Development of a short duration method to assess the envelope thermal performance of multi-family housings

Lorena de Carvalho Araujo, Simon Thébault, Thomas Recht, Alain Sempey, Patrick Schalbart, Michaël Cohen, Laurent Mora

Development of a short duration method to assess the envelope thermal performance of multi-family housings

Lorena de Carvalho Araujo[a,b], Simon Thébault[a], Thomas Recht[b], Alain Sempey[b], Patrick Schalbart[c], Michaël Cohen[c], Laurent Mora[b]

[a] *CSTB, 84 Av. Jean Jaures, Champs-sur-Marne, 77420, France*
[b] *University of Bordeaux, I2M, CNRS, Arts et Metiers ParisTech, 351 Cours de la Liberation, Talence, 33405, France*
[c] *MINES ParisTech, Centre for Energy Efficiency of Systems, 5 rue Léon Blum, Palaiseau, 91120, France*

Abstract

Building energy efficiency is a key factor in reducing CO_2 emissions. For this reason, EU member states have developed thermal regulations to ensure building thermal performance. These results are often based on results achieved with building simulation software during the design stage. However, the actual thermal performance can deviate significantly from the predicted one, and this difference is known as the energy performance gap. Accurate indicators of the actual thermal performance are a valuable tool to guarantee building quality. These indicators, including the heat transfer coefficient (HTC) and the heat loss coefficient (HLC), can be estimated by the application of *in situ* methods. As multi-family housing and tertiary sector buildings are an important part of the building stock, mature methods to measure their thermal performance are needed. This paper presents a short-duration method for assessing the HTC in large building typologies using a sampling approach. The method was applied in a four-storey building model under different conditions to study the limits of the method and to improve indicator bias and uncertainty. Indicator quality was strongly influenced by the external weather conditions, the temperature variation during the protocol and the heat exchange with the adjacent apartments. Under winter conditions and with stable indoor temperatures, the method had a high accuracy when the protocol was applied for half a day. It is recommended that the protocol be used over two days to improve indicator quality under less favorable test conditions.

Keywords: Building envelope thermal performance, HTC estimation, multi-family housing

1. Introduction

The energy-performance requirements stipulated in building standards and regulations have gradually been increasing, and this should translate into lower energy consumption to provide a given level of building services. Despite these efforts, the as-built energy performance does not always meet what was predicted in the design phase (Zou & Alam 2020; Majcen et al. 2013). This discrepancy between the design and reality is known as the energy performance gap, a phenomenon that has been widely reported in the literature, with the extent of the energy gap varying between different studies (Johnston et al. 2015; Pappalardo & Reverdy 2020; Calì et al. 2016).

The analysis of building energy performance can be done at different scales, from building components to the whole building, including the contributions from occupants and HVAC systems (Roels 2017). This work focuses on an intermediate scale, the building envelope thermal performance, which has shown to have a significant influence in indoor human thermal comfort (Liu et al. 2018). Only the building fabric is concerned, other factors that influence final energy consumption are not taken into account, such as systems, weather conditions and the behaviour of the occupants. Two indicators are commonly used to describe the thermal behaviour of the envelope: the heat loss coefficient (HLC) and the heat transfer coefficient (HTC). These are useful indicators to rate building envelope quality, and are extensively discussed in the literature (Rouchier et al. 2019; Díaz-Hernández et al. 2020; Jiménez 2016). The HLC quantifies the amount of energy required in the steady-state to maintain a one-degree temperature difference between the interior and exterior of the building (Bouchié et al. 2015). The HTC is only related to the envelope thermal losses by transmission to the exterior, and is equivalent to the HLC minus the thermal losses by advection, which are related to infiltration (H_{inf}) through the envelope (Bouchié & Ibos 2020). All these indicators have W/K units, and their mathematical formulae are represented in equations 1 and 2. The indices i, j, k represent each point, linear and surface element of the envelope going from 1 to p, q, r, respectively.

$$HTC = \sum_{j=1}^{q} \chi_j + \sum_{i=1}^{p} \psi_i \times L_i + \sum_{k=1}^{r} U_{w_k} \times A_{w_k} \tag{1}$$

where:

- χ_j is the heat flow from a point thermal bridge j (W/K);
- ψ_i is the heat flow of a linear thermal bridge i (W/m.K);
- L_i is length of a linear thermal bridge i (m);
- U_{w_k} is the thermal transmittance of an external homogeneous surface k (W/(m^2.K));
- A_{w_k} is the area of an external homogeneous surface k (m^2).

$$HLC = HTC + H_{inf} \tag{2}$$

Although these indicators can be calculated mathematically , some terms from equation 1, such as the U-values and thermal bridges, may differ between an actual building and its design. Assessment of HTC and HLC values is important for reducing the performance gap and in the long run may increase the quality of building construction. At building handover, knowledge of these values helps to ensure project results and to identify possible malfunctions that could lead to excessive energy consumption when the building is in the operational phase. For construction companies, it could be used for internal quality control and as feedback to support improved building design. Knowledge of HTC and HLC can also be used as a tool to inform occupants, owners and public authorities about the quality of building insulation, which in turn can be useful for public policy management, the allocation of subventions and the awarding of quality labels. (Ziour & Calberg-Ellen 2020). For this reason, the use of tests and agreed protocols for *in situ* measurement of envelope performance after building handover can be an alternative to evaluate these parameters. Different methods methods for assessing HTC and HLC based on various protocols and estimation processes are available

nowadays. Most of them rely on a heat balance for the building treated as a single zone, shown in equation 3 Senave et al. (2019b).

$$C \times \frac{\partial T}{\partial t} = \phi_h + \phi_{int} + \phi_{sol} + \phi_v + \phi_{inf} + \phi_{tr} \tag{3}$$

where:

- T is the average indoor temperature of the zone (K);
- C is the effective heat capacity of the zone thermal mass (J/K);
- ϕ_h is the heat flow from heating systems (W);
- ϕ_{int} is the internal heat gains (W);
- ϕ_{sol} is the solar gains (W);
- ϕ_v is the heat exchange due to ventilation (W);
- ϕ_{inf} is the heat exchange due to infiltration through the building envelope (W);
- ϕ_{tr} is the heat flow due to transmission through the building envelope (W).

Based on this description, two main families of methods can be identified. The first considers the thermodynamic behaviour of the thermal mass of the building, while the second relies on a steady-state hypothesis. To neglect the dynamic term in the heat balance, it is assumed that the total heat stored and released by the thermal mass of the building over a specific period are equal (Senave et al. 2019a). This hypothesis is consistent with the use of long enough time steps in comparison with a thermal cycle (Bauwens & Roels 2014). On the one hand, such a hypothesis reduces the heat balance to an algebraic equation, allowing modelling solutions such as linear regression and mean ratios to be used. On the other hand, it implies longer protocols so that enough data points can be acquired to characterise the envelope performance. When quick results are required, dynamic methods are more suitable (Madsen et al. 2021).

Inside the dynamic methods, two subgroups can be distinguished according to the protocol requirements for building vacancy. As the envelope thermal performance is related to ϕ_{tr} and ϕ_{inf} (equation 3), the other heat flows should be controlled or estimated during a test to avoid biasing the HTC and HLC estimation. The first strategy is to apply the method in a vacant building. This allows to minimise the other heat flows through protocol requirements, such as closing shutters (reducing ϕ_{sol}), stopping the HVAC systems (controlling then ϕ_h and ϕ_v) and avoiding the internal gains from occupants and appliances (ϕ_{int}). Since the thermal flows not related to the building envelope are neutralized, the number of boundary conditions is reduced and the modelling step is simplified. However, these are considered invasive methods, since they interfere with the normal use of the building. They are often more suitably applied immediately after construction or important retrofit actions, while buildings are not yet occupied. However, this type of protocol should be as short as possible due to the vacancy cost (Deb et al. 2021).

Another strategy consists in applying the protocol in an occupied building, which releases the constraint on protocol duration. However, the behaviour of occupants results in a series of complex thermal phenomena in the building during a protocol application, such as windows opening, use and control of HVAC system, use of electrical appliances and lightening, etc. There is a need for more instrumentation to evaluate the relevant parameters related to these heat flows, and their models might be more complex. Further studies on this strategy have

been conducted in the framework of the freshly concluded Annex 71, from the International Energy Agency, "Building Energy Performance Assessment Based on In-situ Measurements", and related works (Bauwens et al. 2021; Rasmussen 2020; Zhang et al. 2022). In addition, in occupied buildings, it is harder to defend the hypothesis of a homogeneous zone temperature presented in equation 3 (Senave et al. 2019b), since occupant actions can affect at any time the temperature gradient inside the building.

Invasive methods allow a higher control of the building state during a protocol application, simplifying the measurement and modelling of the thermal phenomena. In addition, they are a step further in maturity, having been widely deployed by different organisations, specially in the French context. For this reason, this paper focuses on the use of such methods to assess building envelope thermal performance. Different invasive dynamic methods are available nowadays, such as ISABELE (Bouchié et al. 2014; Schetelat 2014; Thébault. & Bouchié 2018), EPILOG (PACTE 2017; de Carvalho Araujo 2018), QUB (Alzetto et al. 2018b; Mangematin et al. 2012; Alzetto et al. 2018a), and Bacher & Madsen Bacher & Madsen (2011).

Although the cited methods present specificities in their protocols and estimation processes, they share common principles. Figure 1 shows a general scheme to represent these methods. The protocol is the part developed *in situ*, consisting of equipment, requirements, scenarios and an adequate duration for acquiring enough informative quality data. In invasive methods, the building is commonly heated to increase the temperature difference with the environment and to enhance the signal of heat losses through the building envelope. Data related to indoor temperature, released power, weather conditions and others are collected during the protocol application. This data can then be pre-processed to meet the necessary format used in the estimation process, where modelling algorithms are applied.

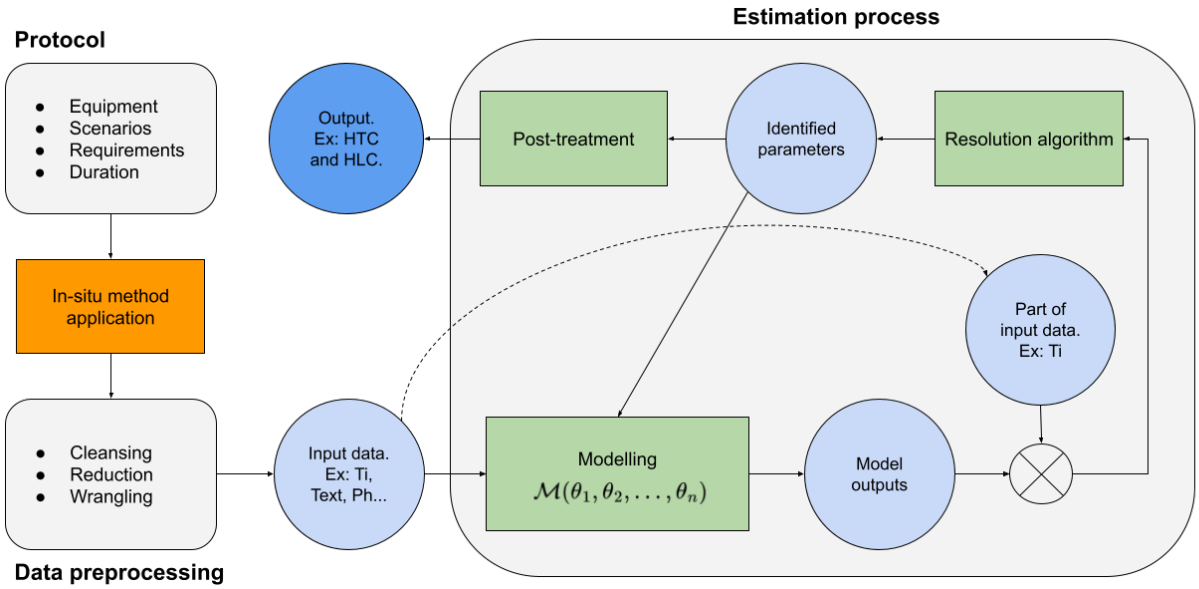


Figure 1: General scheme of the sub-steps of a method to assess building envelope thermal performance with *in situ* measurements.

