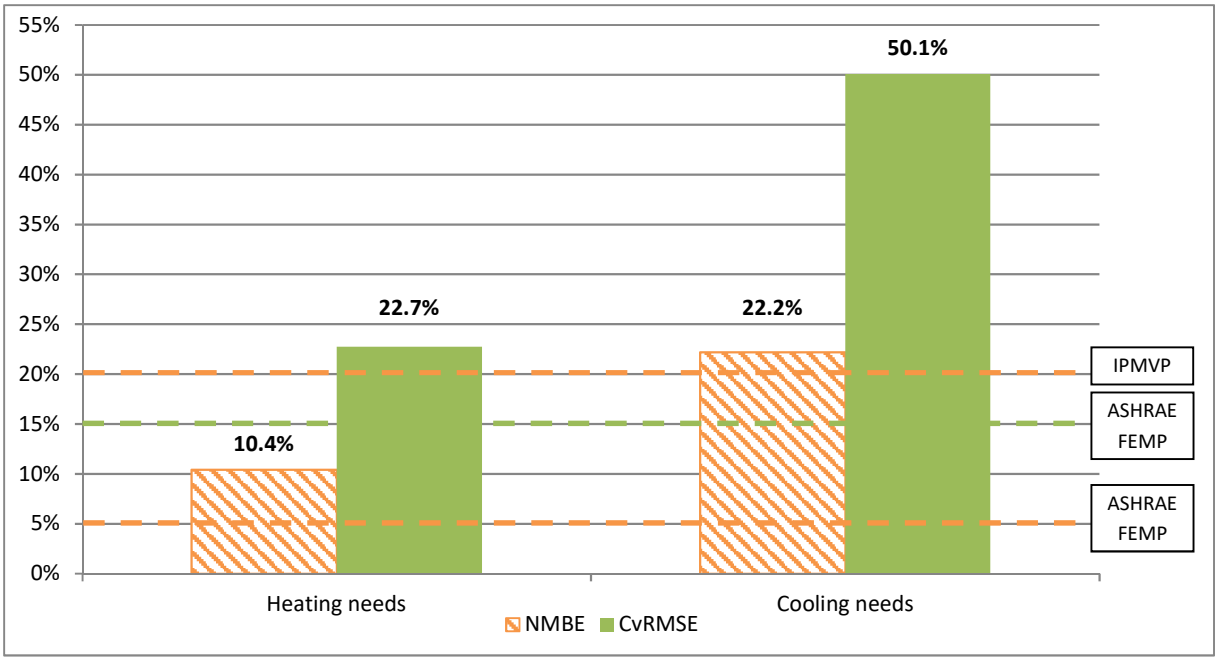

Figure 6: Calibration performance indicators in heating and cooling mode and guidelines thresholds for a monthly calibration

Regarding the IPMVP guidelines (i.e. NMBE error values lower than 20 %), Higashi's model can be considered as being calibrated for heating needs only. ASHRAE's and FEMP's calibration indicators are not respected, whether it is for heating or cooling needs. To understand these results, a comparison between measured and simulated energy need could be made. The first results show a tendency to underestimate energy needs, in both heating and cooling mode, as shown in Figure 7. The same trend can also be observed on a monthly basis in Figure 8, excepted for mid-season months where the energy demand is low.

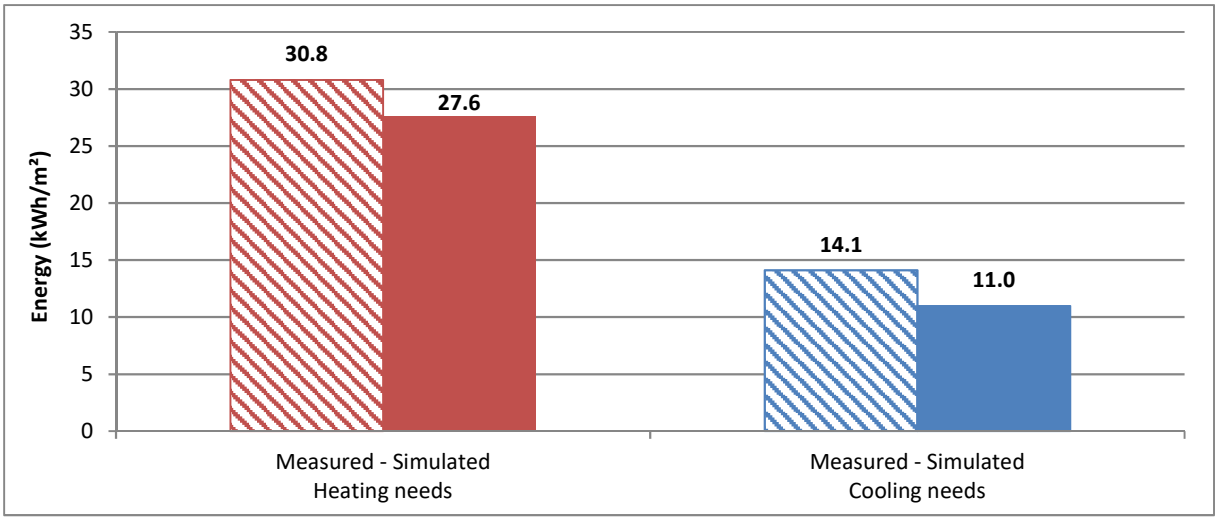

Figure 7: Annual measured and simulated heating and cooling needs

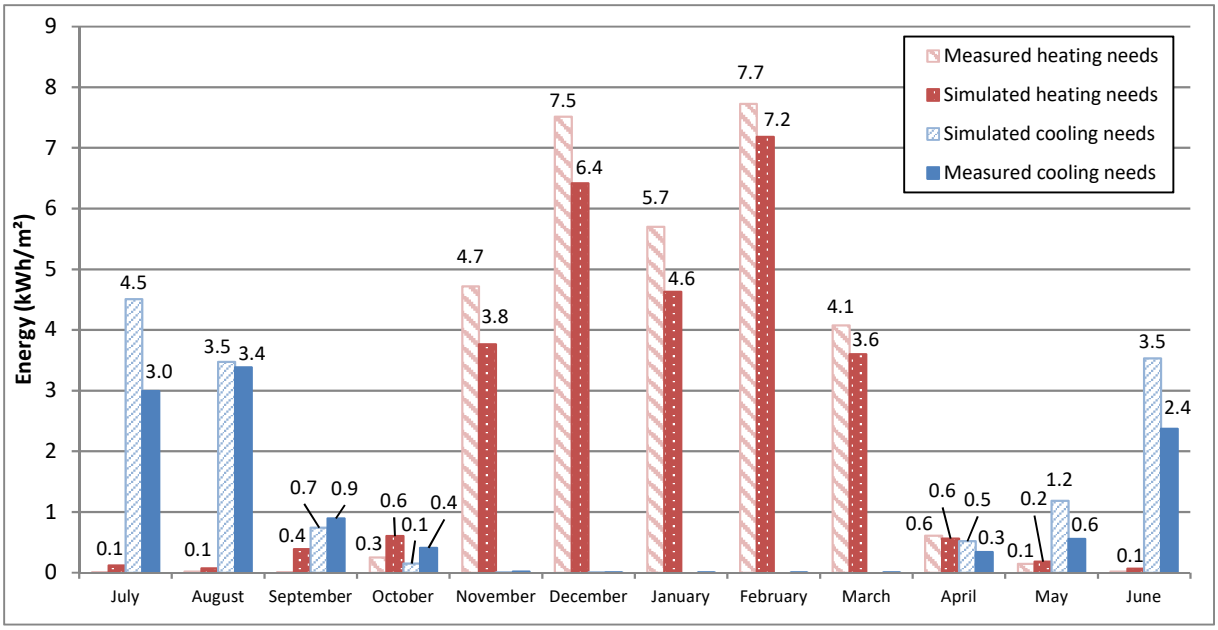

Figure 8: Monthly measured and simulated heating and cooling needs (the energy needs are zero in the absence of a label)

4.1.2. Hourly results

A similar analysis has been performed on an hourly basis to assess the simulation performance according to the hourly guidelines thresholds, results are presented on Figure 9. Similarly to the monthly results, the model cannot be considered as being calibrated, even if, except for IPMVP guidelines, the criteria are less stringent.

