Identifiability of the thermal performance of a building envelope from poorly informative data

Sarah Juricic

▶ **To cite this version:**

Sarah Juricic. Identifiability of the thermal performance of a building envelope from poorly informative data. Thermics [physics.class-ph]. Université Savoie Mont Blanc, 2020. English. NNT : 2020CHAMA014 . tel-03181809

HAL Id: tel-03181809

https://theses.hal.science/tel-03181809v1

Submitted on 25 Mar 2021

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.

UNIVERSITÉ SAVOIE MONT BLANC

THÈSE

Pour obtenir le grade de

DOCTEURE DE L'UNIVERSITÉ SAVOIE MONT-BLANC

Spécialité : **Génie Civil et Sciences de l'Habitat**

Arrêtée ministériel : 25 mai 2016

Présentée par

Sarah JURICIC

Thèse dirigée par **Gilles FRAISSE**
et codirigée par **Simon ROUCHIER**

préparée au sein du **Laboratoire LOCIE**
dans **l'École Doctorale SISEO**

Identifiability of the thermal performance of a building envelope from poorly informative data

Thèse soutenue publiquement le **9 juillet 2020**,
devant le jury composé de :

Christian Ghiaus
INSA Lyon, Président

Patrick Salagnac
Université de La Rochelle, Rapporteur

Ruchi Choudhary
University of Cambridge, Rapporteure

Staf Roels
KU Leuven, Examinateur

Gilles Fraisse
Université Savoie Mont-Blanc, Directeur de thèse

Simon Rouchier
Université Savoie Mont-Blanc, Co-Directeur de thèse

Jeanne Goffart
Université Savoie Mont-Blanc, Co-Encadrant de thèse

Aurélie Foucquier
Université Grenoble Alpes CEA LITEN, Co-Encadrant de thèse

Peder Bacher
Technical University of Denmark, Invité

Acknowledgements

Ces premiers mots achèvent en réalité un travail de longue haleine, fruit de rencontres avec des personnes extraordinaires que je voudrais remercier ici.

En premier lieu, je voudrais remercier Simon Rouchier, Maître de conférences à l'Université Savoie Mont-Blanc, pour la confiance qu'il m'a accordée du début à la fin de ce projet, pour son soucis de ma formation doctorale et pour m'avoir aidé à garder le cap. Simon est plus qu'un encadrant brillant, il est une inspiration à bien des égards.

Je voudrais aussi remercier Gilles Fraisse, Professeur à l'Université Savoie Mont-Blanc et directeur de cette thèse, pour son aide à voir *the big picture*, pour son honnêteté et pour sa patience à encadrer une doctorante somme toute bien obstinée.

Ce travail de thèse doit aussi beaucoup à mes deux encadrantes Aurélie Foucquier et Jeanne Goffart. Je remercie Aurélie d'avoir trouvé le temps d'accompagner mes errances, de s'être rendue disponible à tout moment, et pour ses questions et ses conseils aussi pertinents qu'avisés. Jeanne, merci pour toutes les discussions, les schémas, les cafés, les conseils, les encouragements et l'amitié. Que l'aventure continue !

Ces quelques mots ne sauraient, tout compte fait, rendre justice à ce que ce travail de thèse doit à cette équipe d'encadrement. Peu d'autres doctorants je pense, auront eu la chance d'avoir un encadrement aussi pertinent, drôle, disponible et bienveillant que celui-ci.

I would also like to deeply thank Peder Bacher for the warmest welcome at DTU. I strongly appreciate how Peder took the time, every week, to provide guidance for my research. I also particularly enjoyed his enthusiasm for science... and basically for everything else !

In addition, I would like to express my gratitude to Ruchi Choudhary, Patrick Salagnac, Staf Roels and Christian Ghiaus for accepting to review this work. Their insight and comments have been highly appreciated.

I would like to adress a particular thank to Staf Roels and in general to all participants of the IEA EBC annex 71. The annex has been an unexpected and wonderful opportunity to learn, to get in touch with brilliant reseachers and to put my work in a broader perspective.

La recette d'une thèse réussie m'a dit un jour Rémi Carrié en peu ou prou ces mots, c'est un bon encadrement, un bon cadre de travail et un bon environnement personnel. Le LOCIE y a sans aucun doute beaucoup contribué.

Mille mercis donc à Nicolas Cellier, pour son aide et son amitié. Il a toujours répondu présent pour me remonter le moral, pour "faire de la science", pour coder, pour débugger... Nous sommes nombreux à lui devoir beaucoup.

Merci à Gaëlle Guyot pour sa bienveillance, ses encouragements et son témoignage de maman-chercheuse engagée et motivante. Merci à Mickaël Rabouille, merci aux co-doctorants 2016 Sanghasri et Léa, un merci particulier à Yannis, co-bureau et co-auteur de mémorables *office pranks*, mais aussi à Madina, Ainagul et tous les autres pour votre support, votre amitié et d'avoir été là. Merci à tous ceux qui font ou ont fait du LOCIE un laboratoire chaleureux, vivant et fourmillant d'idées. Enfin bien sûr merci à tous ceux qui ont oeuvrés à acheter et entretenir la machine à café du LOCIE, meilleure amie du chercheur.

Quant à "l'environnement personnel", si j'ai pu aller jusqu'au bout de ce travail, je le dois en effet à ma merveilleuse petite famille. Merci Jean pour ton soutien indéfectible, ta patience, tes encouragements et ton écoute incroyable. Merci Elias d'être venu à nous et d'éclairer de tes yeux verts et rieurs tous les autres moments de la vie.

Contents

Abstract

Accurate thermal diagnosis of buildings is a key to drive adapted retrofit strategies, necessary to reduce the energy use of the building sector. Diagnosis from in situ non intrusive measurements have recently gained a strong interest as they deliver insight on the thermal performance of the building in actual operating conditions. The most accurate protocols rely on controlled indoor conditions and demand therefore that the buildings be unoccupied. In many cases however like hospitals, offices or collective dwellings, the building cannot be left vacant. In such cases, measurements could rather be performed in non intrusive conditions, with a limited number of sensors and uncontrolled indoor conditions, which leads the data to be less informative.

Poorly informative data may either exacerbate the uncertainties of the estimation of the thermal properties or even make the estimation impossible. This thesis proposes to tackle the feasibility of thermal performance estimation from poorly informative data.

Assessing the thermal performance of a building envelope consists in numerically fitting an adequate thermal model to collected data. This work therefore first reviews existing methods and various suitable heat transfer models at building scale. In the light of non intrusive and poorly informative framework, RC models constitute a promising choice. First, they learn from the dynamic nature of the data, albeit poorly informative, and on shorter datasets than other models. Second, they are built from physical knowledge with a reasonable number of parameters. RC models can then learn from light measurement plans whilst having a physical meaning.

The thermal performance estimation itself then consists in calibrating the chosen model, i.e. fitting its parameters such that the model prediction is the most consistent with the observations. The best fitting parameters cannot be calculated analytically: the calibration becomes a numerical problem. The frequentist and Bayesian approaches for parameter estimation and their underlying assumptions are discussed with the perspective of non intrusive experiments. Finally, efficient algorithms are presented to numerically solve the problem.

Once a model fits the data, physical interpretation of its parameters, provided they are unique, does not follow immediately. Clearly, the unicity of the estimation, i.e. the identifiability, is a necessary condition to any interpretation. Identifiability is a property of a given model with a given dataset. The model itself needs to be structurally identifiable. Some tools for such a verification have been applied to form a set of theoretically identifiable RC models. At the same time, identifiability also depends on the quality of the data. A workflow is proposed and constitutes a sequence of necessary steps that should be fulfilled during any model calibration, in this work or in any calibration for the matter.