These methods are used to solve inverse problems, where the physical phenomena should

be initially modelled, describing how the parameters of the model translate into experimentally observable effects. Then the phenomena take place in reality and the approach consists in tuning the parameters so that the model outputs are similar to the experimental data (Xu & Darve 2021). The modelling part can be composed of one or many complementary models, considering a model as "a representation of structure in a physical system and/or its properties" (Hestenes 1997). Once a model is defined, its structure and implications can be analysed to extract valuable information, this process is known as model-based inference (Ford 2009). The resolution algorithms can vary among the modelling techniques, as for the static methods they are the inclination of a line defined by aggregated data (Roels et al. 2017; Uriarte et al. 2019) or by linear regression (Zayane 2011; Bauwens & Roels 2014). Optimisation algorithms are a special type of resolution algorithm, where the parameters are optimised based on a criterion, such as the ordinary least-squares (de Carvalho Araujo 2018) and the maximum likelihood (Roels et al. 2015; Thébault 2017), which are commonly used for dynamic methods. Once the model parameters are estimated, a post-treatment can be applied to calculate the final indicator, such as the HTC or the HLC.

Although different methods are available to assess building envelope thermal performance, most of them have been conceived to be applied to single-family houses. In the case of larger buildings, such as multi-family housing, the scale of the tested area can impose major difficulties to the application of methods requiring extensive building instrumentation. Some simplified methods, such as the energy signature (Nordström et al. 2013) static method, have already been applied to larger buildings. As during protocol application the building envelope is not under steady-state regime, their total measurement duration is longer, ranging from several weeks to several years (Madsen et al. 2021), which can be a constraint when fast validation of construction quality is desired.

The typology composition of residential building stock varies among the EU member states, but in average multi-family housing (MFH) accounts for 50 % of the residential building stock in Europe (European Comission 2018). In France, 44 % of the residential building stock is composed of MFH (ADEME 2013). MFH represents an expressive share of building stock and therefore presents a relevant potential for energy savings. This reinforces the importance of assuring the envelope thermal performance among those building typologies.

Presently, there are no mature fast duration methods adapted to large building typologies, which presents a potential for developing new methods to answer this demand. Considering the extra complexities related to MFH size, a protocol applicable to vacant buildings was developed, thus avoiding secondary thermal phenomena during the test. Invasive methods must be as fast as possible to limit vacancy costs. RC network thermal models are strong candidates to develop a short measurement time method, since they take into account the thermal dynamic phenomena related to the building thermal mass. Among the proposed methods using RC models, EPILOG and ISABELE methods present flexibility regarding the model order and season of protocol application. SEREINE method is an evolution of these methods, presenting characteristics from both. ISABELE and SEREINE methods present a methodology to estimate the uncertainties in the final indicator, giving information on the reliability of the results. Nevertheless, ISABELE was conceived for single-family house applications and SEREINE method is still under development. In order to make them applicable to MFH, it is thus necessary to adapt the method protocol and estimation process.

One of the main challenges facing large building typologies regards their dimensions and

the consequent difficulties to apply an *in-situ* protocol. One approach could be to instrument the whole building to apply the protocol to the whole envelope. This approach has been studied in (de Carvalho Araujo et al. 2021), showing good reliability when applied to a mid-rise MFH. However, its application is limited to cases where the whole building is available and vacant for testing. The volume of equipment necessary for a test can also constrain the application, depending on the heating system available in the building and the heating power required for the test.

An alternative to verify the envelope thermal performance of large buildings is to partially test the building envelope, with a sampling approach. In this approach, sampled apartments in the building are tested, and the tested area has then similar dimensions to single-family houses, which facilitates the protocol application.

This paper shows one strategy to adapt the SEREINE method to the context of large buildings, using a sampling approach to assess the HTC of building apartments. To verify the reliability of the method with a referential indicator, this paper focuses on the method applied through building simulations.

Section 2 presents the sampling approach in the context of the SEREINE method and the scientific tools used to study it with building simulations. Section 3 concerns the building model and the virtual environment used to apply the method. In section 4, the experimental plans used to study the quality of the method are presented and the results are discussed. Finally, in section 5 the main conclusions are drawn and recommendations are given to ensure the reliability of the method.

2. Materials and Methods

To meet the challenge of assessing of the envelope thermal performance of large buildings, an adaptation of the SEREINE method is suggested with the use of a sampling approach. Part 2.1 presents the sampling approach and its main challenges. In part 2.2, the principles of the SEREINE method are presented, highlighting the features of the sampling approach. Part 2.3 shows the tools used to qualify the developed method in a simulation environment.

2.1. Sampling approach

The sampling approach consists of the thermal performance assessment of parts of the building envelope. Figure 2 illustrates the floor of a MFH, in which the protocol is applied to the southeast apartment. The measured perimeter, is composed of a part of the building envelope (in red), shared walls with the common area (in green) and with adjacent apartments (in yellow). Depending on which floor the tested apartment is situated in the building, it can have either the envelope or shared walls above and underneath it. Here, shared walls design the walls, ceiling and floor facing an adjacent area different from the exterior environment.

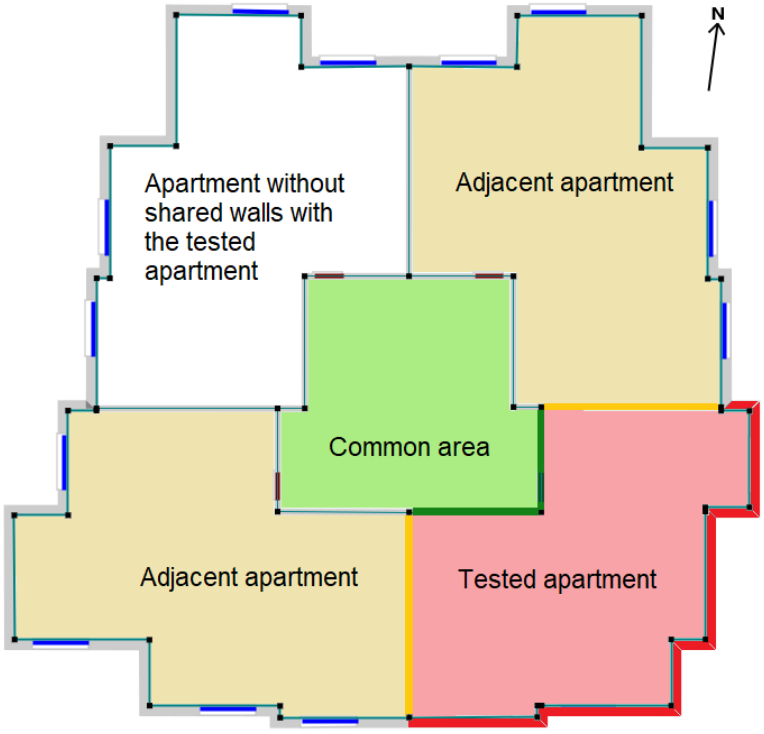


Figure 2: Example of the measured perimeter and the boundaries in a sampling approach method application.

Shared walls facing heated areas are seldom insulated for thermal purposes, although they might have a thin layer of acoustic insulation in recent buildings. Anyway, the thermal transmittance (U-value in $Wm^{-2}K^{-1}$) of these walls would be comparatively higher than the exterior walls in high-performance buildings. For this reason, the heat flow density passing through the shared walls (ϕ_{shar} in W) can be significant during a protocol application, depending on the temperatures applied on both sides of the wall. Since the objective of the method is to characterise the envelope thermal performance, if ϕ_{shar} is not properly treated by the method, it can influence the HTC estimation. The protocol of the sampling approach should be developed to estimate or limit these heat flows to ensure the good quality of the HTC results.

In this work, the estimation of ϕ_{shar} (equation 4), is based on the shared walls surfaces (A_{adj}) and U-values (U_{shar_m}) and the temperature on both sides of the walls (T_{adj} and T_i). The indices m and l stand respectively for the different shared walls in the tested area and the parts of a wall with different thermal properties.

$$\phi_{shar} = \sum_{m=1}^{s} \left(\sum_{l=1}^{r} U_{shar_{m_l}} A_{adj_{m_l}} \times (T_{adj_m} - T_i) \right) \tag{4}$$

2.2. SEREINE analysis method (for sampling approach)

SEREINE analysis tool is a Python software developed for the purposes of the project with the same name, from the French PROFEEL program. Although this tool is still in

development, it is similar to the ISABELE and EPILOG methods that led to encouraging measurement results for the HTC of newly built SFH (Thébault. & Bouchié 2018; Boisson & Bouchié 2014; de Carvalho Araujo 2018). In this section, an analysis method using SEREINE tool will be described specifically for the HTC measurement on MFH with the sampling approach. The following subsections describe the overall analysis method.

2.2.1. Thermal models

In this study, two first-order lumped RC thermal models were used for calibration, respectively *ti* and *tw*. Figure 3 presents these models, where the parameters in red are estimated by an optimisation algorithm, the parameters in blue are measured and the parameters in yellow are estimated through other models.

In both models, the output variable is the indoor air temperature (Ti). The equivalent outdoor temperature (T_{eq}) represents the data collected by special sensors developed in Isabele method with a white and a black plate, called SENS sensors. The resistances are modelled as equivalent transmittances, H_{tr}, H_e and H_i represent the thermal losses through the envelope. C_i and C_w represent the lumped heat capacity of the air and of the thermal mass and P_h stands for the heat power delivered during the test.

Heat losses due to infiltrations (P_{inf}) and their uncertainties can also be estimated by means of a simplified aeraulic model from the EN 16798-7 (method 1) (Standard 2017) that uses blowerdoor and windspeed measurements, main building dimensions and outdoor pressure coefficients to calculate the air flow rate (Thébault & Bouchié 2015; Thebault & Millet 2016).

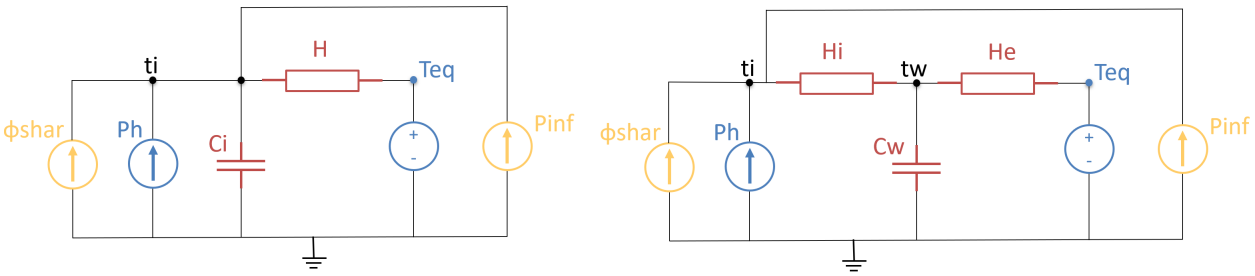


Figure 3: RC models used for calibration : *ti* (left) and *tw* (right)

To take into account the heat transfer through shared walls, which is not part of the measured perimeter (external envelope), a third heat flux is added to the internal space: ϕ_{shar}. This heat flow should be estimated to avoid impacting the HTC estimate, which refers only to thermal losses through the envelope.

2.2.2. Optimisation algorithm

Similar to the refined ISABELE method (Thébault. & Bouchié 2018), the estimated capacitances and resistances (and then, the transmission heat loss coefficient) are calculated employing maximum a posteriori probability (MAP) estimate using a bayesian approach with stochastic state space modelling of the RC models (performed by pySIP algorithm (Raillon et al. 2019)). In this process, model estimations of T_i are fitted to the measured variable. The likelihood ratio test then performs the model selection between models *ti* and *tw*.