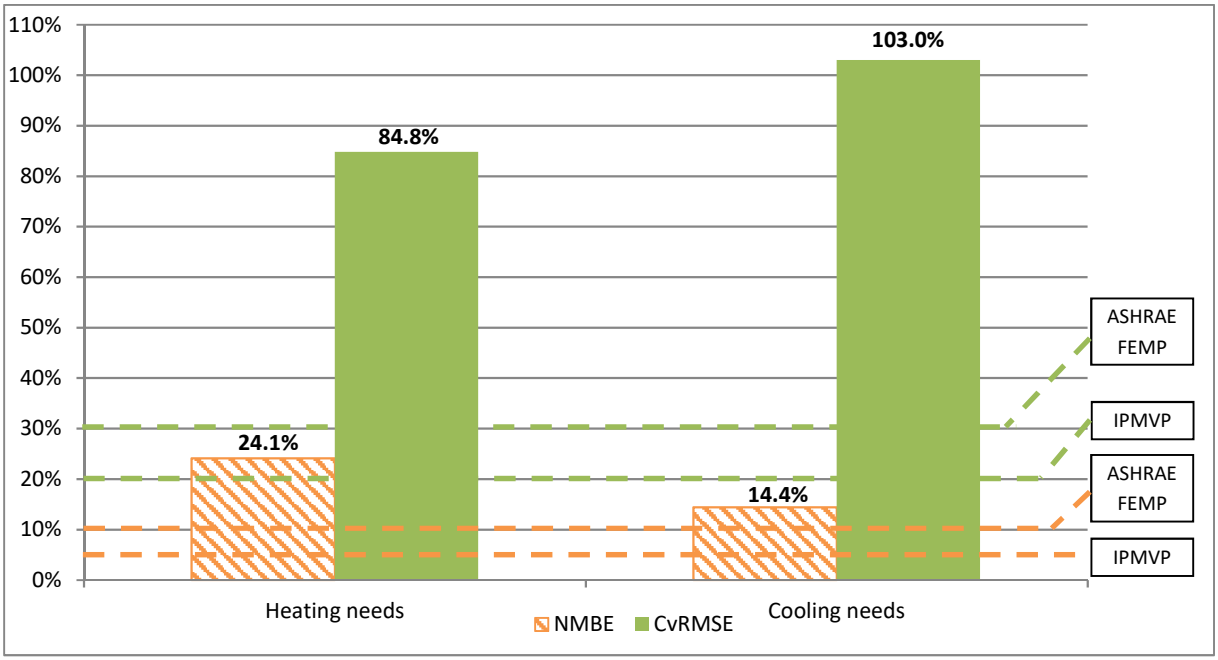

Figure 9: Calibration performance indicators in heating and cooling mode and guidelines thresholds for an hourly calibration

At this stage, in most cases where the information available on the building is limited, leading to level 1 or 2 calibration processes [30], the user goes through a tuning process. The main idea is to change the simulation inputs to match the outputs, in this case the annual or monthly energy consumption. In this study, the amount of available data allow to go into a more detailed dynamical analysis to understand the causes of such discrepancies on the energy consumption.

4.2. Dynamical power analysis

4.2.1. Heating power

In this section, measured and simulated dynamical heating and cooling powers profiles are compared over one week in heating mode and one week in cooling mode. Figure 10 shows the measured (black curve) and simulated (red dotted curve) evolution of the total radiant ceiling's heating power of Higashi building over one week of February. The daily profiles are clearly identifiable, from 6 am to 8 pm, corresponding to the times when the heating set point switches from unoccupied to standby, and vice versa. Heating starts earlier at the beginning of the week (on Sunday at 6 pm) to insure an acceptable temperature on Monday morning after a week-end without heating. This can be observed on the power profiles. The powers peaks around 2 am (green arrows in Figure 10) correspond to the night guard round during which he successively activate the offices occupancy sensors, switching the temperature set point to occupied mode and briefly activating heating. Even if the simulated and measured heating powers are in very good agreement, few discrepancies can be observed:

- The total simulated heating power is calculated by summing the radiant ceiling's powers of each zone of the building. Yet, the measured value is calculated by summing the powers of the ten offices areas measured by heat meters placed at the office area scale, on the main pipes. This means that the heat loss along the pipes supplying each zone are not taken in account in the simulation, resulting in a slightly underestimated heating power.

- The inlet heating water temperature is increased of 7 °C from 6 am to 9 am each morning of the week to have more available heating power during the morning needs peak. This is not represented in the simulation, leading to important discrepancies, framed in black in Figure 10.

- The radiant ceiling's valve is not simulated. Its opening time is not represented in the model. Secondly, the pipes and radiant ceilings' thermal inertia are also not simulated. Consequently, the time and energy needed to warm up the pipes and the radiant ceilings are not considered. Those two simplifications lead to a situation where the heating power is instantaneously available and directly emitted in the zone, which is obviously not true in reality. This can be seen on the power profiles during the morning peak (framed in red dotted line). This difference is coupled to the previously described phenomenon on Tuesday, Wednesday, Thursday and Friday. However, due to the early start of heating, the two phenomena can be dissociated on Monday.

These three main discrepancies tend to minimize the energy consumption, which is consistent with the observations previously made on an annual and monthly basis.

4.2.2. Cooling power

Figure 11 shows the measured (black curve) and simulated (red dotted curve) evolution of the total radiant ceiling's cooling power over one week of June. Daily profiles composed of two distinct night-time and daytime sub-profiles are observed. Night-time profiles correspond to the night cooling mode where temperature set point drops to 23 °C in the zones, inducing an important power shift. The cooling temperature set point is programmed to drop to 23 °C between 11 pm and 6 am (red arrows in Figure 13) if the measured zone's temperature is above 25 °C (hysteresis 0.5 °C) and if the night precede a weekday. This night-cooling mode allows to benefit from the lower outside

temperature and the absence of internal gain to lower the zones' temperatures during nighttime. The simulated and measured cooling powers are correctly correlated, however several discrepancies can be observed:

- The measured cooling power drops to zero during nighttime (downward red arrows). In night cooling mode, cooling power can be either provided by the magnetic levitation chiller or by adiabatic cooling using directly the chiller's condenser as an adiabatic cooling tower. After further investigation, it has been found that adiabatic cooling is not properly working on this building, causing the power to drop to zero. This dysfunction is obviously not represented in the simulation in which night cooling is supposed to work properly. However, night cooling mode is correctly simulated when adiabatic cooling is not used as it can be seen on Monday.

- A time gap between the simulated and measured cooling power can be observed at the end of each day (circled in blue dotted line). This could also be due to a wrong hypothesis regarding the occupancy density at the end of the day. An analysis has been conducted based on the occupancy sensors and people seem to leave their office later than expected, leading to internal gains being produced later in the day in reality than in simulation and to an underestimated cooling power during this period.

- A slight power drop can be observed on the simulated cooling power around 1 pm (upward green arrows). This correspond to the decrease of occupants during the lunch break. However, this phenomenon is not observed on the measured cooling power. This is most likely the result of a wrong assumption made about the occupancy density at this time of the day. People might leave their office for their lunch break more rarely than expected, generating more internal gains in reality than in simulation, leading to an underestimated cooling power during this period.

- The 3-way valve controlling the inlet water temperature of the radiant ceilings' network is not represented in the simulation. Consequently, the periodic oscillations observed on the measured cooling power during the whole week (two examples framed in green) are not visible on the simulated cooling power.

- Similarly, to what is observed in heating mode, the influence of the pipes and radiant ceiling's thermal inertia can also be seen during the night power peak coming with the night cooling mode's start.

The above discrepancies confirm the trend of underestimation observed in the annual and monthly cooling energy consumption data.