Still, physical interpretation does not immediately follow. A fundamental difference is made between identifiability and physical interpretation of the estimated values. This difference indeed exists de facto for the reason that any model has an intrinsic error. There is therefore no guarantee that unique estimates of a slightly erroneous thermal model can be identified to the actual thermal properties. For this reason, a stochastic state space representation of the RC models is chosen to at least propagate the uncertainty due to the model error and reproduce representative uncertainties of the estimation.

This thesis next deals with assessing the ability of stochastic RC models to be physically interpreted. To do so, a numerical methodology is proposed, using a computer based one storey building energy model. This reference model is simulated under known conditions and its actual thermal performance is calculated. Each simulation of the reference model serves as training set for calibrating stochastic RC models.

The numerical methodology is first applied to assess the influence of weather variability on the robustness of the estimation of thermal performance. Indeed, under non intrusive conditions, weather conditions are expected to have a larger influence on indoor temperature and heating power than in controlled experiments. The reference model is simulated multiple times, each with a different synthetic weather dataset, created as to allow sensitivity analysis. Finally, an RC model is trained on each of the synthetic datasets in order to deliver parameter estimates for each scenario.

The relationship between the reproducibility of the estimation and the duration of the experiment is explored. It is indeed expected that below a certain duration, the parameter estimation is not reproducible under different weather conditions, i.e. the calibration would yield a different outcome. Results indicate that for the case study, a minimum of 11 days data is necessary in a non intrusive experiment to achieve reproducibility of the estimation.

At the same time, it is found that estimations are strongly influenced by outdoor temperature and wind speed. Large infiltration rates in the computer model explain this outcome. The overall variability is significantly dampened with longer datasets: it decreases by a factor 4 with 11 days training. Yet, the influence of outdoor temperature and wind speed on the remaining variability keeps significant. Stochastic RC models may be considered appropriate for reproducible experiments but, for the building of this case study, these results call for improvement to better account for ventilation in the RC model.

The numerical methodology is finally applied to explore the physical interpretability of thermal resistance parameters taken individually. In particular, the interpretability of parameters supposedly representing heat transfers through ventilation and heat transfers through unheated neighbouring spaces is assessed. With fixed weather data, the reference building model undergoes changes of thermal properties. Again, each simulation serves as training set for the calibration of appropriate RC models.

Results find that from a non intrusive experiment, neither the estimated transfer coefficient through ventilation nor that towards neighbouring spaces should individually be physically interpreted. The high covariances between all thermal resistance parameters are found to be strong indicators that each parameter could not be individually interpreted. Despite the lack of interpretability of heat transfers decomposition and the high covariances of both models, the estimation of the overall thermal performance is found satisfactorily accurate. These results also suggest that retrofit strategies poorly benefit from non intrusive datasets and need larger and more variate sources of information.

Keywords*: Thermal performance estimation; Inverse problem; RC models; Stochastic state space models; Frequentist calibration; Bayesian calibration; Global sensitivity analysis; Weather influence; Identifiability; Physical interpretability*

General introduction

> *"Under emissions in line with current pledges under the Paris Agreement (known as Nationally Determined Contributions, or NDCs), global warming is expected to surpass 1.5 °C above pre-industrial levels, even if these pledges are supplemented with very challenging increases in the scale and ambition of mitigation after 2030. This increased action would need to achieve net zero CO_2 emissions in less than 15 years. Even if this is achieved, temperatures would only be expected to remain below the 1.5 °C threshold if the actual geophysical response ends up being towards the low end of the currently estimated uncertainty range.*
> *Transition challenges as well as identified trade-offs can be reduced* ***if global emissions peak before 2030*** *and* ***marked emissions reductions*** *compared to today* ***are already achieved by 2030****."*
> (IPCC Special Report: Global Warming of 1.5 °C (Rogelj et al., 2018))

The need of global, systemic and urgent action to keep global warming under 1.5 °C above pre-industrial levels, in the best case scenario, is compelling (Rogelj et al., 2018). As suggested by Figure 1, aside transportation and industry, the residential sector and all the more so the entire building sector has a potential for significant reduction of its energy consumption and thus of its CO_2 emissions.

There are multiple ways for the building sector and in general the built environment to comply to the demanding objectives of reducing energy consumption (Lucon et al., 2014), with in particular highly performant building envelopes, efficient appliances, efficient lighting as well as efficient heating, ventilation, and air-conditioning systems. In overall, Lucon et al. (2014) expect "a decrease by as much as 46 % in heating and cooling energy use as compared to 2005 if today's best practices in construction and retrofit know-how are broadly deployed". While the cost of such large scale building retrofit is significant, estimated around 24 billion USD_{2010}, the savings in cumulative energy costs projected in 2050 exceed them by far.

In addition, Ürge-Vorsatz et al. (2014) show that there are many co-benefits to energy efficiency in buildings:

- ***health effects*** thanks to healthier indoor environments and savings due to avoided sick building syndromes,
- ***ecological effects*** thanks to reduced indoor and outdoor air pollution,
- ***economic effets*** thanks to decreased energy bills, business opportunities as well as employment creation, improved energy security and improved productivity,
- ***service provision benefits*** thanks to reduction of energy losses during transmission and distribution,
- ***social effects*** through fuel poverty alleviation, increased comfort or increased awareness of the energy related issues.

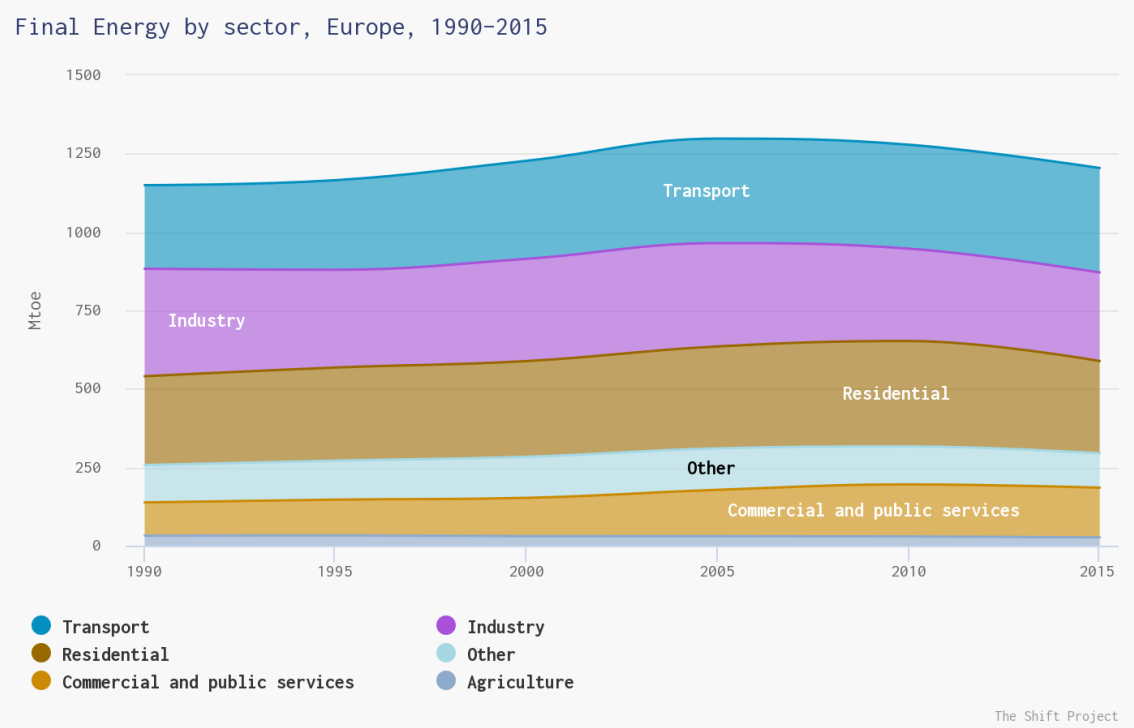


Figure 1 – Final energy consumption by sector in Europe (Sources: IEA and EGEDA, The Shift Project Data Portal). Note: the primary energy consumption remains approximately constant on the same period.