2.2.3. Uncertainty quantification

To consider all sources of uncertainty (input data, output data and model) with both random and systematic components, a hybrid method was developed by combining the pySIP toolbox with a quasi-Monte Carlo method. The corresponding procedure, illustrated in Figure 4, is detailed:

1. Systematic uncertainties are defined, with corresponding probability density functions;
2. Multiple biases are sampled from these and added to the measurement data (using Sobol sampling);
3. Each sample is processed by the pySIP tool, calculating the fitted parameters covariance matrix by MAP estimates. Gaussian HTC estimates are calculated from the covariance matrix;
4. All Gaussian estimates are summed, forming a new probability density function;
5. A 95 % confidence interval is calculated.

A convergence criterion based on the overall variance variation was set to stop the propagation.

The probability densities for systematic uncertainties are considered to be Gaussian in this study, the bounds represent 1.96 times the standard deviation. Two important uncertainty sources were considered for the estimation of ϕ_{shar}:

- A potential offset to the measured adjacent temperatures behind the shared walls T_{adj} (set to ± 0.5 K).
- A bias on the estimated U-values on shared walls U_{shar} (set to ± 20 %, considering that the composition of shared walls is nearly known).

The other uncertainties were applied as usual in the SEREINE method. Errors due to the estimation of outdoor equivalent temperature T_{eq} were grouped into a global systematic uncertainty of ± 0.5 K. Errors due to the fixed height of measurement and insufficient air mixing of the indoor environment, associated with T_i, were of ± 0.5 K. Errors in the measurements of P_h were considered as ± 2 %.

2.3. Quality indicators

Virtual models provide a reference value for HTC (HTC_{ref}), enabling the use of objective criteria to analyse the quality of the method outputs. The bias (equation 5) represents the relative difference of the estimated HTC ($\widehat{HTC}$) to the reference. The closer to zero, the more accurate result.

$$bias = \frac{HTC_{ref} - \widehat{HTC}}{HTC_{ref}} \tag{5}$$

Besides a low bias value, the estimation output should present a limited level of uncertainty to have more precise results. The uncertainty of the indicator is represented as the percentage difference between the middle point and the borders of the result 95 % confidence interval (I95). Different from the bias, the uncertainty does not take into account the reference value, it is only relative to the estimated indicator and the dispersion of its probability

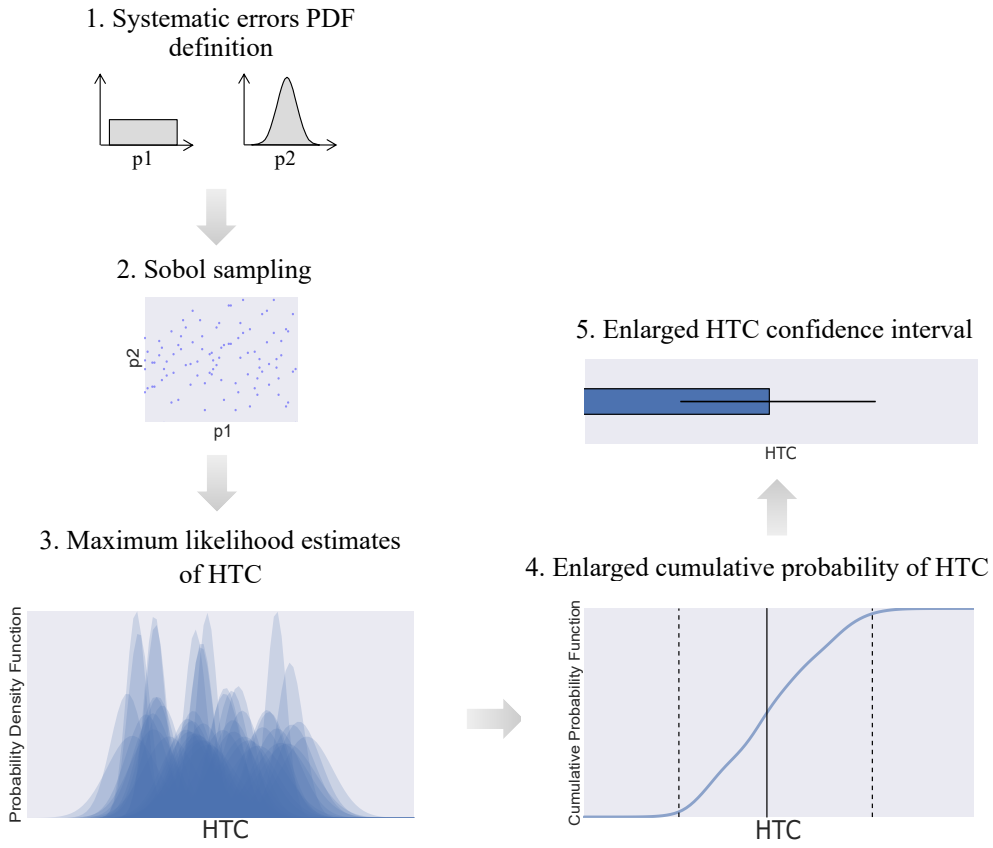


Figure 4: Illustration of the uncertainty propagation procedure, adapted from (Thébault. & Bouchié 2018).

density. A threshold of 15 % in the bias and 35 % in the uncertainty were considered for defining acceptable results.

The combination of these thresholds has the advantage to separate the results into two groups, acceptable and unacceptable, and facilitate its analysis. However, this classification does not give information on the level of the quality of the results. To better take into account the nuances of the results, interpretability (Juricic et al. 2021) was also used as a quality indicator. This represents the intersection of an interval centred in the reference value and the probability distribution of the estimated result. The interpretability is mathematically expressed in equation 6, where *f(x)* is the result probability density function of HTC and i is the interval range that is been analysed. The interpretability can assume continuous values between 0 and 1, since the maximum area under the probability density function is equal to 1.

$$interpretability = \int_{HTC_{ref}(1-i)}^{HTC_{ref}(1+i)} f(x)\,dx \tag{6}$$

Figure 5 shows a graphic representation of the interpretability indicator for a ± 15 % interval of two different tests applied in the same building. The interpretability for each is equal to the value of "b" minus "a" in their cumulative distribution function of the result.

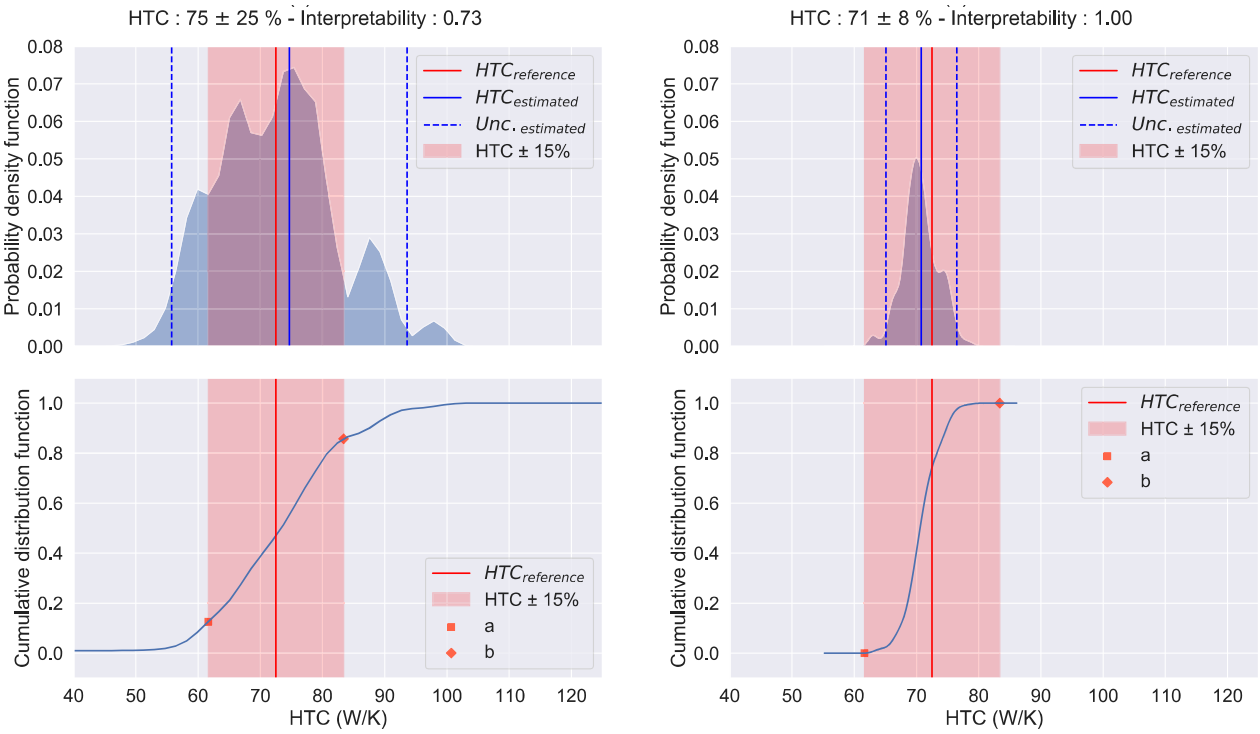


Figure 5: Representation of the interpretability indicator with the cumulative distribution function and the probability density function of the two test results applied in the same building.

It can be seen that both results are acceptable, but the result on the right has better quality with a narrower uncertainty associated with the probability density function. This behaviour is reflected in the interpretability of the results, which associate better quality results with high values with a maximum of 1.

The interpretability indicator has the asset of showing the nuances of quality in the results through a continuous variable. However, when dealing with multiple results, it might be useful to choose a threshold value for this variable to understand better the method behaviour under different conditions. The combination of both criteria can be applied using first the results from the acceptability to define the group of acceptable results. Finally, a minimum value of interpretability could be perceived in this group to represent the threshold for this continuous variable.

3. Description of the case study

Dynamic building simulations were used to apply variations of the sampling method in a virtual building to study the method's validity and limitations. An advantage of using a virtual environment is the possibility of trying the method in different configurations in a more cost and time-effective manner than for real buildings. When adapting a method to a new domain, we should resist drawing conclusions from one single simulation. Although a single simulation can provide a starting point for comparison, it seldom allows a full comprehension of a system's natural tendencies. Model experimentation is preferably done in an exploratory way by trying different simulations to help to understand the trends of a system (Ford 2009).

By comparing multiple simulations, the method limits can be studied, and insights can be given on its behaviour with a specific model. For this reason, multiple experimental plans were performed in an exploratory manner, to see the behaviour of the method when the protocol or the estimation process is modified. The experimental plans present variations in the protocol temperatures and the weather conditions. The idea is both to develop the method and to study its validity domain.

3.1. Simulation software

Software for dynamic building simulation are white-box modelling tools, where very detailed information on building physical properties are combined with descriptive equations to simulate the building's thermal behaviour. Different software are available for these goals and there are recognised methodologies, such as BESTEST - which is based on ASHRAE standards (Berrabah et al. 2021; Soubdhan et al. 2000) - and others (Østergaard Jensen 1995), to validate their capacity to accurately simulate reality.

The virtual experiments performed in this study were simulated with the software PLEIADES and COMFIE. They were created at the beginning of the 90s' (Peuportier & Sommereux 1990) in the present ETB group (Eco-design and thermics of buildings) of the CES (Center for Energy Efficiency of Systems) part of the Mines ParisTech research center (ARMINES 2021). PLEIADES is the software used for creating digital models and COMFIE is used for dynamic thermal simulations. In this work, we call the ensemble of both as "P+C".

This software group has been maintained by the editor IZUBA Energies (IZUBA 2020) and refined through different initiatives (Recht et al. 2018) over the years. It is broadly used in the French BEP context and have been already verified and validated (Salomon et al. 2005; Brun et al. 2009; Spitz 2012). An empirical validation process has also been applied to low-energy buildings, which are more sensible to minor heat gains (Munaretto et al. 2018). It is considered a reliable tool for simulating the SEREINE method application in a renovated building model.