The discrepancies observed in Figure 7 and Figure 8 for the calculation of the annual and monthly energy needs were for the most part physically explained by the dynamical power analysis. However, although the model is tending to underestimate energy requirements, it is able to provide results that are closely correlated with measured values if another variable is considered (i.e. the temperatures of the zones).

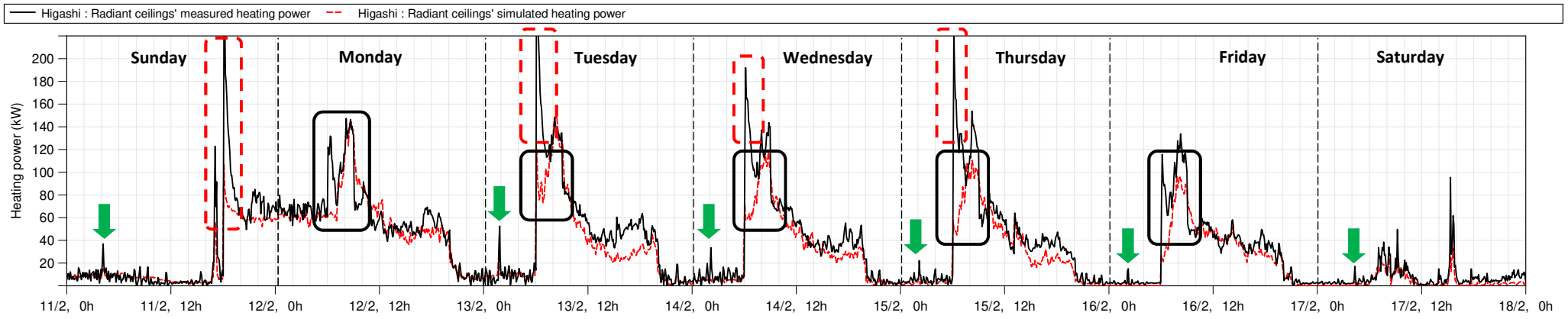

Figure 10: Evolution of the total measured and simulated power of the radiant ceilings over one week in February

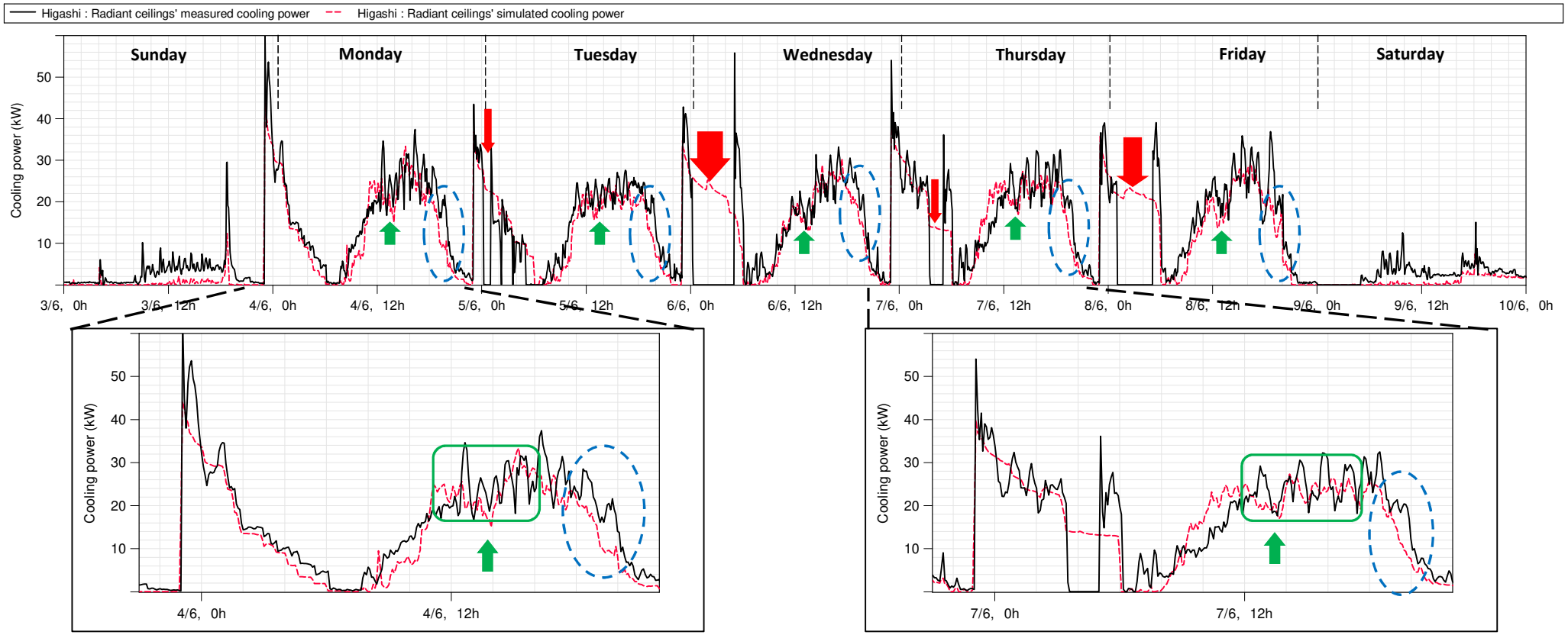

Figure 11: Evolution of the total measured and simulated power of the radiant ceilings over one week in June

4.3. Dynamical and statistical temperatures analysis

4.3.1. Zone scale

In this section, measured and simulated dynamical temperature profiles are compared over the same weeks used for the power analysis. A statistical analysis is also performed over the whole year of simulation.

Figure 12 shows the measured (black curve), simulated (red dotted curve), and set point (thin grey curve) temperatures evolution in heating mode over one week of February for one particular office of the building (office H2N.9, circled in Figure 5). The three heating set point levels, unoccupied, standby and occupied at 18 °C, 21 °C and 23 °C respectively, along with the early start of heating on Sunday are visible. The simulated and measured temperatures are closely correlated and the daily dynamical phenomena are correctly represented in the simulation, however several discrepancies at lower time scale can be identified:

- The temperature elevation during the heating phases in the morning is faster in the simulation due to the fact the pipes and radiant ceiling's inertia is not simulated.

- The temperature drop during the night can be faster in the simulation. This can also be explained by the pipes and radiant ceiling's thermal inertia. In reality, when heating stops, the energy stored inside is released in the zone, slowing down the temperature drop in the zone.

- The oscillations around the temperature set point are not observed in the simulation due to the radiant ceiling's control valve that is not modeled.

A similar analysis is conducted in cooling mode in Figure 13, where the simulated and measured temperatures are correctly correlated, with discrepancies up to 0.8 °C. The simulated and measured temperatures are correctly correlated, with deviations of up to 0.8 °C at most. Standby and occupancy setpoints can be identified around 27 °C and 25 °C respectively. Daytime temperatures are underestimated in simulation. This may be due to higher internal gains in reality, manually forced blinds in open position, or higher occupancy densities than the standard schedule. Indeed, as pointed out in [38] by means of a sensitivity analysis, those factors have a non-negligible impact on the building energy consumption for cooling. On the other hand, the temperature drop that follows a switch to night cooling mode (black arrows in Figure 13) seems slightly overestimated in simulation. Therefore, the night temperatures remain far away from the measured temperatures. This can hardly be explained by the inertia of the pipes and radiant ceilings, as the night cooling mode starts only a few hours after the daytime cooling stops, leaving little time to warm up.