To tackle the issue of energy efficiency of the built environment, consequential efforts must be put in first ambitious energy retrofit of the existing buildings and second highly performant new constructions. On the particular matter of energy retrofit, Sandberg et al. (2016) find that it is very unlikely that the renovation rate of the building stock grows higher than the current 0.6 to 1.6 %, which by 2050 does not cover the entire building stock, at least in France (Sandberg et al., 2016). It is also very unlikely that before 2050 any building would undergo more than a single consequential energy retrofit. This implies that if a retrofit is performed, it must be highly ambitious such as to meet the aforementioned objectives on energy use and CO_2 emissions.

As found by Tuominen et al. (2012), one barrier for building energy retrofit are that the consumers lack knowledge about what energy efficiency is and the extent to which payback time is beneficial for them. Cost optimal retrofit, even in ambitious energy retrofit plans, needs then to be based on an accurate diagnosis of the actual energy performance, notably heat losses of the building envelope and performance of its systems. Heo et al. (2012) for example bases a retrofit analysis on actual estimation of the building envelope and system performance, on the basis of on-site measurements. The presented retrofit analysis includes uncertainty propagation, which can be a useful leverage tool for practitioners in the frame of Energy Performance Contracting (Ligier, 2018).

Meaningful retrofit analysis must however rely on accurate knowledge of the energy performance of the envelope and of the building systems. If such information is unavailable, energy retrofit predictions will have very large uncertainties or even be biased. With a focus on the thermal performance of a building envelope, the most promising ways to make such diagnosis is to exploit on-site measurements in controlled measurements, i.e. with optimal heating sollicitation as to reduce noise and perturbations and thereby enhance data analysis. The coheating test for example is a controlled experiment that intends to accurately estimate the heat losses of a building envelope (Jack et al., 2018).

Controlled experiments for thermal performance of the envelope are however not always

feasible, when for example the building cannot be left vacant at all, such as in hospitals, elderly homes, or not for long such as offices, hotels or industrial buildings. Yet, the energy saving potential of the buildings in these sectors may be significant. There is therefore a need to develop accurate diagnosis methods from on-site measurements compatible with occupancy.

Occupant-friendly experimental set ups imply two major obstacles to accurate thermal performance assessment, that can be categorized as follows:

- poorly informative data:
  - the experiment cannot be controlled in the sense that it cannot follow a specific design of experiment to enhance the diagnosis accuracy. The indoor conditions, designed by the occupants, are then not to be changed.
  - the number of sensors to be placed should be kept at a reasonable number and in reasonable places as to remain occupant-friendly,
- perturbed data:
  - the occupants represent a source of perturbation to the energy behaviour of the building: opening and closing doors or windows, cooking, cleaning, showering and using various energy consuming appliances. The influence of this behaviour is very difficult or impossible to measure unless intrusive sensors are extensively placed in the building.

This work intends to tackle the issue of poorly informative data and how it affects the identification accuracy of thermal losses of a building envelope. This issue is indeed a first step towards thermal performance estimation of the envelope in uncontrolled conditions where measurements may be disrupted by occupants. Indeed, if with poorly informative data exploitation is infeasible, it will not be either with data from measurements under occupancy.

Chapter 1 proposes first a review of existing methods to estimate the thermal performance of a building from on-site measurements. The reasons why such characterisation is an ill-posed problems are exposed, which leads the chapter to then detail how to tackle an inverse problem: finding an accurate model and using appropriate algorithmic tools to solve it numerically. RC models in a stochastic formulation are found to have promising properties for efficient data exploitation. Regardless of the chosen models, two approaches are considered for numerical estimation: frequentist or Bayesian approaches, which advantages and drawbacks are discussed.

Chapter 2 deals with the existence and unicity of the estimation of thermal performance of the envelope, called the identifiability, and exposes how this notion is manifold. Identifiability indeed relies on an adequate structural model definition and on sufficiently informative data. Although the estimation can be proven a posteriori to be unique, the poorly informative nature of the data implies possible bias in the estimation. The chapter therefore discusses how data collection as well as model selection and validation may limit the bias in estimation. A good practice is finally proposed for meaningful calibration. Whether proven identifiability then implies physical interpretability of the results remains however uncertain.

Chapter 3 proposes an original methodology to assess the physical interpretability of a model and proposes a framework to apply it to the assessment of stochastic RC models trained on poorly informative data. The methodology relies on a computer based and detailed reference building energy model. The reference model is simulated in known conditions and the simulation output serve as training data for the model under study, which can afterwards be assessed. A novel assessment indicator is defined and reflects on both the accuracy and the uncertainty with respect

to the targeted value. The novelty of this methodology is also to wrap this assessment procedure in a global sensitivity analysis design. As such, the methodology better grasps how the uncertainty of the estimation and of their interpretability is influenced by the uncertainties in the building envelope environment.

Chapter 4 applies this methodology to the study of variability of weather conditions on the estimation of the thermal performance. Thus, the chapter aims also at determining a minimal measurements duration that achieves repeatable estimations. The model assessment framework therefore simulates the reference model in variable weather conditions. A set of synthetic weather variables is used, stochastically generated as to allow global sensitivity analysis. The selected RC model is found to yield repeatable estimation of an overall thermal resistance.

Chapter 5 proposes to apply the aforementioned methodology to assess the physical interpretability of model parameters individually. Stochastic RC models may indeed be classified as lumped physical models. This implies that, to a certain extent, their parameters have a physical meaning. However, in the context of poorly informative data and of inverse problems in general, whether they may each be interpreted is not assured. In application of the model assessment framework, the reference model is simulated with variable thermal properties of its envelope. Two models are assessed for their ability to determine the heat transfers through ventilation and towards an unheated crawl space.

1

Estimation of thermal performance from measurement data: an inverse problem

The estimation of the thermal performance of a building envelope from measurement data constitutes an inverse problem. Using data from non intrusive measurements makes this problem into a possible particular case of ill-posed inverse problems.
To solve it, one must first consider a proper model for the system under study. Then, one needs to apply an appropriate algorithm to numerically estimate the values of each parameter of the chosen model such as to properly fit a given dataset. This chapter concludes with orientations and objectives for the work to follow.

Contents

1.1 Introduction

Thermal performance of a building has been a scientific interest for decades now. Areas of interest cover as much as the energy efficiency of the building at design stage or during operation, sometimes in relation with quality of the indoor environment. At building scale, thermal performance relates indeed to robust and energy efficient combinations of building envelopes and systems under occupancy.

This work focuses on the estimation of the thermal performance of the building envelope. In this scope, thermal performance of the building envelope is meant as the actual on-site thermophysical characteristics of the envelope. Characteristics of interest may be the actual heat loss or thermal resistance, thermal capacities and air change rates through infiltrations and ventilation. Semantically, there does not seem to be a consensus as to whether these should be designated as characteristics or as thermal performance as authors tend to use with the same meaning at least at envelope scale: "thermal performance" (Bagheri et al., 2017; Chambers, 2017; Farmer et al., 2017; Ji et al., 2019), "thermal performance characterization" (Madsen et al., 2015), "thermal characterization" (Bouache et al., 2013; Chambers, 2017; Rendu et al., 2019) and as the estimation of "thermophysical characteristics" (Gori et al., 2017; Jiménez et al., 2008).

Regardless of the semantics used, estimation of the thermal performance of the envelope is a data based inference process. This chapter develops first in section 1.2 how existing methods perform thermal performance estimation. This review brings into light how the estimation of the thermal performance in a non intrusive framework constitutes a particularly difficult problem to solve, which is further detailed in section 1.3. To make an attempt, both an appropriate model (section 1.4) and adequate numerical tools (section 1.5) need to be chosen.