3.2. Model

A PLEIADES model of a mid-rise MFH with 16 apartments was studied in this work. This model was based on a real four-storey building from a French city close to Lyon. The original building was built in 1978 and has undergone a massive retrofit in 2012. The energy consumption was initially classified from the French energy performance certificate as class E (231 to 330 kWh/m^2/year); after retrofitting it became class B (51 to 90 kWh/m^2/year) (ALEC 2020).

Regarding the PLEIADES model, its structure and the type of materials were kept the same as the original one. However, the level and position of the insulation and the thermal bridge values were modified to correspond to a high-performance building envelope in the present standards for renovation. The building model has 16 floors, with a total of 1048 m^2 of living area and 1508 m^2 of envelope area.

For the components of the building, the material properties (such as λ, C and ρ) proposed by the standard library of P+C were used. The wall compositions are based on references for high-performance retrofitted buildings (ADEME 2018). In the beginning, the values of the thermal bridges were kept from the original model, with optimistic values that could be thought of for a new building. Later, the prescriptions from the French thermal regulation

for existing buildings (CSTB 2007) were used to define these values, according to each type of insulation, thickness and thermal resistance of components. The details on the building components' thermal properties are presented in Appendix A.

3.2.1. Tested apartment

The virtual tests are carried out on a corner apartment on the top floor (apartment 15 in figure 6), with two exposures, east and south. In this location, the proportion between exterior and shared walls is equal to one, the maximum normally reached when testing a single apartment in a building. This proportion plays an important role since the better test conditions are related to a high ratio between the heat flow towards the exterior and that towards the neighbours.

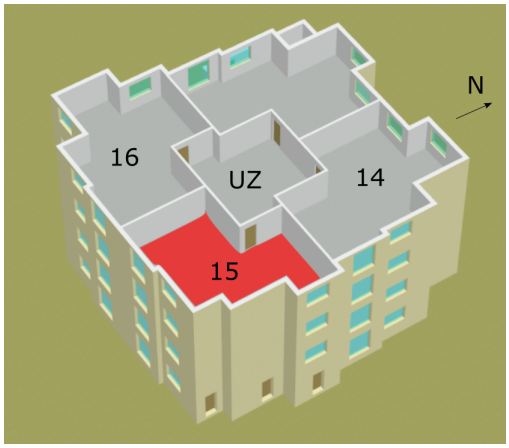


Figure 6: Internal view of the building model top floor, with floor of apartment n° 15 in red.

The apartment has a 63 m² floor area, 107 m² of exterior walls and an equivalent area of shared walls. The shared walls give to four different areas in the building: one unheated zone (UZ) located in the centre, two apartments in the same floor (14 and 16) and one apartment on the second floor (11). Equation 7 describes the heat flow passing through shared walls in an instant t. It should be noticed that the ϕ_{shar} is considered as a steady state heat flow.

$$\phi_{shar;t} = U_{shar,hor} S_{11}(T_{15;t} - T_{11;t}) + U_{shar,vert}\left(S_{14}(T_{15;t} - T_{14;t}) + S_{16}(T_{15;t} - T_{16;t}) + S_{UZ}(T_{15;t} - T_{UZ;t})\right) \quad (7)$$

3.2.2. Reference indicators

A very important asset of virtual method applications is the possibility to have a known value for the final indicator. In a virtual test, the environment can simulate both real and controlled conditions. The combination of a controlled heating scenario with a stable outdoor conditions allows the building reaching a thermal steady state, and therefore enables the calculation of a reference value for HTC and HLC.

An infiltration model has already been developed in the SEREINE method (Thébault & Bouchié 2015; Thebault & Millet 2016) and this paper is focused on developing a sampling approach, where ϕ_{shar} brings the major challenges. Since the use of a realistic infiltration

scenario would not add extra information about the influence of ϕ_{shar} on the method, this was simplified and P_{inf} was neglected, therefore HTC and HLC are equivalent. For this reason, only the HTC indicator is further analysed in this work.

In order to calculate the reference value of HTC, HTC_{ref}, the outdoor environment and the adjacent spaces were set at the same constant temperature and solar radiation was set equal to zero. The ratio between the power delivered and the temperature difference gives the HTC value. The HTC_{ref} of the part of the envelope associated with this apartment is of *72.5* $W.K^{-1}$. This value is used as a reference throughout this work.

Having an absolute HTC_{ref} is only feasible with building models since it is impossible to control the weather conditions to achieve steady-state in a real building, which brings the main advantage of developing a method with virtual protocol applications. This enables the verification of the outcomes from the different tested strategies and gives a powerful tool for quality control.

3.3. Dynamic energy simulations

All the simulations were performed using conventional weather files from the French Thermal Regulation 2012 (RT2012), corresponding to French cities such as Trappes, Nancy, La Rochelle and Nice. Outdoor equivalent temperature (T_{eq}) was calculated to make the simulation closer to the methods' *in situ* applications. The calculation of T_{eq} was simplified by reducing the term related to the form factor, as in equation 8.

$$T_{eq} = T_{out} + \frac{\alpha_s \times I_s}{h_c + h_r} \tag{8}$$

The absorptivity coefficient of the exterior wall α_s is set at 0.6, the same as the PLEIADES model exterior walls. The solar irradiance (I_s) was calculated for each direction and inclination of exterior walls and roof. In COMFIE algorithms documentation manual (IZUBA 2014), radiant and convective heat transfer coefficients (h_c and h_r) are combined into a global coefficient that depends on the surface orientation, its emissivities and wind exposure. The values of these coefficients were based on the COMFIE manual and can be found in Appendix A.

4. Results and Discussion

The sampling method was applied to the tested apartment with variations in the simulation parameters. The experimental plan, presented in part 4.1, intended to improve the estimation method through the variation in the protocol temperatures. The second study, in part 4.2, was used to verify the performance of the method under different weather conditions, throughout the year and for different cities in France.

4.1. Effect of test temperature and preheating

An experimental plan was built to investigate the advantages of preheating the apartment before the protocol application. A two-week fixed-temperature test period was used before testing the variants with preheating. Even when no preheating was applied, the apartment presented an initial test temperature of 15.8 °C, which was due to heat exchanges with the

neighbours that were heated throughout the year. The temperature difference between the beginning and the end of the test ($\Delta\ T_{beg-end}$) was also studied, as it could be an influential parameter on the quality of the test results. For this reason, different setpoint temperatures were used applying the same $\Delta\ T_{beg-end}$ of 0 and 4 Kelvins. Its values, as well as those to which the tested housing is heated during the protocol, are given in table 1.

Table 1: Preheating, setpoint temperatures and calculated $\Delta\ T_{beg-end}$ of the experimental plan.

Variant number	Preheating temperature [°C]	Setpoint temperature [°C]	$\Delta\ T_{beg-end}$ [K]
1	no preheating		2.2
2	14	18	-14
3	18		-18
4	no preheating		6.2
5	18	22	4
6	22		0
7	no preheating		10.2
8	22	26	4
9	26		0
10	no preheating		14.2
11	26	30	4
12	30		0

All these tests were performed under winter conditions, with the previously cited weather station of Nancy starting on the 15th of January. The temperature of the adjacent spaces was variable over the day, considering that the adjacent apartments were occupied during the test and that their temperatures could not be controlled. The scenario chosen to represent the temperature in the apartments is cyclical over a period of one day. Figure 7 shows the evolution of the setpoint temperature of the apartments that are adjacent to the tested area (number 11, 14 and 16) and of all other apartments in the building that are not adjacent to the tested area. In this scenario, one of the neighbours keeps a constant temperature over time and two of them vary the setpoint temperature between night and day. The unheated zone in the middle does not have any temperature setting. However, it is heated by the surrounding apartments.

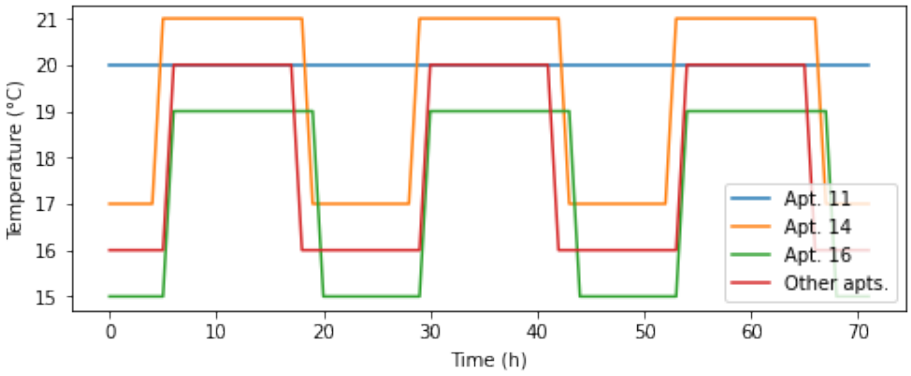


Figure 7: Setpoint temperatures for the scenario with variation in the neighbours' temperature

The equivalent outdoor temperature takes into account the solar gains per building facade. Figure 8 presents the solar radiation for each direction (G_S, G_E, G_W, G_N), the outdoor temperature (Text) and its equivalent temperatures per facade (Teq_S, Teq_E, Teq_W, Teq_N) over the test duration.

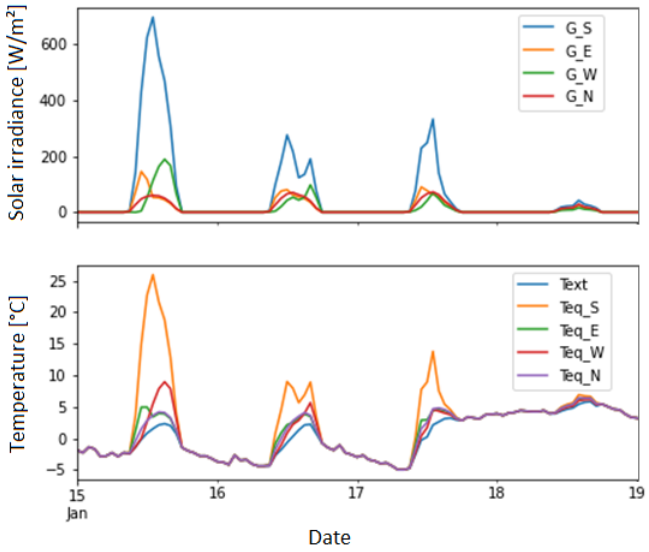


Figure 8: Solar radiation per facade (top) and outdoor and equivalent outdoor temperatures (down) during the virtual protocol application of the second experimental plan.

The twelve variants were analysed with SEREINE algorithms using a test duration from half a day to four days. The results were considered acceptable if they presented a bias inferior to 15 % and an uncertainty inferior to 35 %. Table 2 shows the statistics on the interpretability indicator, inner 15 % interval, for the results considered acceptable and unacceptable. It can be seen that just one-third of variants were acceptable with a very short duration of half day. For this short duration also, the variants diverged, while for longer durations all tests converged. With the increment of test duration, more results became acceptable and from three days all twelve variants were acceptable. Regarding the interpretability of the results, 0.5 is the maximum interpretability value presented among the unacceptable results. Even

though an interpretability of 0.4 can be found among the acceptable results, the value of 0.5 is going to be used as a threshold of result quality in this study.

Table 2: Statistical description of the interpretability of the results according to acceptability and test duration.