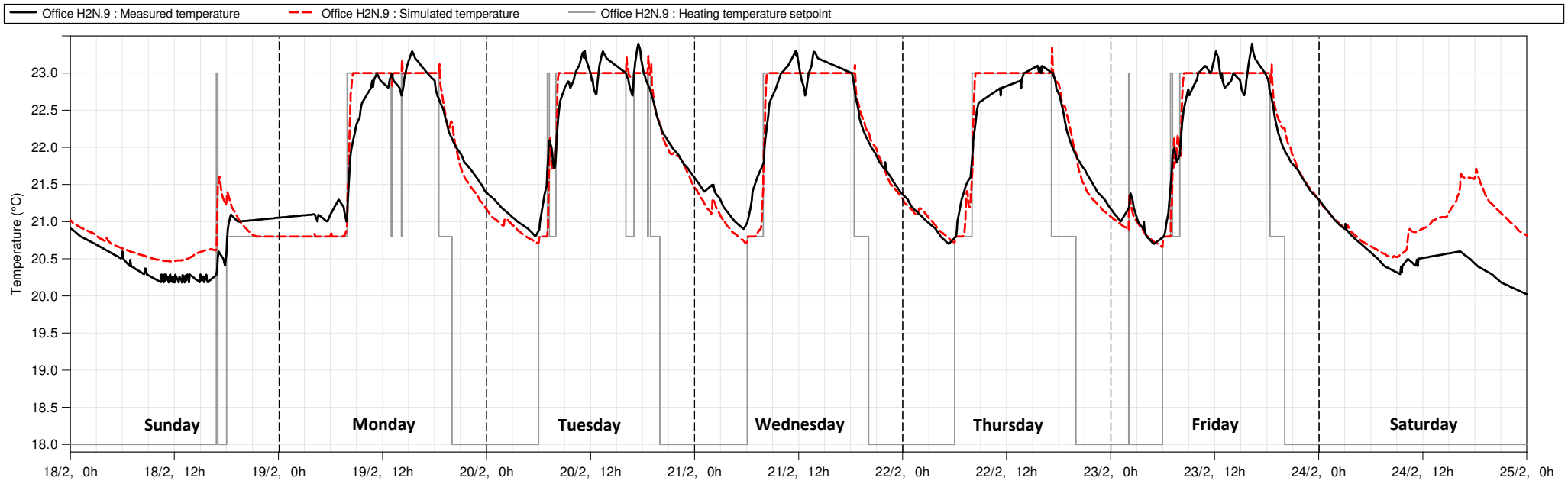

Figure 12: Measured and simulated temperatures evolution over one week of February in heating mode for Higashi's office H2N.9

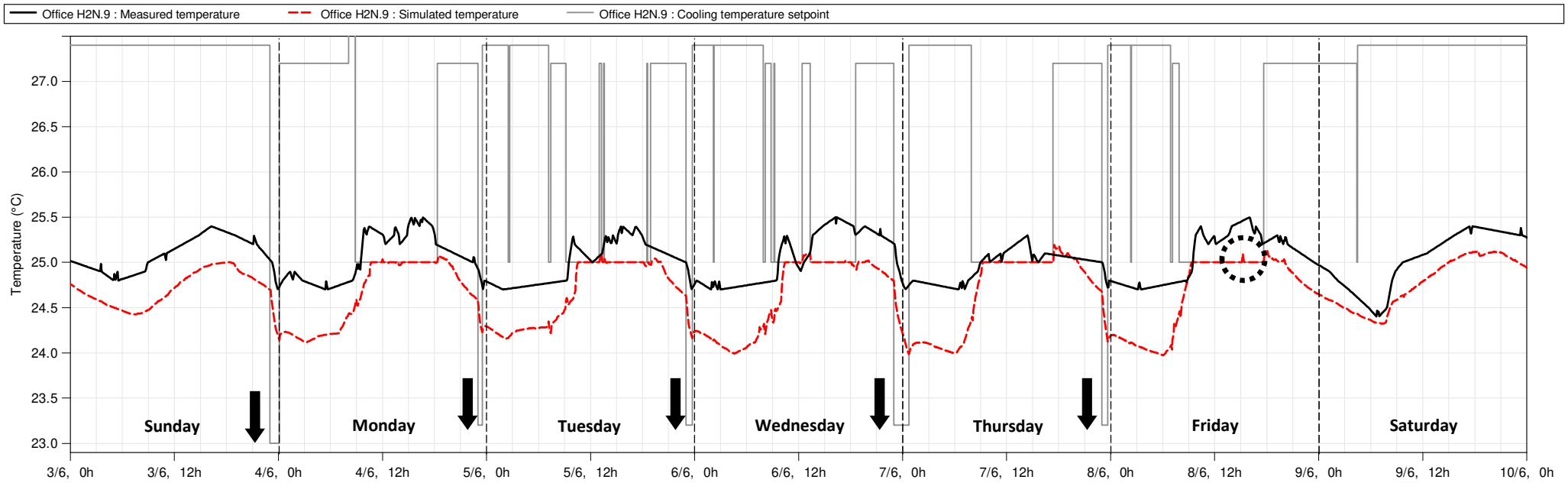

Figure 13: Measured and simulated temperatures evolution over one week of June in cooling mode for Higashi's office H2N.9

A satistical analysis has been done over the complete year of simulation, using the statistical indicators mentionned by the guidelines, within the limit of available temperature measurements. Results are gathered in Table 6. After identification and suppresion of non reliable data, 71.6 % of the year is available for comparison with simulated data. The NMBE and CvRMSE have been calculted using tempertaures in Kelvin, resulting in low values (0.007 % and 0.151 %, respectively). This questionned the use of NMBE and CvRMSE as performance indicators when working with variables that can be expressed in relative metrics such as temperature in degrees Celsius. In this case, rather than using an absolute unit (here the degree Kelvin) for the calculation of relative performance indicators, it may be more meaningful to keep using the relative metric and calculate an absolute error.

The error probability density of Higashi's office H2N.9 is represented in Figure 14 where 100 % of the points lie betwen -1.5 °C and 1.5 °C and 90 % are situated between -0.65 °C and 0.82 °C. Note that using absolute error is much more suitable for interpreting temperature data. Relative density scatter plot is shown in Figure 15. Several observations can be done regarding the data points repartition:

- Different groups of points can be identified depending on the temperature. In heating mode, data points are concentrated around the standby and occupied heating temperature setpoint, at roughly 21 °C and 23 °C respectively (circled in red). In cooling mode, data points are concentrated around the occupied cooling temperature setpoint, at roughly 25 °C (circled in blue dotted line).

- Simulation is in good agreement with the measurements for the points ranging around the heating standby temperature at 21 °C.

- Very good agreement is observed at 23 °C. The slight shifts correspond to the measured temperature oscillation around the setpoint due to the regulation operated by the V2V visible in Figure 12.

- An horizontal shift towards the right is observed at 25 °C. This means that the temperature setpoint is difficult to reach in reality, as observed in Figure 13. The installed cooling capacities have been dimensioned for a 28 °C setpoint. However, the setpoint has been modified to 25 °C during operation phase to meet the occupants' demands for summer comfort. Therefore, in case of high cooling requirements, the installed capacity may sometimes be insufficient to reach the setpoint temperature.

Generally speaking, energy needs and consequently zones' temperature tend to be more complicated to predict in cooling mode than in heating mode since the absolute powers values involved are far less important in summer (around a factor 3). This means that the building is much more sensitive to any internal or external perturbation like an internal or solar gain variation.