1.2 Existing methods for the estimation of the thermal performance of a building envelope

To estimate the thermal performance of a building from measured data has raised multiple data-driven methods. One may divide them into two main categories: steady-state or dynamic methods. At the same time, the methods may rely on destructive/intrusive or non-intrusive experiments and may be performed in controlled or uncontrolled conditions.

Although the thermal properties of a building envelope may be studied at wall scale with heat flux methods or thermography, and literature is abundant on this matter (Biddulph et al. (2014); Bienvenido-Huertas et al. (2019); Chaffar (2012); Gori et al. (2017); ISO 9869-1 (2014); Rasooli and Itard (2018); Rodler et al. (2019); Yang et al. (2019) just to name a few recent), this section will only detail methods estimating an overall thermal performance of the building envelope.

With the term *methods* is actually meant the combination of:

- a specific design of experiment (acquisition of *Data* in Figure 1.1);
- and its numerical methodology to infer the estimation from the collected data:
  - the type of model used for the energy balance or for the heat dynamics in the building (*Model* and *parametrization* in Figure 1.1);
  - as well as the numerical tools to solve the inverse problem (*solver* and *inference* in Figure 1.1).

The usual distinction between steady-state and dynamic methods refers to the choice of model and relies on whether the hypothesis of thermal equilibrium of the energy balance equation

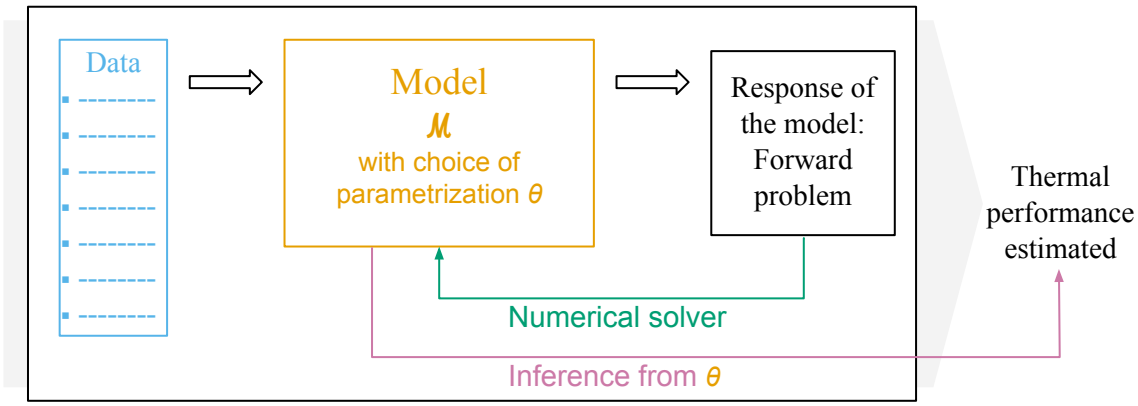


Figure 1.1 – Estimation of the overall thermal performance of a building: a combination of collected data, model choice and numerical tools to infer the quantity of interest

is accepted or not. In other words, upon considering the energy balance of the building, one may consider the boundary conditions sufficiently invariant over a certain chosen time span and consider the heat storage in the building elements negligible. At equilibrium, the latter is indeed very low. In steady-state conditions, the energy balance then constitutes a linear regression model, solved through least square methods. On the contrary, dynamic conditions will consider heat storage and variations of boundary conditions non negligible and incorporate ad hoc dynamic terms in the formulation of the model. Let us finally underline that the methods described hereafter are based on or inspired by the physical framework described in the PSTAR procedure Subbarao (1988), which is itself more a proper framework to a thorough analysis of the energy performance than a method.

Very basically, the steady state methods used for parameter interpretation rely on a simplified static energy balance of a building in Equation 1.1 where the heat charging in and out of the building elements $Q^*_{storage}(t)$ are considered null.

$$\begin{aligned} HTC(T_i(t)-T_o(t))+Q^{in}_{storage}(t)+Q^{out}_{storage}(t)-Q_{sun}(t) \\ +Q_{ventilation}(t)-Q_{heating}(t)+Q_{ground}(t)-Q_{internal}(t) \quad = \quad 0 \end{aligned} \tag{1.1}$$

Depending on the measured variables and simplifications, this equation may result in a simple static equation as in Hammarsten (1987), with Equation 1.2. It merely becomes the relationship between the energy consumption Q, the temperature difference between indoors and outdoors $T_i - T_o$, the solar heat gains through S the solar aperture and I the solar irradiation and Q_0 constant indoor heat gains:

$$Q = Q_0 + HTC(T_i - T_o) + SI_{sol} \tag{1.2}$$

Ferlay (2012) proposes a variation of Equation 1.2 by replacing the temperature difference by degree-days. Degree-days then allows to use monthly averaged energy consumptions, which are easier to collect.

Applications of this steady-state building energy model with among others Day et al. (2004); Fels (1986); Ferlay (2012); Hammarsten (1984, 1987); Lee et al. (2014); Sonderegger (1978) use data of buildings under occupancy, which constitutes a major advantage as it may be easily used by any facility manager and with basic knowledge of the physics nor accurate measurement of the building.

Solving the equation is done through ordinary least square methods, to which Subbarao (1988) suggests using generalized least squares to improve the auto-correlated residuals.

If some of the applications are meant for energy prediction, parameter interpretation should be done cautiously (Hammarsten, 1987). The author underlines that with a static equation such as 1.2, the time step of the data should not be shorter than 24h. Additionally, parameters estimates need to be unbiased, by which the author means parameters to have insignificant covariances. Sources of error for physical interpretation are, according to the author, due to a rough representation of the physics and/or multi-collinearity of the input variables T_o and I_{sol} and the author proposes to preferably use more advanced models such as dynamic energy-balance models as in Sonderegger (1978).

While the above mentioned applications use data from occupied buildings, i.e. without particular conditions for the measurements, co-heating test method rely on a particular design of experiment (Johnston et al., 2013). As reported in Bauwens and Roels (2014), the usual setting is to uniformly heat a building at a constant temperature, reaching at least a $10K$ temperature difference between indoor and outdoor, for example 25 $^{\circ}C$. The building is to be left unoccupied for several days to several weeks as to reduce uncertainties and perturbations. Weather variables are measured, if available: temperatures, wind speed and direction, solar irradiation, relative humidity. The collected data is aggregated with at least 24 h time steps, but measurements are done preferably with lower time steps as to catch high frequency dynamics that would create uncertainties in the data. Suitable starting time for averaging is suggested in Everett (1985) with a start around 6 am, thus asserting the steady state hypothesis. The co-heating is coupled to an estimation of the heat losses by infiltration through a tracer gas test or a blower-door test. Knowing the infiltration losses, it becomes possible to decompose the overall heat transfer coefficient obtained by the direct exploitation of the data into transmission heat transfers by deducing the infiltration heat transfers.

The data from a co-heating test is exploited by a linear regression analysis of Equation 1.2. As suggested by Bauwens and Roels (2014), performing a linear regression analysis instead of simply averaging over the total dataset allows to deal with outliers more easily. Indeed, according to the adaptation of the energy balance equation to the building configuration, it is possible to take more than one confounding factors, like solar gains or wind effect (Jack et al., 2018). Various techniques to plot the results of the regression analysis directly on a graph are available (Bauwens and Roels, 2014; Jack et al., 2018) and make the interpretation of the results and their uncertainty easier.

Let us conclude on the co-heating method by underlining that controlling the indoor temperature holds the advantage to increase the accuracy of the estimation result, see Jack et al. (2018). The same authors established a ± 8 to 10 % error, pinpointing the residual variability to the difficulty to correctly estimate the solar gains.