Interpretability \ Acceptability	Acceptable results						Unacceptable results			
statistical \ Duration (days)	0.5	1	1.5	2	3	4	0.5	1	1.5	2
description \ Amount	4	9	10	10	12	12	5	3	2	2
Mean	0.8	0.7	0.7	0.8	0.8	0.8	0.1	0.2	0.4	0.5
Standard deviation	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Minimum value	0.8	0.5	0.4	0.5	0.6	0.7	0.0	0.1	0.3	0.4
25th percentile	0.8	0.5	0.7	0.7	0.7	0.7	0.2	0.1	0.3	0.4
50th percentile	0.8	0.6	0.8	0.8	0.8	0.8	0.2	0.2	0.4	0.5
75th percentile	0.8	0.7	0.8	0.8	0.9	0.9	0.2	0.2	0.4	0.5
Maximum value	0.8	0.9	0.9	0.9	0.9	0.9	0.2	0.2	0.5	0.5

Figure 9 presents the interpretability of the 12 variants according to their setpoint temperature and $\Delta\ T_{beg-end}$. Only the diagonal of the area without preheating was tested. The areas in green are those not tested and those in grey did not converge. The blue areas correspond to the tests with acceptable results and the red areas to those with unacceptable results. The areas with light colour are tests with interpretability close to 0.5. The higher the interpretability of a test the better is its result.

For very short durations, 0.5 to 1 day most results were already acceptable if no $\Delta\ T_{beg-end}$ was applied. It indicates that preheating the building at the same temperature as the test can be a good strategy. This is mainly coherent during winter conditions since the test setpoint temperature can be close to that considered comfortable for occupants.

However, for 0.5 days, even without $\Delta\ T_{beg-end}$, the variant with 30 °C test temperature diverged. This test temperature is also the one furthest from the adjacent spaces' temperatures (T_{adj}). Considering the heated zones, T_{adj} is on average 17 °C during the night and 20 °C during the day, with a global mean of 18.8 °C. Although the variants with the highest setpoint temperature present an important signal towards the outside, they are also those with higher φ_{shar} and the test conditions are therefore less optimal.

An unexpected result is presented for the test duration of 2 days: the extremes of setpoint temperature had better results. Maybe the contribution of the increased heat flow towards the exterior compensated the effect of the φ_{shar} estimation errors. From 3 days, all results were acceptable and those with zero $\Delta\ T_{beg-end}$ and low $\Delta\ T_{int-adj}$ presented high interpretability values. The results for 4 days are not presented, since it has a similar behaviour to that of 3 days test duration.

In general, preheating the building at the same temperature as the test seems to be a good strategy to decrease the vacancy period of the test application, considering the place can be heated for one week before the test beginning with the local heating system. Under this condition and with test temperatures close to the T_{adj}, the method performed well even for very short durations, such as half a day. These conclusions are drawn for winter conditions and may be different when the outside temperature is higher.

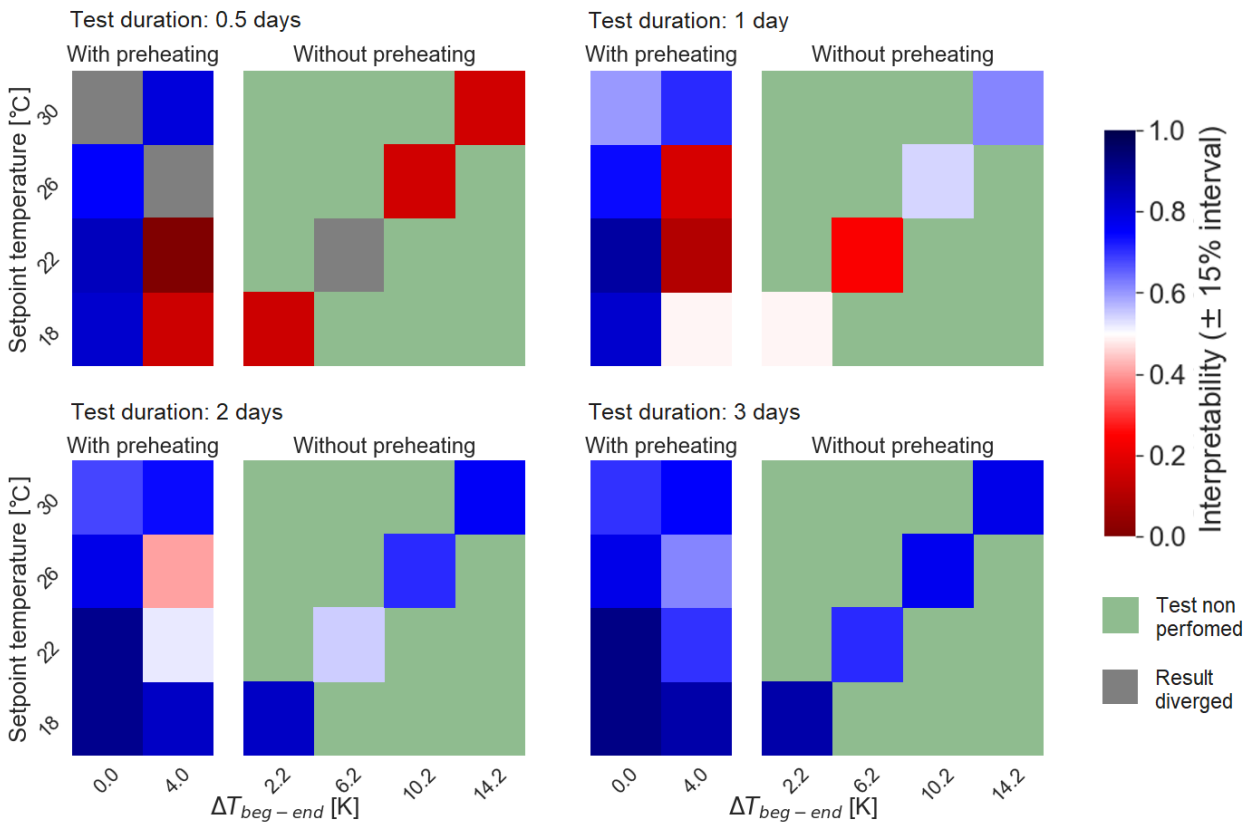


Figure 9: Interpretability of the results for different protocol setpoint temperatures, preheating conditions and test duration.

4.2. Effect of the weather

Three main experimental plans were developed, named A, B and C, to study the method behaviour over the year and at different locations in France. They were carried out in four French cities located in different climate zones: La Rochelle, Nancy, Nice and Trappes. In each experimental plan, one test was conducted per city and per month of the year. All tests started at 8 p.m on the 15th day of the month tested.

All tests followed a 2-week preheating period with a temperature as the setpoint temperature. In this way, a zero $\Delta\ T_{beg-end}$ was used in all these tests, since it related to the most favourable test condition, observed in subsection 4.1. The incident solar radiation on each of the building facades is extracted from the software and used in the SEREINE estimation algorithm. A time step of five minutes was used for the estimation process of all tests.

The experimental plan A was performed initially, to represent a protocol where the neighbours do not have their temperature controlled, bringing fewer constraints in the method application. The study presented good results for the winter period, however, tests applied during the mid-season and summer presented low-quality results. As an attempt to improve the estimation for other periods than winter time, a more invasive protocol was tried in experimental plan B. In this case, the adjacent spaces are heated at the same temperature as the tested area, except for the unheated zone. This condition is expected to improve the results since it decreases the φ_{shar} level, however, a level of heat flow towards the unheated

zone still exists.

A second attempt to improve the method behaviour during warmer seasons was to allow higher values of indoor temperature. The experimental plan C was thus performed with variable temperatures of the neighbours, but with indoor temperatures limited to a higher temperature of 35 °C. Also, a higher temperature difference with the exterior ($\Delta\ T_{int-ext}$) was used to determine the test temperature: instead of adding 10 K to the monthly mean outdoor temperature, a $\Delta\ T_{int-ext}$ of 15 K was used, based on the mean outdoor temperature of the four days following the beginning of the test.

The exact criteria used for defining the indoor temperature (T_{int}) and the adjacent space's temperature (T_{adj}) in each of the three experimental plans are presented below. The difference between these two temperatures is noted $\Delta\ T_{int-adj}$. Figure 10 shows the temperatures used for each test applied in all three experimental plans.

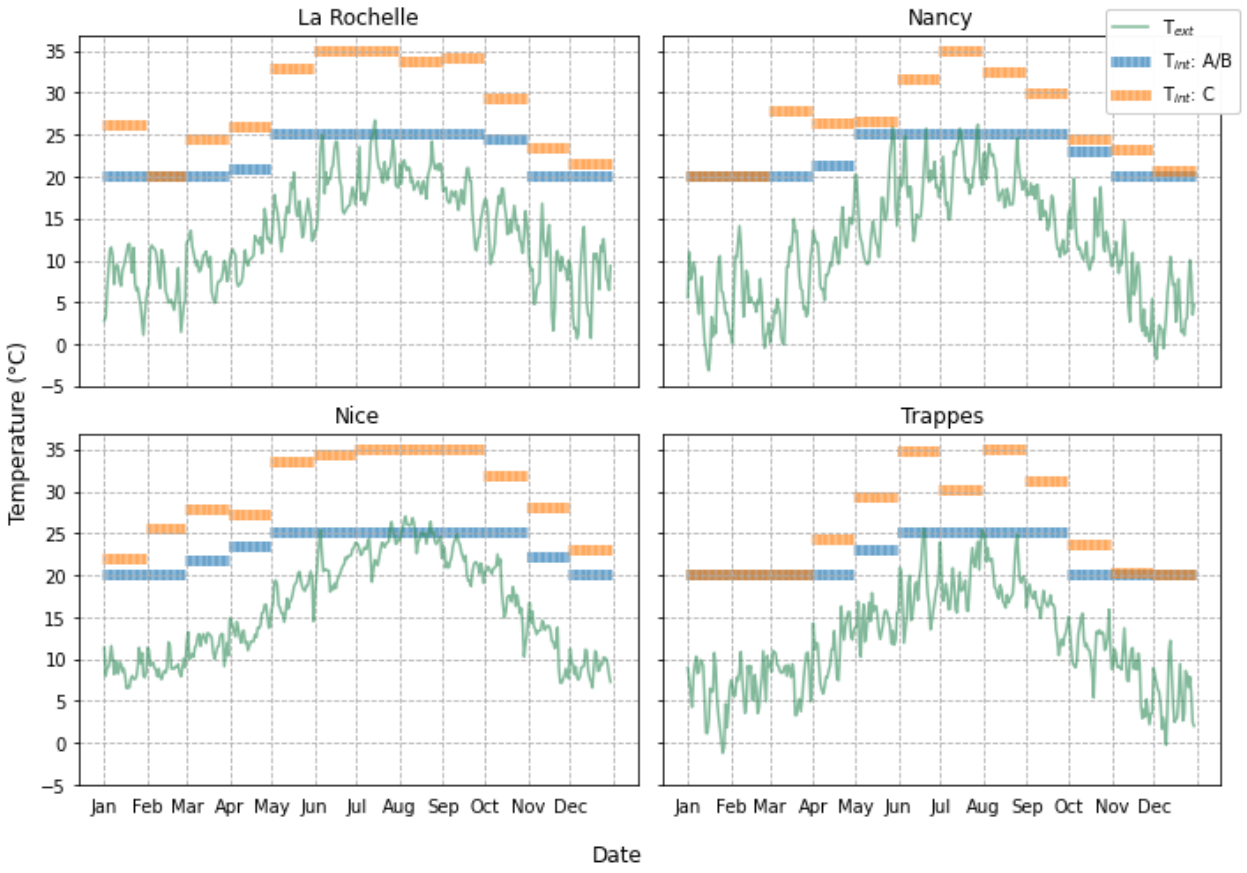


Figure 10: Tint for experimental plans A, B, C for the four different weather.

- Experimental plan A
  - $T_{int} = T_{ext_{mean_{month}}} + 10$ K, within the interval of 20 °C and 25 °C;
  - T_{adj} follows a variation pattern, different among neighbours and between day and night (figure 7).

- Experimental plan B
  - T_{int} is the same as experimental plan A;
  - $T_{adj} = T_{int}$.
- Experimental plan C
  - $T_{int} = T_{ext_{mean_{4days}}} + 15$ K, within the interval of 20 °C and 35 °C;
  - T_{adj} is variable, same as experimental plan A (figure 7).