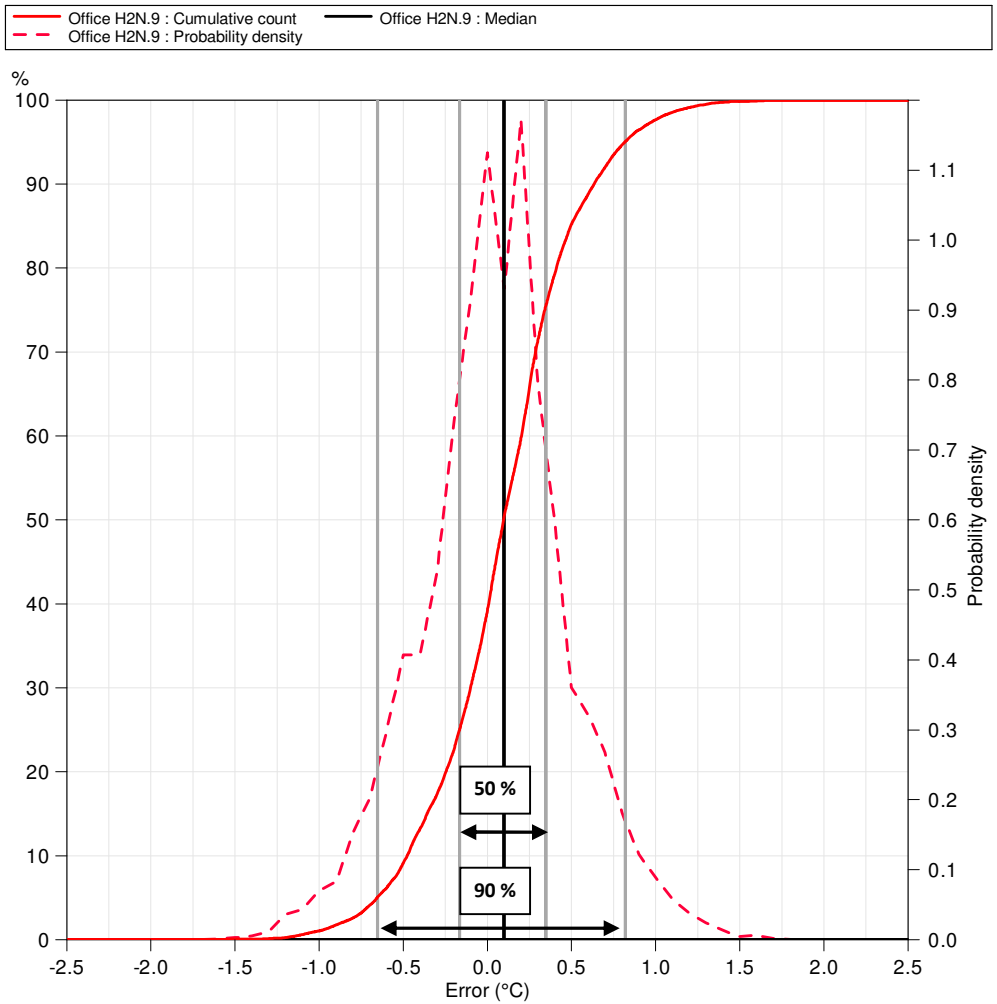

Figure 1: Distributed and cumulative probability density

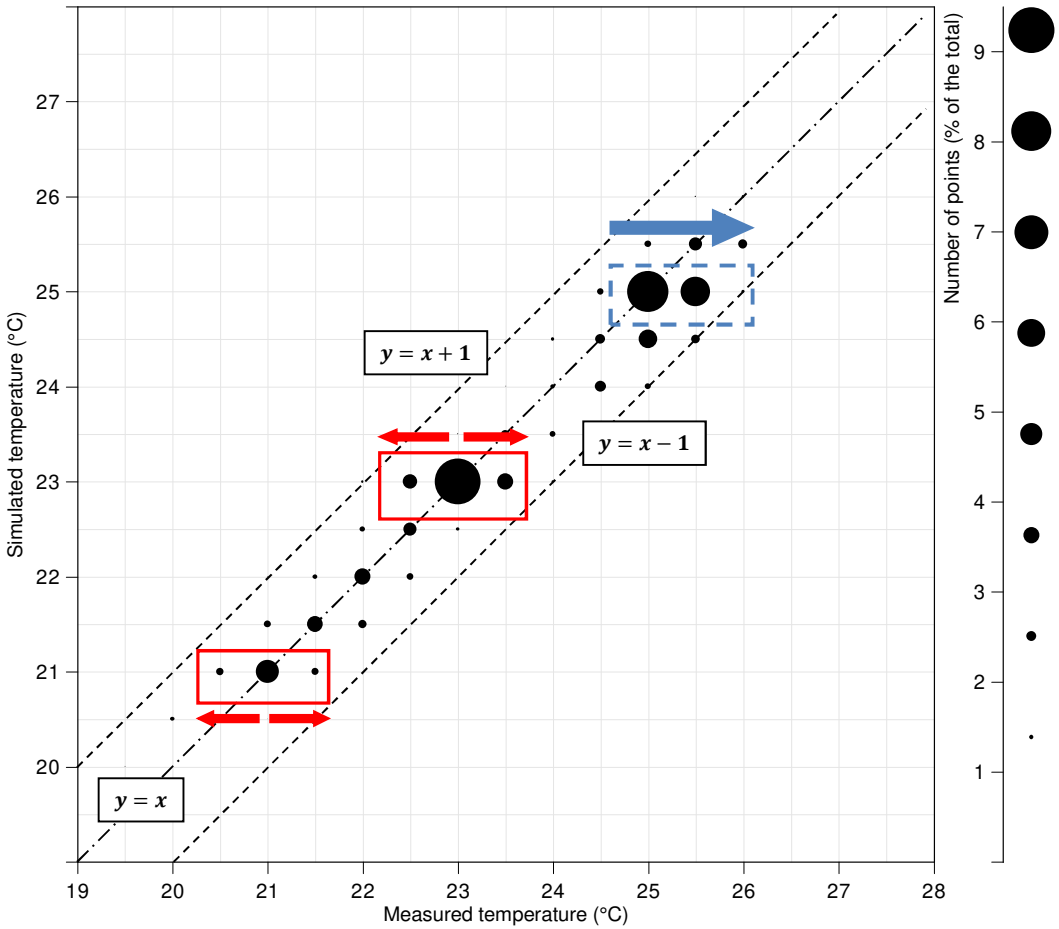

Figure 2: Simulated temperature according to the measured temperature

4.3.2. Building scale

A similar statistical analysis has been made at the building scale, meaning that every zone's simulated and measured temperatures have been gathered in one global dataset. Results are gathered in Table 6. In total, 63 % of the points are available for comparison among the 132 calibrated zones of the building. The 37 % that are not available are due to periods when the quality of the data does not allow its use due to malfunctions of the measuring equipment. NMBE, and CvRMSE values of 0.037 % and 0.394 % have been found. The same observations as before can be done regarding the use of those indicators for temperature data.

	Data available (%)	*NMBE (%)*	*CvRMSE (°C)*
Office H2N.9	71.6	0.007	0.151
Higashi building	63.0	0.037	0.394

Table 6: Statistical analysis results of the simulated and measured temperatures

The error probability density of Higashi's zones are represented in Figure 16 where 90 % of the data points are situated between -1.72 °C and 2.02 °C of error and 50 % lie between -0.41 °C and 0.84 °C. Relative density scatter plot is shown in Figure 17. Several general observations can be done regarding the data points repartition :

- In heating mode (circled in red), a majority of data points are centered around the $y = x$ curve and are included between -1 °C and +1 °C of error. This means that the zones's temperature are correctly simulated in heating mode.

- A shift towards the right is observed in cooling mode (circled in blue dotted line), meaning that the simulation tends to underestimate the zones's temperatures. This probably corresponds to situations where the cooling setpoint is reached in simulation and not in reality, due to several reasons previsouly detailed.

Box plots for each zone have been generated and are represented in Figure 18. The median value is represented by the box's center thick line, 50 % of the data set is contained in the box and 90 % lie between the lower and upper whiskers. For clarity purposes, the remaining 10 % of points lying outside the whiskers are not represented. The boxes are organized in ascending order according to the zone's RMSE, and their width depends on the number of measured temperature points available for comparison with the simulation. The zone previously analyzed (offices H2N.9) is represented in red (first box plot).

60 % of Higashi's zones have at least 90 % of data points lying between -2 °C and +2 °C. Most of the remaining zones still show acceptable results with error values very close from this interval. However, there are a few areas with significant errors, mainly due to specific uses (lunch areas, meeting rooms, rest rooms, hairdresser, and photocopying service) that are not taken into account in the simulation. Indeed, as previously mentioned, there are no electrical energy meters at the zone level. Therefore, the approach used consists in allocating to each zone a share of the total measured power at the office area level based on the area of the zone is not appropriate. Internal gains are therefore not properly taken into account. Further analysis, including on-site audits, would be essential to draw up a detailed inventory of the electrical equipment used in order to properly calibrate these zones.