As made clear in Bauwens and Roels (2014), the previously mentioned methods are steady-state assumptions, considered true in the conditions in which the data is collected and exploited, of a fundamentally dynamic system. Other methods take advantage of the dynamic nature of the heat transfers in buildings. These methods rely then on a dynamic formulation of the energy balance equation as in Sonderegger (1978), where the state of the dependence of the system at $t = t_n$ to $t = t_{n-1}$ is literally accounted for. The overall advantage of using a dynamical energy balance model is to achieve satisfactory accuracy in fewer days. In overall, methods based on dynamic models use state space models in a statistical form like auto-regressive models or state space models inspired by the heat transfer dynamics like RC models.

Auto-regressive with exogenous inputs (ARX) models are based on an input-output relation: a chosen output is considered function of some influential inputs. For example, as is described in Madsen et al. (2015), the heating power can be considered as a time dependant function of indoor temperature, outdoor temperature and solar radiation, which results in the functional relationship

of Equation 1.3 where B is the backshift operator and $\phi(B)$ a polynomial of order p such as $\phi(B) = 1 + \phi_1 \cdot B + \phi_2 \cdot B^2 + ... + \phi_p \cdot B^p$. Put very simply, the backshift operator links, at order p, the variable Φ at state t to its p previous states, i.e. function of the states $t-p$ to $t-1$. The model order is inferred iteratively aiming at ε a Gaussian white noise, which means that the ARX prediction would fit properly the data.

$$\phi(B)\Phi_t^h = \omega_i(B)T_t^i + \omega_e(B)T_t^e + \omega_{sol}(B)I_t^{sol} + \varepsilon_t \tag{1.3}$$

The indoor temperature may also be considered as the output variable. Its dynamic behaviour is then described by Equation 1.4.

$$\phi(B)T_t^i = \omega_h(B)\Phi_t^h + \omega_e(B)T_t^e + \omega_{sol}(B)I_t^{sol} + \varepsilon_t \tag{1.4}$$

Although ARX models have been used for prediction purposes, an overall Heat Transfer Coefficient (HTC) can be estimated from its parameters (Madsen et al., 2015). For example, with the heating power as output, the HTC is estimated by $H_{tot}(\lambda) = \lambda H_i + (1-\lambda)H_e$ such that the variance of $H_{tot}(\lambda))$ is minimal, with $H_i = \frac{\omega_i(1)}{\phi(1)}$ and $H_e = \frac{-\omega_e(1)}{\phi(1)}$ (again see Madsen et al. (2015) for the detailed procedure).

In Senave et al. (2019) and Senave et al. (2020a), the authors establish promising performance of ARX models in various occupant-friendly conditions, depending on the availability of measurements of some boundary conditions such as solar irradiation or heat fluxes through poorly insulated ground floor slabs. Error goes from a few percents up to almost 25% in the case of solar irradiation approximations (Senave et al., 2020a). In overall, within a few weeks ARX parameters yield robust estimates of heat loss coefficients.

Next to ARX based methods are methods based on linear state space models, also called RC models (Madsen et al., 2015). Unlike ARX models which is a simple input-output relation, RC models naively approximate the heat transfer physics in the building. Under some assumptions, they can be described by linear terms using time invariant parameters. As such, RC models can be written in a state space representation: a set of first order differential equation that presents convenient properties.

Existing methods using RC models are the Quick U-Building (QUB) method (Alzetto et al., 2018a,b; Mangematin et al., 2012; Meulemans, 2018; Meulemans et al., 2017) or the ISABELE method (Boisson and Bouchié, 2014; Schetelat and Bouchié, 2014; Thébault, 2017; Thébault and Bouchié, 2018). There are also other punctual applications to heat transfer estimation (Rouchier et al., 2019, 2018) or building thermal behaviour prediction (Bacher and Madsen, 2011). These methods can be distinguished by the design of experiment used to control measurement conditions and by the RC models used.

The QUB method relies on a rather short data acquisition duration, compared to the other methods. Performed during one night as to avoid solar irradiation, the heating power is programmed to follow a two phase scenario with first large constant heating power followed by a close to free floating phase, during which the indoor temperature slowly decays as the heating power is set to minimal. The heating power step of the first phase is designed as to result in an exponential curve of temperature. With sufficient duration of each phase, a few hours, it is expected that the largest time constant, the one of interest, is visible in the response of the building.

The data is exploited by the simplified dynamic model of Equation 1.5:

$$P = HLC \times (T_{in} - T_{out}) + C \times \frac{dT_{in}}{dt} \tag{1.5}$$

This equation can be considered accurate for each phase, and the overall heat loss coefficient is then defined by Equation 1.6, where P_i, ΔT_i and a_i are respectively the total power, the inside-outside temperature at the end and the derivative of the indoor temperature at the end of phase i.

Strictly speaking, the QUB method does then not actually make use directly of the RC model but rather uses the derivatives at

$$HLC = \frac{P_1 a_2 - P_2 a_1}{\Delta T_1 a_2 - \Delta T_2 a_1} \tag{1.6}$$

Results of HLC estimations in various buildings show deviation from target reference value up to 15%. Variable weather conditions such as strong wind and large outdoor temperature variations may lead to larger errors in HLC estimations, as suggested in Alzetto et al. (2018a).

The ISABELE method was developed to characterize the thermal properties of the building envelope when commissioning a newly built building. The goal is to determine heat transfers through the envelope, excluding the air infiltrations. The latter are measured prior to the HTC measurement and is included in the data exploitation.

For the HTC estimation, the ISABELE method also relies on a specific controlled experiment: after a short free floating phase, the indoor temperature is set to follow a long step signal. During the experiment, the window blinds or shutters are closed and the hypothetical ventilation shut. The experiment lasts longer than for the QUB method, as it has been found that more than 2 days led to more robust results (Thébault and Bouchié, 2018).

The collected data is used to fit an RC model, which takes into account the infiltration priorly measured. Although uncertainties of the results are calculated by the calibration method (Nelder-Mead (Boisson and Bouchié, 2014) or MCMC (Schetelat and Bouchié, 2014)), Thébault and Bouchié (2018) propose to correct their widths by propagating through a Monte-Carlo the most influential systematic measurement errors, such as temperature or heat power errors. Systematic errors are indeed not taken into account in the calibration process by any means other than uncertainty propagation, whereas they have a significant influence on accuracy.

Finally, Bacher and Madsen (2011); Rouchier et al. (2019, 2018) also suggest dynamic models to exploit in situ controlled data. Bacher and Madsen (2011) proposes a model selection process to determine what stochastic RC model is best fit for a given set of data. The authors have however heat dynamics prediction in view rather than thermal performance inference. The collected data is that of an unoccupied office building which heat power is controlled to follow a so-called pseudo-random signal.

Rouchier et al. (2018), on basis of a heating power step signal of two days, compare stochastic RC models with deterministic RC models and conclude that parameter estimation with stochastic models is not closer to the target values, but that at least the associated uncertainties are more realistic. In Rouchier et al. (2019), a pseudo-random heating power signal in a $1m^3$ box serves as training for an on-line algorithm called sequential Monte-Carlo. It provides RC model training with an update of the parameter estimation at each new observation. The authors show how a sequential Monte-Carlo yields similar parameter estimation than a Marginal Monte-Carlo algorithm and that the results are in agreement with the target values for both the HTC and the solar aperture coefficient (i.e. the coefficient bringing correction of the solar irradiation heat gains).

Interestingly, the in-line algorithm revealed the most influential events that enhanced parameter estimation. In particular, it uncovered that the model learns a lot when heating power is finally turned on. This outcome brings into light how useful information in the collected data can be to model training, in particular for dynamic models.

Table 1.1 lists as a conclusion the aforementioned methods. The most accurate estimations of the thermal performance unsurprisingly result from controlled measurements conditions. When the building cannot be left vacant, some existing methods are compatible with uncontrolled conditions. Methods relying on steady-state models require almost no data acquisition material, but need data over several months, seemingly without guarantee of physical interpretability. Methods relying on dynamic models require shorter duration datasets, and seem to reach more

satisfactory results.