The model selection process from the SEREINE algorithm was initially used, in which the model noted *ti* was more commonly chosen. However, the model selection process was still in development in the SEREINE project, while these analyses were conducted. So finally, both models, *ti* and *tw*, were applied to each case, disregarding the model selection process. The model *tw* presented more convergent results and, among them, there is a higher percentage of results that finished the uncertainty propagation process. The results for a same test also presented a higher interpretability value compared to results using *ti* model, with an average of 11 % higher values. In addition, it has a larger number of acceptable results for 2 and 3 days duration, with 66 acceptable results against 26 of model *ti*. For this reason, the following results presented here are solely for the estimation process with the model *tw*.

In experimental plan A, 10 % of the tests diverged while, in experimental plan B, this level decreased to 4 %. For experimental plans C, all the tests converged. Although a test might converge, it does not imply that the uncertainty propagation process has been correctly performed. Whenever a test does not complete this process, the result is not going to be further analysed. This choice was taken to avoid misinterpretation by comparing results with different levels of information. The first has a higher level of uncertainty since they have passed through the uncertainty propagation process. The results that did not finish this process have an uncertainty close to a single fit in the optimisation process, which is much smaller.

Once a result converged and went through the uncertainty propagation step, it was studied regarding the acceptability and interpretability criteria. Most of the results presented a stabilisation of the values from 2 days of test duration, some of them presented stable behaviour earlier with 1 day of test duration. To verify this, almost half of the tests in the experimental plans were conducted for up to one week of duration. Figure 11 shows the mean interpretability value for these tests with total duration of 7 days. It can be seen that from half of a day to one day there is an important improvement in the interpretability values, among all groups. From one to two days the interpretability still shows a significant increase. After that, although some improvements can be reached, they are probably not enough to justify an extra day of protocol application *in situ*.

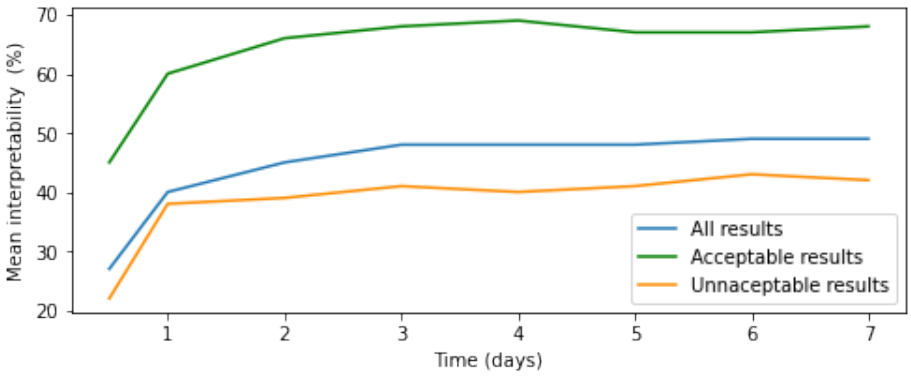


Figure 11: Mean interpretability for tests with maximum duration of 7 days by results acceptability

It should be considered that this duration depends on the context. In this study, the building has internal wall insulation and the protocol requires preheating the tested area. The duration could be longer for cases with external wall insulation and without preheating. An example of a test with acceptable results is shown in figure 12.

The whole experimental plan was conducted with at least three days of test duration since it seems to be enough to stabilise the test results. A total of 144 results is expected per test duration in the ensemble of the three experimental plans. From these, only 25 % of the tests converged and had the uncertainty propagation process completed when the protocol was applied for half a day. This value increases to 70 % for a test duration of one day, and to 85 % for a test duration of two and three days. Considering this, a really short duration of half a day should be avoided for a test application when similar experimental conditions are encountered *in situ*. One day of test duration can be considered, but the ideal would be to test a building for at least two days.

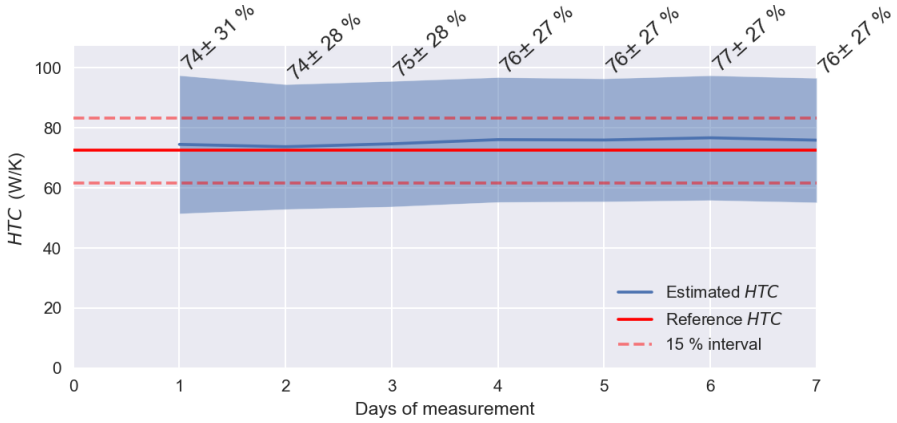


Figure 12: Results progression for the numerical test performed in February in La Rochelle with experimental plan B.

Figure 13 presents the results of interpretability for the tests that converged in the experimental plans A, B and C for a two days duration. The empty values correspond to tests that did not converge or complete the uncertainty propagation process, for reasons which would require further investigation. The values without highlight failed to reach the criteria

of bias inferior to 15 % and uncertainty inferior to 35 % and those in purple met both criteria. The values in blue met only the bias criterion, while those in red met only the uncertainty criterion.

City	La Rochelle			Nancy			Nice			Trappes		
Month\Experimental plan	A	B	C	A	B	C	A	B	C	A	B	C
January	52%		50%	85%		85%	60%	63%	60%	69%	76%	69%
February	68%	65%	68%	66%	61%	66%	37%	42%	46%		68%	67%
March	46%	48%	50%	45%	44%	45%	33%	43%	44%	56%	53%	56%
April	47%	51%	43%	41%	37%	42%	42%	50%	44%	38%	38%	42%
May	18%	25%	36%	43%	68%	41%	23%	28%	37%	27%	27%	
June	12%	9%	27%	34%	51%		19%				33%	
July			34%	9%	19%	24%		17%		20%	23%	
August	24%	30%	36%	30%	41%	39%			29%	31%	39%	43%
September	31%	41%	40%	40%	49%	43%	14%	8%	40%	31%	36%	
October	37%	54%	41%	54%		58%	38%	44%	42%	42%	44%	50%
November	33%	30%	39%	43%	44%	51%	29%	33%	44%		52%	53%
December	76%		77%	73%	68%	73%		62%	61%	75%	74%	74%

Reached criteria: Bias | Uncertainty | Bias and uncertainty

Figure 13: Interpretability and acceptability for 2 days of duration

The maximum interpretability was 85 % with Nancy's weather in January, which is the month with the lower external temperatures during the test. The city of Nice presented the lower amount of acceptable tests, and it is the city with the highest average outdoor temperatures. The external temperatures seem to be a determinant factor in the quality of the results, which is an expected behaviour, since it affects ϕ_{ext} levels. It can be seen that applying the test during winter improves the probability of achieving higher interpretability values and having acceptable results. Although the attempts of heating the adjacent spaces and increasing the indoor temperature improved the interpretability of the results, these strategies were not significant enough to reach acceptable results out of the winter season, with few exceptions.

Another tendency observed is that it was more challenging to reach the uncertainty criterion than the bias. The bias criterion was commonly met in tests performed during the mid-season and even during some summer months. However, the uncertainty criterion was mainly met during the colder months. The increased level of uncertainty is probably due to the increased level of input uncertainty in the U-value of the shared walls. In addition, this building presents fairly good insulation on the exterior walls and no insulation on the shared walls, which can lead to an unfavourable test condition, where ϕ_{shar} has a relatively high magnitude compared to ϕ_{ext}.

In figure 13, the results are ordered by month of application, where a trend of convergence and interpretability level can be seen according to the seasons. However, the days on which the tests were performed might have a hotter or colder outdoor temperature than the tested month average. To better understand the method behaviour, this data was analysed according to the temperature differences with the outdoor environment and adjacent spaces (figures 14 and 15). Besides the colour describing the interpretability level, the shape indi-

cates the ranges of outdoor temperatures during the test. The triangular points represent cold weather conditions, in which the mean outdoor temperature is below 6 °C. The circles and stars represent respectively mild and hot weather conditions.

The uncertainty and bias of the same tests are presented in figure 14. It can be seen that a bias inferior to 15 % is often met when $\Delta T_{int-ext}$ is higher enough. At least 12 K of temperature difference would be desirable for achieving this criterion. The uncertainty criterion is more restrictive and it is mainly determining which results are acceptable. It presents a similar behaviour to the interpretability, with better results for high $\Delta\ T_{int-ext}$ values. At least 15 K of temperature difference would be desirable to let the uncertainty inside the acceptable zone in addition to low temperature differences with the outdoor environment.

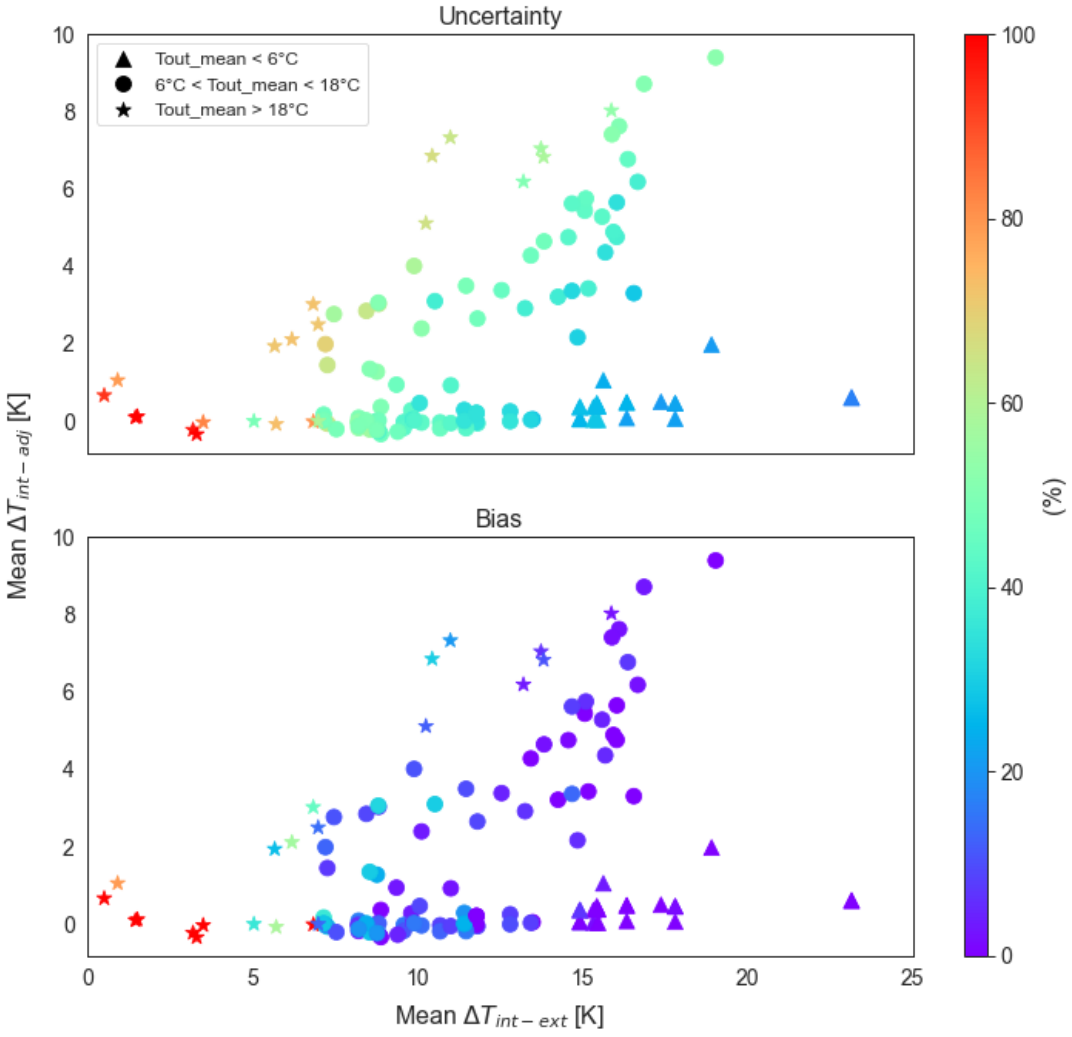


Figure 14: Bias and uncertainty of the results according to the mean $\Delta\ T_{int-ext}$ and $\Delta\ T_{int-adj}$ for 2 days of test in experimental plans A, B and C.