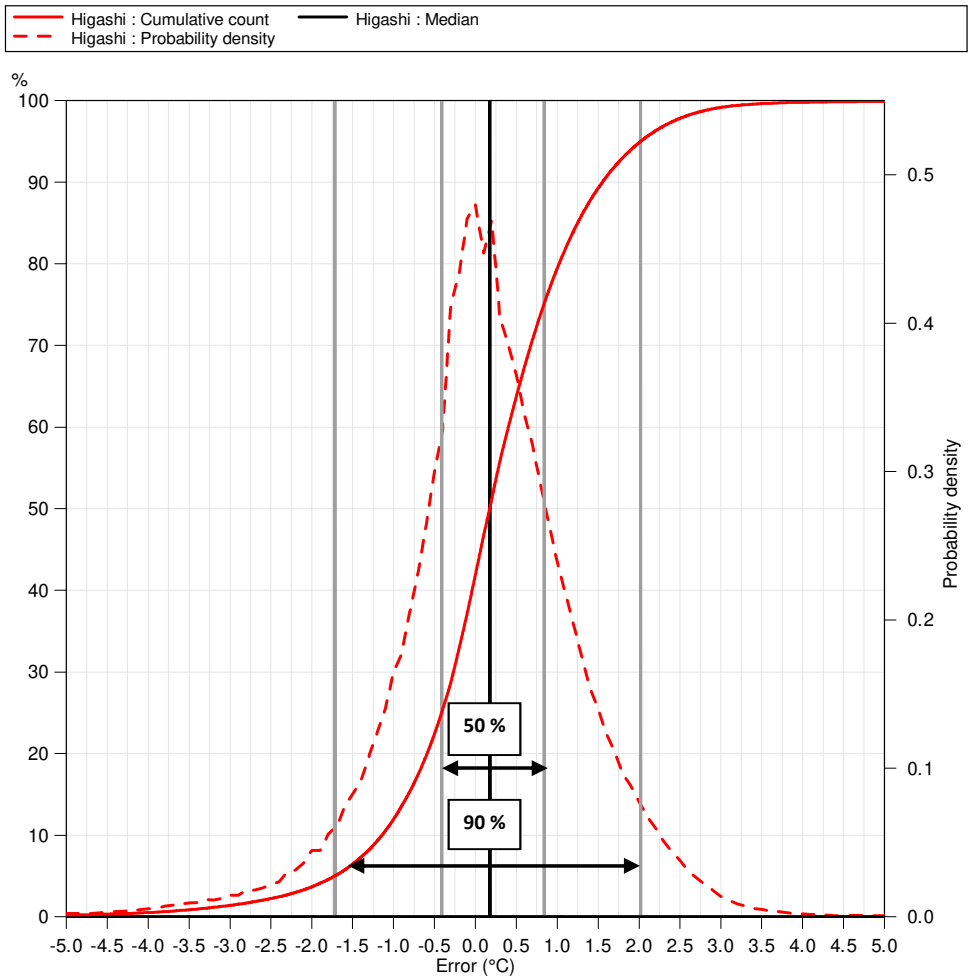

Figure 3: Distributed and cumulative probability density

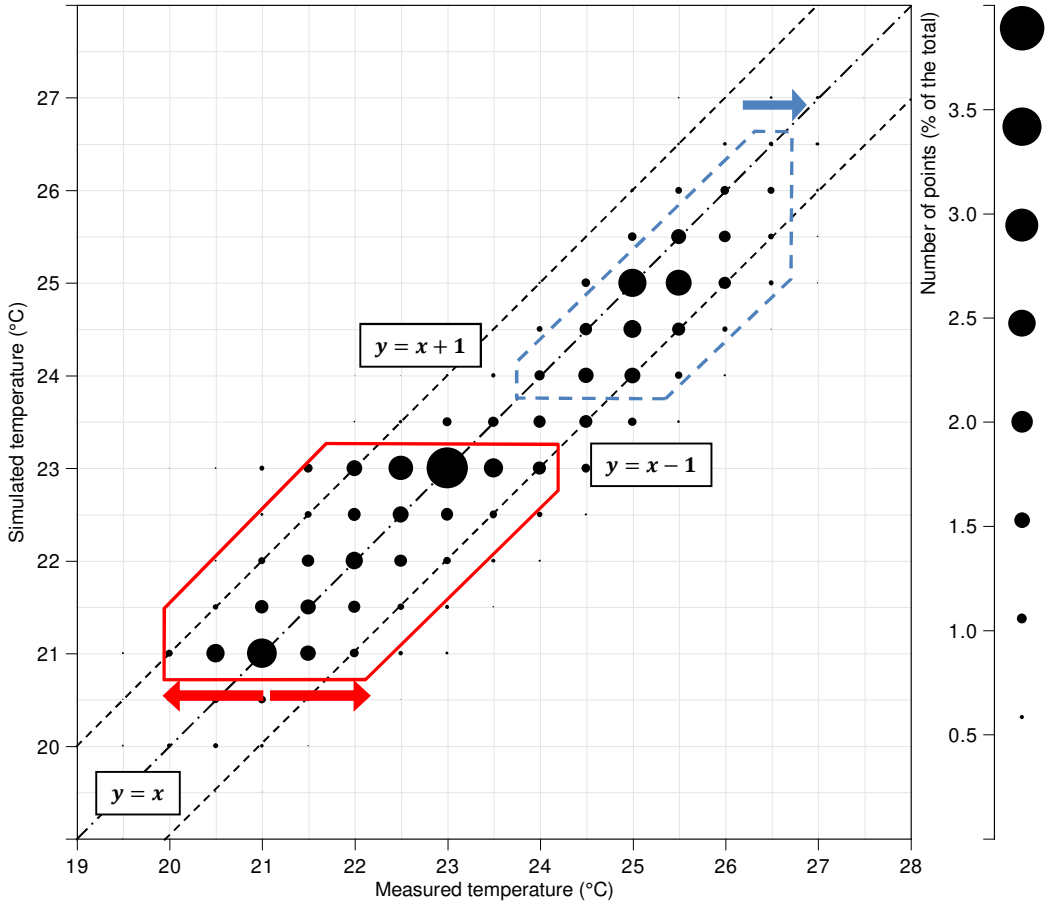

Figure 4: Simulated temperature according to the measured temperature

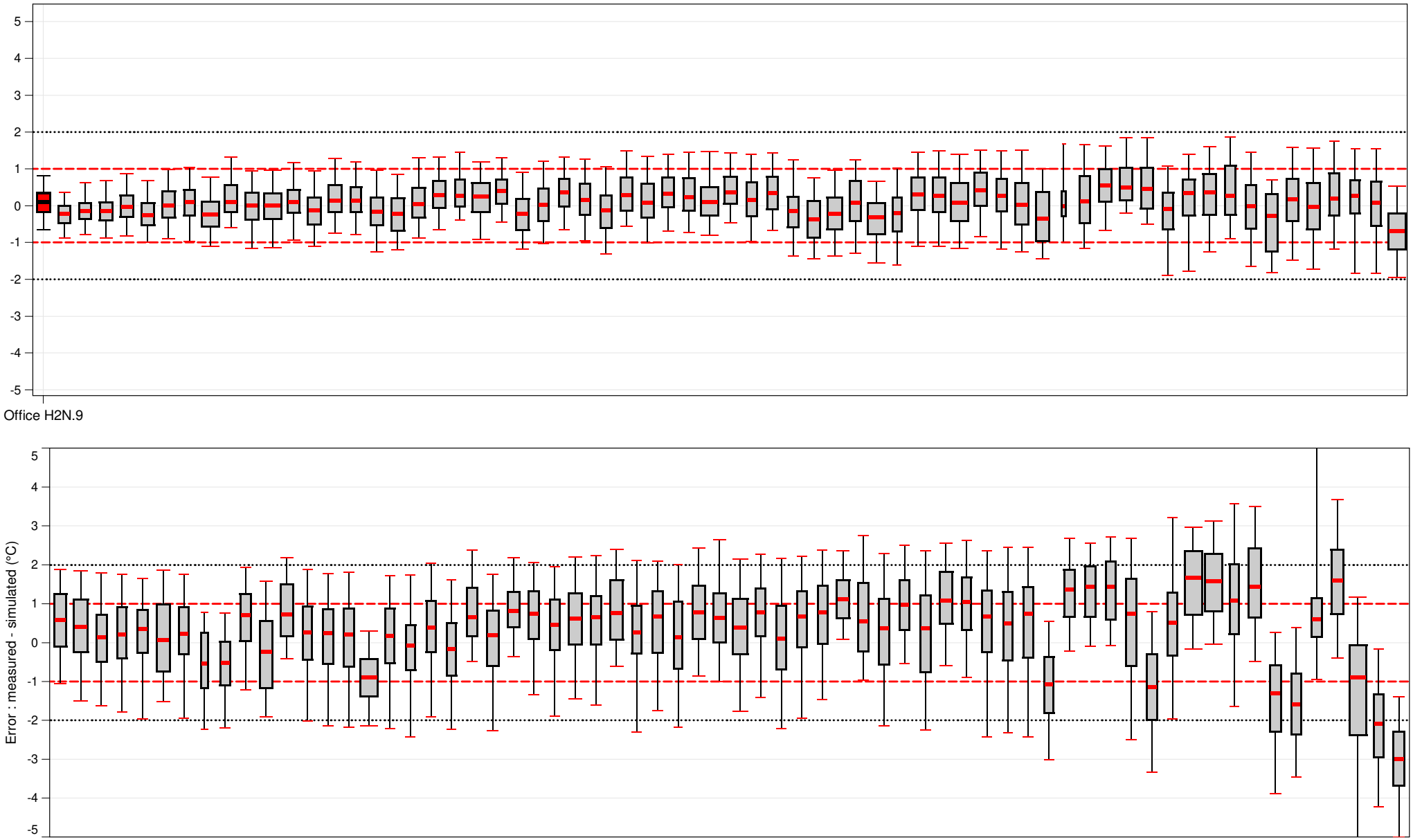

Figure 18: Box diagram of temperature errors in Higashi building zones

4.4. Discussion on the calibration results

The model showed a good correlation with measured values when considering zones temperatures along with heating and cooling powers. However, as shown in Figure 6, NMBE and CvRMSE value calculated on the energy needs where too high to consider the model as begin calibrated. The input parameters identified in the literature as having a significant influence have been determined based on measurements and/or construction data verified by a commissioning process. Even if the model is already very constrained resulting in a reduced margin for calibration, a few important parameters remain partially known and could potentially be adjusted by pursuing the calibration process to the second step where input parameters are tuned:

- Zone occupancy is calculated by coupling the theoretical number of occupants with a weekly occupancy density schedule. Knowing the exact number of occupants at each time and for each office would allow the internal gains to more accurate in the simulation and would probably improve the results. In fact, it was planned during the design phase that the occupancy sensors would be able to measure the number of people. However, this functionality was abandoned for privacy reasons. As a result, the exact number of occupants remains unknown.

- The internal gains generated by the electrical equipment are not known at the zone level, therefore they are determined according to the electrical power measured at the office area level and the surface of the zone. For office buildings with similar use from one zone to another, this assumption is acceptable. However, the simulation results for the few zones with a use that differs from the main use would certainly be improved if the internal inputs were specifically known during the simulation. This would require the installation of permanent electrical energy meters in each zone, making the monitoring installation more complex. A less invasive solution, such as a measurement campaign during a few weeks, would allow a better knowledge of the electricity consumption profiles of these particular zones. As their use is quite similar from one month to another, the results could be extrapolated over the whole year.

- As specified in section 3.3, the temperature is controlled on the operative temperature, calculated with a radiant fraction equal to 0.2. A higher radiant fraction would tend to attenuate the operating temperature variations. Given its influence on the dynamic behavior of the zones temperatures, this assumption could be subject to a tuning process. However, the temperature sensors used for control are placed in portable devices which can be moved as desired by the user. As the radiant temperature measured by the sensor is dependent on the configuration factors and on the radiant temperatures of the walls with which the device can exchange radiation, the latter is therefore highly dependent on the geometry at the time of the measurement, which is dependent on the device position in the zone that can change frequently. While the assumption that the position of the device would remain invariant over time is acceptable for a given area, it seems difficult to apply it to the scale of the building. Therefore, the tuning of the radiant fraction would have to be done zone by zone and could not be based on additional in-situ measurements. Moreover, this process would be difficult to control given the presence of other non-considered first-order parameters also influencing temperature dynamics, such as the inertia of the heat emitters.

Among the particularly influential input parameters identified in the literature and mentioned in section 3.2, three parameters have been identified as being adjustable during a calibration process in order to improve the performance of the model. Therefore, pursuing the calibration process on other parameters than those previously cited does not seem judicious, since three first-order parameters need to be adjusted first.

In addition, as stated in the previous sections, the majority of discrepancies are caused not by wrong or incomplete hypothesis but by the non-modeling of several physical phenomena related to energy production, distribution, and control systems (radiant ceilings and pipes thermal inertia, 2-ways valves opening time, control action of the 3-way valve on the main distribution pipe, change in the water inlet temperature, free-cooling dysfunction).Consequently, the modeling of those systems could really improve the simulation accuracy and thus helping to meet calibration indicators.