This section is not viewed as an exhaustive literature review of all methods for characterising the thermal performance of a building. To dig further into that subject, the curious will find reader-friendly sources in Zayane (2011), Bauwens and Roels (2014), Bauwens (2015), Janssens (2016), Thébault (2017) and Raillon and Ghiaus (2018).

	References	Duration of data acquisition	Compatible with occupancy	Accuracy to reference value
Steady state under occupancy	Hammarsten (1987)	months	yes	*not quantified*
Co-heating	Johnston et al. (2013) Bauwens and Roels (2014) Jack et al. (2018)	$<$ 3 weeks	no	$\pm$10% (Jack et al., 2018)
Auto--regressive models	Senave et al. (2019) Senave et al. (2020a)	a few weeks	yes	$\pm$5% (Senave et al., 2019) $\pm$15% (Senave et al., 2020a)
QUB	Mangematin et al. (2012) Alzetto et al. (2018a) Alzetto et al. (2018b) Meulemans et al. (2017) Meulemans (2018)	$< 48h$	no	$\pm$15% (Alzetto et al., 2018a)
ISABELE (EVAREPE)	Thébault (2017) Thébault and Bouchié (2018) Schetelat and Bouchié (2014) Bouchié et al. (2014)	2-5 days	no	$\pm$15% (Thébault and Bouchié, 2018)
State space with PRBS	Bacher and Madsen (2011) Rouchier et al. (2018) Rouchier et al. (2019)	$<$ 1 week	no	*not quantified*

Table 1.1 – Overview of existing methods used for thermal performance estimation as described in Section 1.2: duration of data acquisition, compatibility with occupancy and accuracy to reference value of thermal performance are variable.
Note: the mentioned accuracies are extracted from different case studies and in different conditions, methods might score differently if compared against a common benchmark.

1.3 From controlled to uncontrolled conditions: an ill-posed inverse problem

1.3.1 Inverse problems: some definitions

For clarity, let us first define some concepts that will thoroughly be used in this work. Definitions are freely inspired by those in Walter and Pronzato (1997) or Muñoz-Tamayo et al. (2018).

A *system* $\mathfrak{S}$ is a part of the universe, chosen more or less arbitrarily, considered as the entity to study and that has interaction as a whole with some external variables. The system is delimited by spatial and temporal boundaries.

The external variables that act on the system and internal state variables are most probably not all observable, i.e. measurable. The observable quantities that influence the system's behaviour are usually called *inputs* of the system and noted $\boldsymbol{u}$. The remaining influential variables can then be considered as perturbations or simply noises. They are not controllable and sometimes difficult to apprehend. The internal and observable quantities may then be called *output* of the system $\boldsymbol{y}$. Some other quantities might be of interest, but are not directly observable.

The system and its boundaries as well as the observables and the input variables then drive the construction of a *model* of the system. A model, as defined by Walter and Pronzato (1997), is "a rule to compute from quantities observed from the system other quantities which are hoped to be close to the actual values as observed in the system". A model can be understood as a set of equations that describe the physics of interest in the system. The model links the observable inputs with the observable outputs of the system. The model may be linear or non-linear (in its inputs): it is linear if for all λ and μ, $y(\lambda \cdot u_1 + \mu \cdot u_2) = \lambda \cdot y(u_1) + \mu \cdot y(u_2)$.

Some models, such as the RC models mentioned earlier in section 1.2, describe quantities of the system by their derivatives with respect to time as a function of their previous state. Such models can be called state space models describing state variables $\boldsymbol{x}$. RC models are incidently linear in their inputs.

The set of equations describing the system uses *parameters* usually named $\boldsymbol{p}$ or θ. They are scalars (not vectors) and usually time-independent. When all parameters are known, one may speak of model $\mathscr{M}(\theta)$. Before actual values are attributed to the parameters, one talks about a model *structure* $\mathscr{M}(.)$. A structural property is then a property that is valid for almost all values of parameters (Bellman and Åström, 1970). Linearity is an example of structural property of a model (Muñoz-Tamayo et al., 2018).

Model calibration is the "action of using a numerical routine, algorithm, for finding the value of unknown parameters of a model that best fit an experimental data set" of observable quantities (Muñoz-Tamayo et al., 2018). Model calibration is also called *parameter identification* (or *estimation*) and model fitting (Muñoz-Tamayo et al., 2018).

Walter and Pronzato (1997) expressly underline the importance of estimating parameter *uncertainty* along with parameter identification. Indeed, it would be foolish to consider that given the measurement uncertainties, the result of model calibration would yield a unique set of parameters. On the contrary, there is a set of acceptable models and that translates into parameter estimates and their uncertainty.

Upon developing an appropriate model to fit some collected data, the model can possibly contain more parameters than needed to describe the system output. This is called over-parametrization (Muñoz-Tamayo et al., 2018) and results in the model over-fitting the data (Rouchier, 2017).

When applied to the thermal performance estimation of a building, the *system* may be the building in its whole, or simply the building envelope. Observable *inputs* and *outputs* are for example weather conditions, heating or cooling power, indoor temperatures, wall surface temperatures, etc...

When estimating the overall thermal performance, one is then interested in a quantity that is not directly observable. This whole procedure is called solving an inverse problem. Upon considering any system $\mathfrak{S}$, here the building envelope, Tarantola (2005) considers that inverse problems are part of three arbitrary scientific procedures to study such a system and underlines that they obviously are strongly linked:

- parametrization of the system: determining a minimal set of parameters that completely characterize the system from a given point of view. Characterization here means almost perfectly describing the physical phenomena at hand in the system.
- forward modelling: discover the physical laws that allow predictions of the quantity of interest.
- inverse modelling: use actual measurements to infer the actual values of parameters (or variables) of interest.

Let us underline that a distinction is made between forward and inverse modelling, which implies that as the goals are different, the appropriate models to the one or the other might be different as well. In other words, a valid model for forward modelling might be misleading in inverse modelling. The choice of an appropriate model for solving this particular inverse problem is therefore further discussed in section 1.4.

Very generally, among the issues raised upon solving an inverse problem, the most crucial might be the questions raised by Hadamard (1907): does the solution exist? is the solution unique?

- if data collection were ideal, **existence** would not be an issue: given that the thermal performance estimation here is based on actual physics and given an appropriate model, the solution exists de facto. Now abandoning the assumption that the collected data are ideal, the issue resides in the fact that the measurements are not exhaustive, discrete and noisy, which is why it is difficult to find a model that perfectly characterizes the system from imperfect data. Stuart (2013) argues that with data perturbed by noise, the inverse problem becomes $y = G(\theta, u) + \varepsilon$ where ε is the measurement noise and where $G : \mathbb{R}^p \times \mathbb{R}^n \mapsto Im_G$ is assumed to have an inverse. There might arise two problems:
  - $y \notin Im_G$ because of the noise ε, so inverting G from y will not be possible. And as the noise ε can only be statistically defined, it is not possible to subtract it to y to inverse the equation hence the existence issue.
  - even with $y \in Im_G$, the noise ε causes uncertainty and in the best case scenario, the solution to the inverse problem is inaccurate.
- the question of **unicity** relates to the possibility that multiple sets of parameters lead to the exact same measured output for a given input. Non unicity leads to identical input-output behaviour for multiple sets of parameters. It is then not simply an issue of parameters and their uncertainty, where there would still be one single most plausible set of parameters. Let us here also emphasize that identical input output behaviour is theoretically different than identical states at all times. In continuous time, two models might have different states but as soon as they are measured, they could yield the same output in a discrete and finite sampling. It follows that the question of unicity of the solution is not just a question of model but also a question of data, and that it becomes all the more difficult that the data is limited.

If the solution to an inverse problem exists and is unique, the problem is said to be well-posed (Hadamard, 1907; Maillet et al., 2011). The inverse problem is all the more well-posed that its solution is not too sensitive to small variations in the data. If not, it is said to be ill-posed and the estimation of the solution will become difficult.