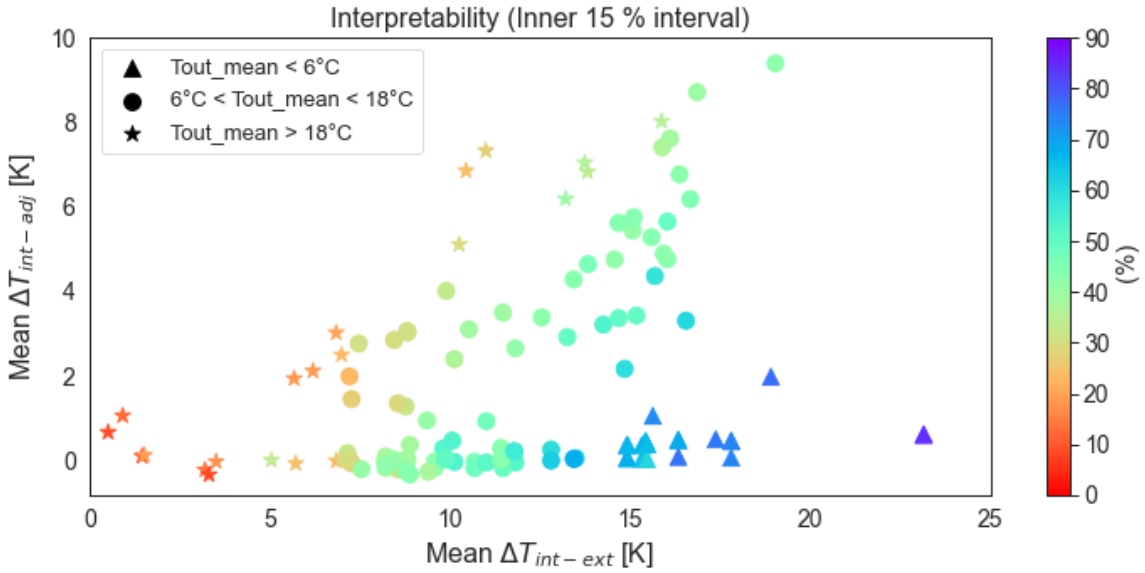


Figure 15: Interpretability of the results according to the mean $\Delta\ T_{int-ext}$ and $\Delta\ T_{int-adj}$ for 2 days of test in experimental plans A, B and C.

The group of points sharing the same colour have a general diagonal shape, showing that both differences in temperature influence the quality of the results. It can be seen that lower differences in temperature with the neighbours and higher differences with the outdoor environment provide higher levels of interpretability, which reinforces the importance of having more heat flow passing through the envelope than the shared walls.

Increasing the test temperature is a strategy to consider to improve the results, however, it can also increase the ϕ_{shar} if the adjacent spaces are not heated similarly. Since the shared walls are usually less insulated than the exterior walls, a lower temperature difference could lead to important heat losses to the adjacent spaces. The effectiveness of the increment of the test temperature thus depends on the temperature of the adjacent spaces. During the heating season, it is less relevant since an important temperature gradient with the outdoor environment can be reached with temperatures inside the thermal comfort range. If adjacent spaces are occupied, adjacent temperatures may be close to that of the tested area, leading to a low level of ϕ_{shar}.

In this case, there would be no need to control the adjacent spaces' temperatures, which avoids the application of a more invasive protocol. However, during mid-season, the T_{ext} increases and temperatures inside the thermal comfort zone might not provide an enough high $\Delta\ T_{int-ext}$ to test. To compensate for the higher external temperatures, T_{int} should be increased to improve the heat signal passing through the building envelope. However, this will possibly increase $\Delta\ T_{int-adj}$ at the same time, which is associated with less favourable testing conditions. This compensation gets critical during summer when the level of indoor temperature demanded to have an important $\Delta\ T_{int-ext}$ can exceed the safety of materials in a dwelling.

For mid-season, a possibility would be the use of both strategies together: increasing the test temperature and controlling the adjacent spaces' temperatures close to the test temperature. It can be seen in figure 15 that the bottom right zone of the graph, with higher

interpretability, is only composed of tests performed with mean outdoor temperature below 6 °C. There are no points representing the mid-season months in this area. In experimental plan B, the test temperatures were limited to 25 °C. It would be interesting to perform an experiment with higher temperatures and controlled neighbour temperatures to verify the improvement it could bring for mild temperature weather.

5. Conclusion

Numerical simulations were used to investigate a sampling approach for estimating a building HTC. The variation in indoor temperature before and during the test application showed that a protocol with preheating reaches stable results faster. Preheating the building to temperatures similar to that of the test setpoint is recommended to decrease the vacancy time for protocol application.

Forty-eight different weather conditions were used with three variations in the protocol parameters related to indoor, outdoor and neighbour temperatures. The simulations in which the heat flow delivered during the test traversed mainly the envelope, instead of the shared walls, presented more precise results. For this reason, higher temperature differences with the exterior, which were mainly reached during winter, led to less uncertain results. Decreasing the temperature differences with the neighbours, slightly improved the HTC accuracy, but did not yield acceptable results under mild and hot weather. Overall, the uncertainty criterion was the most challenging aspect for reaching acceptable HTC estimations, while the bias criterion was commonly reached disregarding weather conditions.

The proposed method is viable in winter from two days of measurement, favouring a test setpoint temperature equal to the preheating. All the simulations had a controlled and stable preheating temperature, which facilitates the short-duration HTC estimation. In real life, this condition might be difficult to reach, since the preheating would be applied by occupants with the local heating system and not the test equipment. If preheating is not applied, a longer test duration will be recommended to ensure acceptable results.

As a perspective of this work, other ways of taking into account the heat flow going towards neighbours could be considered. The use of direct measurements on the shared walls is a strategy that started to be studied in the context of the SEREINE project. Another possibility would be the integration of this parameter in the RC thermal model to be identified, imposing new boundary conditions to the system. This work could also be extended by the application of the method to buildings with different architectures and construction techniques.

Appendix

A. Building model characteristics

The global radiant and convective superficial heat transfer coefficients are used to calculate the walls' U-values and the equivalent outdoor temperatures. In P + C manual, this coefficient depends on the surface orientation, exposition and emissivity IZUBA (2014). The surface interior and exterior coefficients (hint and hext) associated with the different building models, were based on P + C manual considering a wall emissivity of 0.9. For exterior horizontal walls, a hext of 22.2 W/(m²/K) and a hint of 9.43 W/(m²/K) were considered, for vertical walls, these figures were respectively 18.2 W/(m²/K) and 8.13 W/(m²/K). In the case of shared walls, both sides of the walls are under interior conditions and the hint for horizontal walls was 8.00 W/(m²/K) and for vertical it was 8.13 W/(m²/K).

The composition of the walls and floors regarding thermal resistance are presented in table A.1. The model's windows have a U-value of 1.9 W/(m².K) and a solar factor of 50 %, while the doors have a U-value of 3.5 W/(m².K).

Table A.1: Thermal resistance of walls of the building model after retrofitting

Component	*Material (ext → int)*	*thickness [cm]*	*λ [W/(m.K)]*	*R [K.m²/W]*	*U-value [W/K.m²]*
Low floor	Concrete	16	1.600	0.10	0.314
	Expanded polystyrene	12	0.039	3.08	
Intermediate Floor	Heavy concrete	20	1.750	0.11	2.778
High floor	Heavy concrete	20	1.750	0.11	0.209
	Polyurethane	14	0.030	4.67	
Exterior wall	Heavy concrete	15	1.750	0.09	0.254
	Expanded polystyrene	15	0.039	3.85	
Shared wall	Plaster	1	0.350	0.03	2.525
	Concrete	15	1.750	0.09	
	Plaster	1	0.350	0.03	

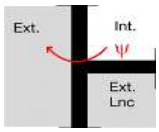
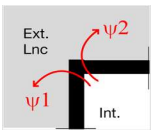
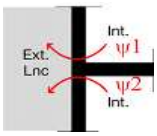
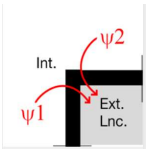
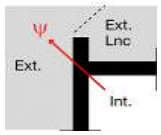
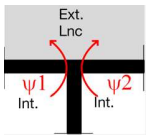
The thermal mass features of the building model are presented in table A.2.

Table A.2: Heat capacity of the building model per component material after retrofitting

Material (order) - Component	*Density kg/m3*	*Specific heat J/kg K*	*Volume m3*	*Heat capacity J/K*
Plaster (1) - Shared wall	1000	800	4.94	3.95E+06
Concrete (2) - Shared wall	2300	920	74.05	1.57E+08
Plaster (3) - Shared wall	1000	800	4.94	3.95E+06
Concrete (1) - Exterior wall	2300	920	96.72	2.05E+08
Expanded polystyrene (2) - Exterior wall	25	1380	96.72	3.34E+06
Concrete (1) - Low floor	2300	920	55.11	1.17E+08
Expanded polystyrene (2) - Low floor	25	1380	41.33	1.43E+06
Concrete (1) - High floor	2300	920	68.59	1.45E+08
Polyurethane (2) - High floor	35	837	48.01	1.41E+06
Concrete (2) - Intermediate Floor	2300	920	202.78	4.29E+08
Total				1.07E+09

The types of thermal bridges with their thermal characteristics are presented in table A.3. The graphical representation of the thermal bridges come from P+C software.

Table A.3: Thermal properties of thermal bridges of the building model after retrofitting

Wall connection	ψ[W/m.K] Retrofitted	ψ[W/m.K] New	Wall connection	ψ[W/m.K] Retrofitted	ψ[W/m.K] New
Exterior wall / low floors 	0,06	0,05	Outward angles 	0,03	0,03
Exterior wall / intermediate floors 	1,1	0,77	Inward angles 	0,03	0,03
Exterior wall / high floor 	0,97	0,05	Shared wall / high floor 	0,95	0,68

References

ADEME (2013). Chiffres clés. https://www.ademe.fr/sites/default/files/assets/documents/chiffres-cles-batiment-edition-2013-8123.pdf. Online; accessed on 2021-10-12.

ADEME (2018). *Enquête TREMI – Travaux de Rénovation Energétique des Maisons Individuelles - Campagne 2017.* Technical report, ADEME.

ALEC (2020). Résidence vignette figuière - FEYZIN. https://carto.infoenergie69-grandlyon.org/projet/residence-vignette-figuiere/. Online; accessed on 2021-10-27.

Alzetto, F., Meulemans, J., Pandraud, G., & Roux, D. (2018a). A perturbation method to estimate building thermal performance. *Comptes Rendus Chimie*, 21(10), 938–942. The power of synthesis towards new materials. Joint Symposium of the National Academy of Sciences Leopoldina and the "Académie des Sciences", Paris Halle (Saale), Germany, 23 & 24 November 2017.

Alzetto, F., Pandraud, G., Fitton, R., Heusler, I., & Sinnesbichler, H. (2018b). QUB: A fast dynamic method for in-situ measurement of the whole building heat loss. *Energy and Buildings*, 174, 124–133.

ARMINES (2021). Les centres de recherche énergétique et génie des procédés. https://armines.net/fr/centres-de-recherche-departements-thematiques/%C3%A9nerg%C3%A9tique-et-g%C3%A9nie-des-proc%C3%A9d%C3%A9s. Online; accessed on 2021-10-28.