Based on the calibration conducted, several observations can be done:

- A calibration performance assessment should not be based solely on statistical indicators, but should be completed by a dynamic analysis that allows for a physical interpretation of the results to guide decision-making during the calibration process.

- Such analysis can only be possible with a precise knowledge of the building characteristics and its functioning as well as detailed monitoring data to consider transient phenomena. Indeed, several sources of discrepancies between measurement and simulation have been identified as transient phenomena involving a fast dynamic. As a result, data measured at an hourly time-step appears to be a minimum to conduct a calibration process based on physical analysis. Time-steps around the minute would be ideal for capturing system dynamics as well.

- Given the property of equifinality of BES models [35], considering only one criteria for calibration can be very misleading. Energy consumption may closely match measured data without being certain that other variables like zones temperatures are consistent. Consequently, calibration performance assessment should be conducted on several variables representing different aspects of the building's dynamic such as heating and cooling powers, temperatures, or humidity.

- Performance indicators recommended by the guidelines are not well suited for temperature values for which we recommend the use of indicators based on absolute error that are easier to understand and interpret.

Those observations are based on this case study only but are similar to those formulated in [37]. These are obviously not relevant in the case of calibrations carried out with a limited amount of data and information on the building (levels 1&2 according to [30]). Establish standardized methods remain a challenging task, given the diversity of BES models used in the industry along with the heterogeneity of the buildings, of their monitoring quality and consequently of the available data for calibration. With the current trend towards increased building instrumentation, building professionals would greatly benefit from further similar calibration work. Such work would support these observations in order to make the recommendations of the organizations evolving towards more appropriate methods and criteria.

Conclusion

This paper details the manual calibration of a 5 434 m² office building model simulated over one year at a five minutes time-step. Thanks to a complete instrumentation, the majority of input parameters where known, making a level 5 [30] calibration possible. Thus, the approach simply consists in feeding the model with as much measured data as possible instead of day-typed schedules without tuning unknown parameters to match measured data. The authors followed a similar approach as the one described in [36], including the creation of a version control file to store each revision of the model along with the corresponding modifications made and the obtained results, allowing to keep a complete history of the calibration process.

At first, in spite of a model mostly feed with measured data, it showed poor performances when considering maximum value on performance indicators (NMBE and CvRMSE) recommended by the guidelines. Instead of going through a classical tuning process to match monthly energy consumption data, the choice was made to focus on explaining those discrepancies using available building information and monitoring data. A dynamical heating and cooling powers analysis has been made, resulting in closely correlated measured and simulated data, especially in heating mode. It has been followed by a dynamical temperature analysis performed at the zone scale, completed by a global statistical analysis at the building scale for the 132 zones. Results showed very good agreement between measured and simulated temperature values with 90 % of data points ranging between -1.72 °C and 2.02 °C of error.

In the end, the model showed a good correlation with measured values when considering zones temperatures along with heating and cooling powers. The main discrepancies have been physically explained and are the result of the non-modeling of several physical phenomena, systems and control strategies. Consequently the model performance could hardly be improved though a classical tuning process acting on existing input data. At light of this calibration work, several observations have been done regarding the current calibration guidelines.