1.3.2 Loss of information in uncontrolled conditions

The review of existing methods in Section 1.2 showed how inference from controlled conditions yielded the most accurate results in shorter experiments. Although outdoor conditions remain variable, controlled experiments induce a specific signal for the heating power. These type of signals have the advantage of being non destructive and at the same time enhance accuracy of the parameter estimation. Usual signals are:

- highly informative in optimized signals like the PRBS (Pseudo-Random binary signals) or ROLBS (Randomly Ordered Logarithmically distributed Binary Sequence) signals (Godfrey, 1979). These types of signals are synthetically created to optimally excite the system under study while reducing correlations between input variables. To cover a sufficient number of frequencies, the experiment usually lasts at least a few days.
- sufficiently informative like in the QUB (Alzetto et al., 2018b) or ISABELE (Boisson and Bouchié, 2014) tests. Heating signals are not designed to cover a wide range of frequencies, but cover one or some of the essential frequencies as to guarantee satisfactory estimation of the parameters of interest. As less frequencies are included on the heating signal, amplitudes are larger and make the indoor temperature rise up to $35°C$.

When the building cannot be left vacant, such controlled conditions can however not be achieved. At the same time, uncontrolled conditions are advantageously less demanding in terms of measurement devices, such as heating supply and fans, used to homogenize temperatures across the building. The experimental setup is then obviously non destructive, but also reduced to occupant-friendly heating schedules and indoor temperatures. Without characterizing in detail usual heating schedules, it can be expected that the achieved indoor temperatures vary between $17°C$ and $23°C$ in winter, with possible lower temperature set points when absence or at night (Huebner et al., 2013). In overall, the temperature amplitudes will certainly not exceed a few degrees. Such heating schedules have then probable correlation with outdoor weather variables, as absence during the day as well as day and night alternation have approximately the same frequencies. In comparison to controlled conditions, occupancy-friendly experiment deliver therefore poorly informative data.

As a consequence, occupancy-friendly experimental data turns the inverse problem of estimating heat transfers of a building envelope into a particularly ill posed one as they accumulate several issues:

- data collection is limited;
- not exhaustive, regarding the complexity of the heat transfers at building scale;
- even possibly not representative or biased.

All in all, data from non intrusive and occupant-friendly can be considered as poorly informative.

These measurement conditions question the accuracy that can be obtained from such data as poorly informative data could bias the inference. They also question whether satisfactory uncertainties can be obtained after all.

1.4 Appropriate models for thermal performance estimation

Regardless of the methods mentioned in section 1.2, solving the inverse problem of estimating the thermal performance relies on choosing an appropriate thermal model, fit to the constraints

of this particular work. This section reviews types of modelling in the field of heat transfers and focuses in particular on RC modelling.

For clarity, let us remind here that, with the choice of semantic from Walter and Pronzato (1994), *characterization* of a system $\mathfrak{S}$ means determining a sufficient mathematical model, i.e. a set of equations describing the system, or at least describing the phenomena under study of the system. It is merely a choice of *structure* for the mathematical model. Let us then note $\mathscr{M}(.)$ the structure characterizing the system. A *model* $\mathscr{M}(\theta)$ is the model from structure $\mathscr{M}(.)$ which parameters take the value θ.

1.4.1 From data-driven to physics-driven modelling

Upon choosing an appropriate model $\mathscr{M}(\theta)$ for the estimation of the thermal performance of a building, one may consider rather data-driven or physics-driven models.

Purely data-driven models, also called *black box models*, describe some particular measured output variables given some measured input variables without the structure reflecting on the actual physical system (Bohlin and Graebe, 1995). Machine learning models or time series linear models such as ARMAX structures are black box models (Bohlin and Graebe, 1995; Rouchier, 2017). The parameters, states and model structures are very much unlikely to bear any physical meaning. This type of models is used for its prediction properties rather than for its ability to infer actual properties (Rouchier, 2017).

On the opposite side, *white box models* are completely physics-driven. It requires a fine knowledge of the physics at work in the system to drive an appropriate choice of structure and model to describe it (Bohlin and Graebe, 1995). In the field of building physics, this category covers complex models such as coupled heat and moisture transfer models as much as heat only transfer models (Rouchier, 2017). Bohlin and Graebe (1995) underlines that white box models are purely deterministic models: discrepancies between the measured output and the model prediction are attributed to the measurement noise (provided the model perfectly characterizes the system). White-box models are rather chosen for the purpose of forward modelling as intended in Tarantola (2005): model, evaluate and/or predict the behaviour of building components and systems, such as their energy consumption, the indoor comfort or indoor temperatures at different scales and times (Foucquier et al., 2013b; Li and Wen, 2014). However, applications to thermal property inference from measured data is limited because of its extensive number of parameters. In particular upon exploiting poorly informative data, the risk of overfitting arises (Bertsekas and Tsitsiklis, 2008): the white box model might extremely well fit the data but its parameters lose their physical meaning. In addition, the large number of parameters of a white box model also raises the curse of dimensionality: finding the best fitting model with n parameters is exponentially harder from a computational point of view than finding the best fitting model with $n-1$ parameters. To simplify the process, it has been proposed to first perform a sensitivity analysis on all parameters to determine which one are the most influent, which reduces the number of parameters to fit (Strachan, 1993). Still, this implies to set some of the non influential parameters to an arbitrary value. If thermal performance inference is the objective, it barely makes sense to set values that are yet to be interpreted.

An alternative to extensive physics-driven models are simplified models with yet physical meaning (Kramer et al., 2012), such as RC models or linear regression models. For the latter, the heat transfers are, under some conditions, roughly supposed to be at equilibrium (see section 1.2), which derives from a naive understanding of the physics. Whenever the assumption of steady-state, i.e. equilibrium holds, the parameters of the model are physical and represent aggregated thermal properties.

As for RC models, also called lumped capacitance models (Kramer et al., 2012), they are derived from the linearisation of the otherwise non linear heat transfers in the building. Linearisation induces a simplification of the set of equations describing the system and requires

fewer parameters. As underlined in Afram and Janabi-Sharifi (2015), simplified white box models present the advantage of good generalization capability and high accuracy thanks to the physics-driven model structure, as well as ease to determine the model parameters from measured data.

RC models may be used as deterministic models and as such may be considered as white box models too (Kramer et al., 2012; Wang and Xu, 2006a) although less detailed in their description of the building physics. But Macarulla et al. (2018) stress out that any deterministic model cannot deal with the uncertainties from modelling shortcomings or biased measurements. An alternative is to add a stochastic term in the formulation of the RC model, which introduces a part of data-driven modelling in the physics-based RC model. This stochastic term takes the form of a diffusion term in the differential equations and represents the uncertainty of the pure deterministic drift term of the RC model (Bacher and Madsen, 2011; Kramer et al., 2012). Stochastic RC models enter the category of *grey-box* or *hybrid* modelling which combine physical knowledge and data-embedded information (Bohlin and Graebe, 1995; Foucquier et al., 2013b; Kramer et al., 2012; Li and Wen, 2014; Macarulla et al., 2018; Madsen et al., 2010).

As a conclusion, black box models are in general not physically interpretable, except for ARX models which may provide estimates of steady-state properties. White box detailed models are particularly fit to forward modelling. But as already underlined (Tarantola, 2005), models fit for forward modelling might not be fit for inverse modelling. Simplified white box models and grey-box models seem to be a reasonable choice.

1.4.2 RC models: physics-driven simplified models

The general form of RC model structures derives from a discretization and further simplification of the equations describing the heat transfers in a building. For the sake of intelligibility, this section will use the term *RC models* in plural form to designate *RC model **structures***. It would be more correct to mention structures (see section 1.3.1), as any values of parameters are considered possible so far.