Bacher, P. & Madsen, H. (2011). Identifying suitable models for the heat dynamics of buildings. *Energy and Buildings*, 43(7), 1511–1522.

Bauwens, G., Ritosa, K., & Roels, S. (2021). *Annex 71 final report - Building energy performance assessment based on in-situ measurements: Physical parameter identification.* Technical report, IEA-EBC.

Bauwens, G. & Roels, S. (2014). Co-heating test: A state-of-the-art. *Energy and Buildings*, 82, 163–172.

Berrabah, S., Bouhssine, Z., El Maakoul, A., Degiovanni, A., & Bakhouya, M. (2021). Towards a quadrupole-based method for buildings simulation: Validation with ASHRAE 140 standard. *Thermal Science and Engineering Progress*, (pp. 101069).

Boisson, P. & Bouchié, R. (2014). ISABELE method: In-Situ Assessment of the Building EnveLope pErformances. In *9th International Conference on System Simulation in Buildings* (pp.$\tilde{2}$0). Liège, Belgium.

Bouchié, R., Alzetto, F., Brun, A., Boisson, P., & Thébault, S. (2014). Short methodologies for in-situ assessment of the intrinsic thermal performance of the building envelope. In *Sustainable Places 2014* Nice.

Bouchié, R., Alzetto, F., Brun, A., weeks, c., Preece, M., Ahmad, M., & Sisinni, M. (2015). *Methodologies for the assessment of intrinsic energy performance of buildings' envelope.* Technical report, Project: PERFORMER.

Bouchié, R. & Ibos, L. (2020). *MPEB: Inventaire des méthodes applicables à la caractérisation de la performance énergétique de l'enveloppe.* Technical report, Fondation Batiment Energie.

Brun, A., Spitz, C., Wurtz, E., & Mora, L. (2009). Behavioural comparison of some predictive tools used in a low-energy building. In *Building Simulation 2009* Glasgow.

Calì, D., Osterhage, T., Streblow, R., & Müller, D. (2016). Energy performance gap in refurbished german dwellings: Lesson learned from a field test. *Energy and Buildings*, 127, 1146–1158.

CSTB (2007). *Règles Th-U – Fascicule 5/5 – Ponts Thermiques. Réglementation Thermique des Bâtiments Existants.* Technical report, CSTB.

de Carvalho Araujo, L. (2018). Study of a methodology for measuring the energy performance of building. Master's thesis, PSL Research University. France.

de Carvalho Araujo, L., Thébault, S., Mora, L., & Recht, T. (2021). Measurement of the building envelope thermal performance in collective housings. In *17th International Conference of the International Building Performance Simulation Association, BS2021* (pp. 1–8). Bruges, Belgium.

Deb, C., Gelder, L., Spiekman, M., Pandraud, G., Jack, R., & Fitton, R. (2021). Measuring the heat transfer coefficient (HTC) in buildings: A stakeholder's survey. *Renewable and Sustainable Energy Reviews*, 144, 111008.

Díaz-Hernández, Torres-Hernández, Castro, K. M., Macias-Melo, & Jiménez, M. J. (2020). Data-based RC dynamic modelling incorporating physical criteria to obtain the HLC of in-use buildings: Application to a case study. *Energies*, 13, 313.

European Comission (2018). EU buildings factsheets. https://ec.europa.eu/energy/eu-buildings-factsheets_en. Online; accessed on 2021-09-10.

Ford, A. (2009). *Modeling the Environment*. Washington D.C: Island Press. (2nd edition).

Hestenes, D. (1997). Modeling methodology for physics teachers. In *International Conference on Undergraduate Physics Education*, volume 399 College Park.

IZUBA (2014). *Cahier d'algorithmes de COMFIE*.

IZUBA (2020). Optimisation énergétique et environnementale dans le secteur du bâtiment. https://www.izuba.fr/. Online; accessed on 2021-10-28.

Jiménez, M. J. (2016). *Reliable building energy performance characterisation based on full scale dynamic measurements - Report of Subtask 3, part 1: Thermal performance characterization based on full scale testing - description of the common exercises and physical guidelines*. Technical report, IEA - EBC - Annex 58.

Johnston, D., Miles-Shenton, D., & Farmer, D. (2015). Quantifying the domestic building fabric 'performance gap'. *Building Services Engineering Research and Technology*, 36.

Juricic, S., Goffart, J., Rouchier, S., Foucquier, A., Cellier, N., & Fraisse, G. (2021). Influence of natural weather variability on the thermal characterisation of a building envelope. *Applied Energy*, 288, 116582.

Liu, Y., Yang, L., Zheng, W., Liu, T., Zhang, X., & Liu, J. (2018). A novel building energy efficiency evaluation index: Establishment of calculation model and application. *Energy Conversion and Management*, 166, 522–533.

Madsen, H., Bacher, P., Bauwens, G., Deconinck, A.-H., Reynders, G., Roels, S., Himpe, E., & Lethé, G. (2021). Annex 58 report of subtask 3, part 2: Thermal performance characterisation using time series data – statistical guidelines.

Majcen, D., Itard, L., & Visscher, H. (2013). Theoretical vs. actual energy consumption of labelled dwellings in the netherlands: Discrepancies and policy implications. *Energy Policy*, 54, 125–136. Decades of Diesel.

Mangematin, E., Pandraud, G., & Roux, D. (2012). Quick measurements of energy efficiency of buildings. *Comptes Rendus Physique*, 13, 383–390.

Munaretto, F., Recht, T., Schalbart, P., & Peuportier, B. (2018). Empirical validation of different internal superficial heat transfer models on a full-scale passive house. *Journal of Building Performance Simulation*, 11(3), 261–282.

Nordström, G., Johnsson, H., & Lidelöw, S. (2013). Using the energy signature method to estimate the effective U-Value of buildings. In A. Hakansson, M. Höjer, R. J. Howlett, & L. C. Jain (Eds.), *Sustainability in Energy and Buildings* (pp. 35–44). Berlin, Heidelberg: Springer Berlin Heidelberg.

PACTE, A. Q. C. (2017). EPILOG - evaluation de la performance intrinsque de logements. http://www.programmepacte.fr/epilog-evaluation-de-la-performance-intrinseque-de-logements. Online; accessed on 2018-09-23.

Pappalardo, M. & Reverdy, T. (2020). Explaining the performance gap in a french energy efficient building: Persistent misalignment between building design, space occupancy and operation practices. *Energy Research & Social Science*, 70, 101809.

Peuportier, B. & Sommereux, I. B. (1990). Simulation tool with its expert interface for the thermal design of multizone buildings. *International Journal of Solar Energy*, 8, 109–120.

Raillon, l., Rouchier, S., & Juricic, S. (2019). pySIP: an open-source tool for bayesian inference and prediction of heat transfer in buildings. In *congrès annuel de la Société Française de Thermique* Nantes.

Rasmussen, C. (2020). *Data-driven Methods for Reliable Energy Performance Characterisation of Occupied Buildings*. phdthesis, DTU Compute Department of Applied Mathematics and Computer Science.

Recht, T., Goffart, J., Mora, L., Woloszyn, M., & Catherine, B. (2018). Méthodologie pour la comparaison des performances simulées et mesurées de maisons « à énergie positive ». In *IBPSA France 2018* Bordeaux.

Roels, S. (2017). *EBC Annex 58 Project Summary Report*. Technical report, International Energy Agency.

Roels, S., Bacher, P., Bauwens, G., Castaño, S., Jiménez, M. J., & Madsen, H. (2017). On site characterisation of the overall heat loss coefficient: Comparison of different assessment methods by a blind validation exercise on a round robin test box. *Energy and Buildings*, 153.

Roels, S., Bacher, P., Bauwens, G., Madsen, H., & Jiménez, M. J. (2015). Characterising the actual thermal performance of buildings: Current results of common exercises performed in the framework of the IEA EBC annex 58-project. *Energy Procedia*, 78, 3282–3287. 6th International Building Physics Conference, IBPC 2015.

Rouchier, S., Jiménez, M. J., & Castaño, S. (2019). Sequential monte carlo for on-line parameter estimation of a lumped building energy model. *Energy and Buildings*, 187, 86–94.

Salomon, T., Mikolasek, R., & Peuportier, B. (2005). Outil de simulation thermique du bâtiment, comfie. *Journée thématique SFT-IBPSA, La Rochelle*.

Schetelat, P. et Bouchié, R. (2014). ISABELE : a method for performance assessment at acceptance stage using bayesian calibration. *System Simulation in Buildings*, (pp. 594–608 (P34)). 9th International Conference on System Simulation in Buildings - SSB2014.

Senave, M., Roels, S., Reynders, G., Verbeke, S., & Saelens, D. (2019a). Assessment of data analysis methods to identify the heat loss coefficient from on-board monitoring data. *Energy and Buildings*, 209, 109706.

Senave, M., Roels, S., Verbeke, S., Lambie, E., & Saelens, D. (2019b). Sensitivity of characterizing the heat loss coefficient through on-board monitoring: A case study analysis. *Energies*, 12, 3322.

Soubdhan, T., Mara, T., Boyer, H., & Younes, A. (2000). Chapter 376 - use of BESTEST procedure to improve a building thermal simulation program. In A. Sayigh (Ed.), *World Renewable Energy Congress VI* (pp. 1800–1803). Oxford: Pergamon.

Spitz, C. (2012). *Analyse de la fiabilité des outils de simulation et des incertitudes de métrologie appliquée à l'efficacité énergétique des bâtiments*. Phd thesis, Université de Grenoble.

Standard, E. (2017). NF EN 16798-7 : 2017 - Energy performance of buildings - Ventilation for buildings - Part 7 : calculation methods for the determination of air flow rates in buildings including infiltration (Modules M5-5).

Thébault, S. & Bouchié, R. (2015). Estimating infiltration losses for in-situ measurements of the building envelope thermal performance. *Energy Procedia*, 78, 1756–1761.

Thébault., S. & Bouchié, R. (2018). Refinement of the ISABELE method regarding uncertainty quantification and thermal dynamics modelling. *Energy and Buildings*, 178, 182–205.

Thebault, S. & Millet, J.-R. (2016). Cost-effective air flow rate estimations using blowerdoor and wind speed measurements to assess building envelope thermal performances. *J. Build. Phys.*

Thébault, S. R. (2017). *Contribution à l'évaluation in situ des performances d'isolation thermique de l'enveloppe des bâtiments*. Phd thesis, Université de Lyon.

Uriarte, I., Erkoreka, A., Giraldo-Soto, C., Martin, K., Uriarte, A., & Eguia, P. (2019). Mathematical development of an average method for estimating the reduction of the heat loss coefficient of an energetically retrofitted occupied office building. *Energy and Buildings*, 192, 101–122.

Xu, K. & Darve, E. (2021). Solving inverse problems in stochastic models using deep neural networks and adversarial training. *Computer Methods in Applied Mechanics and Engineering*, 384, 113976.

Zayane, C. (2011). *Identification d'un modèle de comportement thermique de bâtiment à partir de sa courbe de charge*. Phd thesis, École Nationale Supérieure des Mines de Paris.

Zhang, X., Rasmussen, C., Saelens, D., & Roels, S. (2022). Time-dependent solar aperture estimation of a building: Comparing grey-box and white-box approaches. *Renewable and Sustainable Energy Reviews*, 161, 112337.

Ziour, R. & Calberg-Ellen, P. (2020). *Atelier FBE MPEB - Mesure de la Performance énergétique des Bâtiments: Les différentes configurations de MPEB*. Technical report, Fondation Batiment Energie. Online; accessed on 2021-10-22.

Zou, P. X. & Alam, M. (2020). Closing the building energy performance gap through component level analysis and stakeholder collaborations. *Energy and Buildings*, 224, 110276.

Østergaard Jensen, S. (1995). Validation of building energy simulation programs: a methodology. *Energy and Buildings*, 22(2), 133–144.