Although very time-consuming, this work emphasis the importance of a detailed calibration process to address the issue of “knowledge drift” encountered by the modelers [6]. Indeed, the knowledge of the calculation tool and the physical analysis of the results implied by a calibration work force the user to review his skills while refining his knowledge. Detailed calibration case studies also reaffirm the capacity of modelling tools to produce accurate results, whereas the many manifestations of the performance gap tend to produce the opposite.

References

[1] Ahmad AS, Hassan MY, Abdullah MP, Rahman HA, Hussin F, Abdullah H, et al. A review on applications of ANN and SVM for building electrical energy consumption forecasting. Renew Sustain Energy Rev 2014;33:102–9. doi:10.1016/j.rser.2014.01.069.

[2] Crawley DB, Hand JW, Griffith BT. Contrasting the capabilities of building energy performance simulation programs. Build Environ 2008;43:661–73. doi:10.1016/j.buildenv.2006.10.027.

[3] Foucquier A, Robert S, Suard F, Stéphan L, Jay A. State of the art in building modelling and energy performances prediction: A review. Renew Sustain Energy Rev 2013;23:272–88. doi:10.1016/j.rser.2013.03.004.

[4] Karlsson F, Rohdin P, Persson ML. Measured and predicted energy demand of a low energy building: Important aspects when using building energy simulation. Build Serv Eng Res Technol 2007;28:223–35. doi:10.1177/0143624407077393.

[5] Scofield JH. Do LEED-certified buildings save energy? Not really... Energy Build 2009;41:1386–90.

doi:10.1016/j.enbuild.2009.08.006.
[6] Imam S, Coley DA, Walker I. The building performance gap: Are modellers literate? Build Serv Eng Res Technol 2017;38:351–75. doi:10.1177/0143624416684641.
[7] Saltelli A, Tarantola S, Campolongo F, Ratto M. Sensitivity analysis in practice: a guide to assessing scientific models (Google eBook). 2004.
[8] Macdonald IA. Quantifying the Effects of Uncertainty in Building Simulation Iain Alexander Macdonald B . Sc ., M . Sc . A thesis submitted for the Degree of Doctor of Philosophy Department of Mechanical Engineering University of Strathclyde July 2002 2002.
[9] Wei T. A review of sensitivity analysis methods in building energy analysis. Renew Sustain Energy Rev 2013;20:411–9. doi:10.1016/j.rser.2012.12.014.
[10] Pang Z, O'Neill Z, Li Y, Niu F. The role of sensitivity analysis in the building performance analysis: A critical review. Energy Build 2020;209:109659. doi:10.1016/j.enbuild.2019.109659.
[11] Tian W, Heo Y, de Wilde P, Li Z, Yan D, Park CS, et al. A review of uncertainty analysis in building energy assessment. Renew Sustain Energy Rev 2018;93:285–301. doi:10.1016/j.rser.2018.05.029.
[12] Reddy TA. Literature Review on Calibration of Building Energy Simulation Programs : Uses , Problems , Procedur ... ASHRAE Trans 112(1) 2006;112:226–40.
[13] Coakley D, Raftery P, Keane M. A review of methods to match building energy simulation models to measured data. Renew Sustain Energy Rev 2014;37:123–41. doi:10.1016/j.rser.2014.05.007.
[14] De Wit S, Augenbroe G. Analysis of uncertainty in building design evaluations and its implications. Energy Build 2002;34:951–8. doi:10.1016/S0378-7788(02)00070-1.
[15] Menberg K, Heo Y, Choudhary R. Sensitivity analysis methods for building energy models: Comparing computational costs and extractable information. Energy Build 2016;133:433–45. doi:10.1016/j.enbuild.2016.10.005.
[16] Clarke JA, Yaneske PP, Pinney AA. The harmonisation of thermal properties of building materials. BEPAC Res Rep 1990;2:1–87.
[17] Heo Y, Choudhary R, Augenbroe GA. Calibration of building energy models for retrofit analysis under uncertainty. Energy Build 2012;47:550–60. doi:10.1016/j.enbuild.2011.12.029.
[18] Qinpeng W, Godfried A, Yuming S. The Role of Construction Detailing and Workmanship in Achieving Energy-efficient Buildings. Constr Res Congr 2014 2020:2224–33. doi:doi:10.1061/9780784413517.226.
[19] Wang L, Mathew P, Pang X. Uncertainties in energy consumption introduced by building operations and weather for a medium-size office building. Energy Build 2012;53:152–8. doi:10.1016/j.enbuild.2012.06.017.
[20] Tabares-velasco PC. Time-step considerations when simulation dynamic behavior of high-performance homes. 2013.
[21] Santos GH, Mendes N. Analysis of numerical methods and simulation time-step effects on the prediction of building thermal performance 2004;24:1129–42. doi:10.1016/j.applthermaleng.2003.11.029.
[22] Guyon G. Role of the model user in results obtained from simulation software program. Int. Build. Simul. Conf., 1997.
[23] Haberl JS, Abbas M. Development of Graphical Indices for Viewing Building Energy Data: Part I. J Sol Energy Eng 1998;120:156–61. doi:10.1115/1.2888064.
[24] Haberl JS, Bou-Saada TE. Procedures for Calibrating Hourly Simulation Models to Measured Building Energy and Environmental Data. J Sol Energy Eng 1998;120:193–204. doi:10.1115/1.2888069.
[25] Haberl JS, Abbas M. Development of Graphical Indices for Viewing Building Energy Data: Part II. J Sol Energy Eng 1998;120:162–7. doi:10.1115/1.2888065.
[26] Liu G, Liu M. A rapid calibration procedure and case study for simplified simulation models of commonly used HVAC systems. Build Environ 2011;46:409–20. doi:10.1016/j.buildenv.2010.08.002.
[27] Asadi S, Mostavi E, Boussaa D, Indaganti M. Building energy model calibration using automated optimization-based algorithm. Energy Build 2019;198:106–14. doi:10.1016/j.enbuild.2019.06.001.
[28] Yang T, Pan Y, Mao J, Wang Y, Huang Z. An automated optimization method for calibrating building energy simulation models with measured data: Orientation and a case study. Appl Energy 2016;179:1220–31. doi:10.1016/j.apenergy.2016.07.084.
[29] Chong A, Lam KP, Pozzi M, Yang J. Bayesian calibration of building energy models with large datasets. Energy Build 2017;154:343–55. doi:10.1016/j.enbuild.2017.08.069.
[30] Fabrizio E, Monetti V. Methodologies and advancements in the calibration of building energy models. Energies 2015;8:2548–74. doi:10.3390/en8042548.
[31] IPMVP New Construction Subcommittee. International Performance Measurement & Verification Protocol: Concepts and Options for Determining Energy Savings in New Construction. vol. III. 2003.
[32] ASHRAE. ASHRAE Guideline 14-2002: Measurement of Energy and Demand Savings. 2002.
[33] FEMP. M & V Guidelines: Measurement and Verification for Federal Energy Projects. Version 3.0. 2008.
[34] Ruiz GR, Bandera CF. Validation of calibrated energy models: Common errors. Energies 2017;10. doi:10.3390/en10101587.
[35] Beven K. A manifesto for the equifinality thesis. J Hydrol 2006;320:18–36. doi:10.1016/j.jhydrol.2005.07.007.
[36] Raftery P, Keane M, O'Donnell J. Calibrating whole building energy models: An evidence-based methodology. Energy Build 2011;43:2356–64. doi:10.1016/j.enbuild.2011.05.020.

[37] Raftery P, Keane M, Costa A. Calibrating whole building energy models: Detailed case study using hourly measured data. Energy Build 2011. doi:10.1016/j.enbuild.2011.09.039.
[38] Gaetani I, Hoes PJ, Hensen JLM. Estimating the influence of occupant behavior on building heating and cooling energy in one simulation run. Appl Energy 2018;223:159–71. doi:10.1016/j.apenergy.2018.03.108.