To begin with, Fraisse et al. (2002) shows how simplifications of the otherwise non linear heat transfers lead to simple linear models at building scale, reproduced in Equations 1.7. Conduction, under the hypothesis of 1-dimension transfers, can be written as in Equation 1.7a, with a the thermal diffusivity ($m^2 s^{-1}$). When discretized and considering the closest nodes, one may re-write the equation in an electrical analogy as in Equation 1.7b. When only a few nodes are considered in each wall, these simplifications ultimately yield, for each wall, simple linear models such as $2R1C$, $3R2C$, etc... Fraisse et al. (2002) also shows that it is possible to aggregate different RC models, if each represents a wall, provided that the walls are sufficiently similar.

$$\frac{\mathrm{d}^2 T(x,t)}{\mathrm{d}x^2} = \frac{1}{a}\frac{\mathrm{d}T(x,t)}{\mathrm{d}t} \tag{1.7a}$$

$$\frac{T(x+\frac{\Delta x}{2},t)-T(x,t)}{R/2} - \frac{T(x,t)-T(x-\frac{\Delta x}{2},t)}{R/2} \approx C\frac{\mathrm{d}T(x,t)}{\mathrm{d}t} \tag{1.7b}$$

$$\Phi_{conv} = \frac{1}{R_{conv}}(T - T_{air}) \tag{1.7c}$$

$$\Phi_{rad}^{ext} = \frac{1}{R_{vc}}(T - T_c) + \frac{1}{R_{env}}(T - T_{env}) \tag{1.7d}$$

$$\Phi_{rad}^{int} = \frac{1}{R_{rad}}(T - T_{mr}) \tag{1.7e}$$

Convection (Equation 1.7c) as well as long wave (Equation 1.7d) and short wave radiative transfers (Equation 1.7e) can also be linearised and modelled in an electric analogy through resistances. Each resistance, respectively convection R_{conv}, sky R_{vc}, nearby environment R_{env} or indoor radiative R_{rad}, depend respectively on temperatures of the air along the surface, the sky, the

environment in average and the indoor surfaces in average. As an example, R_{rad} equals $\frac{1}{4\varepsilon\sigma T_m^3 S}$, ergo the temperature dependency. Under the assumption that these temperatures do not vary a lot and that they take values close to the temperatures of the building elements, one may consider the resistances yet invariant.

All heat transfers in the building can therefore be considered linear and written in an electrical analogy. Figure 1.2 shows three examples of combinations of the linearised equations from 1.7 to form RC models of walls.

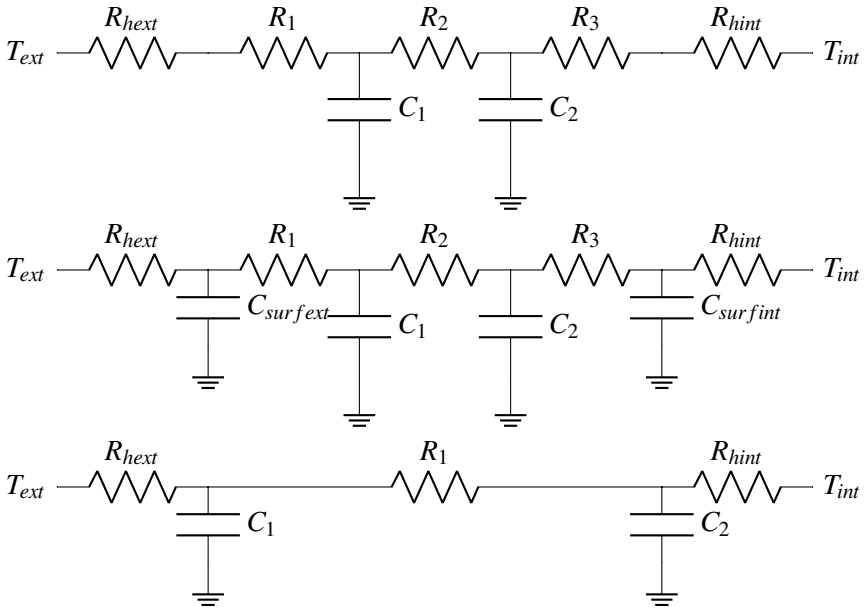


Figure 1.2 – Linearized heat transfers at wall scale in an electrical analogy (Fraisse et al., 2002)

Fraisse et al. (2002) shows how to combine models of separate walls into one single RC equivalent model which allows further simplification. All in all, it is then possible to combine a few RC models as branches to simulate heat dynamics in a whole building (Wang and Xu, 2006b). Let us also note that upon considering a single branch, the order of the RC model relates to its number of thermal capacities. The order of the model is therefore also the number of time constants taken into account in the model.

A separate branch can also model with a linear expression the heat losses through ventilation and infiltration. As visible in Equation 1.8, the sensible heat load in the overall heat balance is function of the temperature difference between indoors and outdoors (ASHRAE, 2013) :

$$\Phi_s(t) = Q(t)\rho c_p \Delta T(t) = \frac{1}{R_{vent}(t)}\Delta T(t) \tag{1.8}$$

where $\Phi_s(t)$ is the sensible heat load (W)

$Q(t)$ the airflow rate (m^3/s)

ρ the air density (kg/m^3)

c_p the specific heat capacity of air ($J/(kg \cdot K)$)

$\Delta T(t)$ the temperature difference between indoors and outdoors (K)

(t) marks the time dependency

Properties ρ, c_p are constant, whereas Q is driven by wind pressure, stack effect and mechanical air change and is therefore time dependant.

Multi-branched RC models, based on concatenations of 2^{nd} and 3^{d} order RC models such as in Wang and Xu (2006b) and Naveros and Ghiaus (2015) have admittedly proven to be satisfactory on basis of their prediction accuracy, i.e. as a solution to forward modelling, but in inverse modelling applications and also in some forward modelling applications, simpler models has been used to describe the heat dynamics at building scale (Madsen and Holst (1995), Madsen et al. (2010), Bacher and Madsen (2011), Andersen et al. (2014), Madsen et al. (2016), Madsen et al. (2015), De Coninck et al. (2016), Thébault and Bouchié (2018), Brastein et al. (2018)).

To keep it short, let us only detail how the 2^{nd} order model of Figure 1.3 can serve for heat dynamics of the whole building (Brastein et al., 2018; De Coninck et al., 2016) and how the model is written in a state space expression.

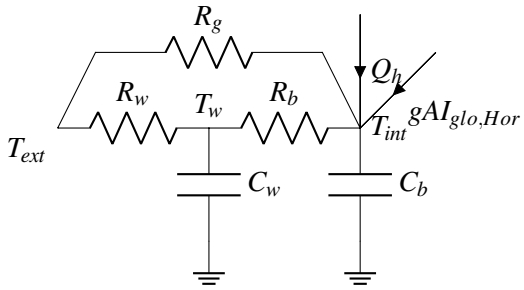


Figure 1.3 – A second order RC model (such as in Brastein et al. (2018))

Very basically, this model has a $2R1C$ branch for the overall conduction, surface convective and radiative heat transfers of walls, all walls aggregated, as well as a parallel branch with a single resistance, aiming at modelling heat flow through parts of the envelope that are not included in variations of node T_w. Brastein et al. (2018) argues that resistance R_g represents an aggregation of heat transfers through windows and low thermal inertia walls. It could be argued that it also represents infiltration and ventilation losses.

A model such as in Figure 1.3 can be easily translated into a state space representation, where must it be recalled, the derivatives of the time dependent states of the system are function of the states themselves. Equation 1.9 shows such representation of the second order model previously shown as example. In this form, it can be considered as a white box model, in the sense carried in Kramer et al. (2012) as it is purely physics-driven. Some authors however add a stochastic term (ω_w, ω_i) to it and form a grey-box model that accounts better for model uncertainties, such as in Equation 1.10, where σ is the covariance matrix of the stochastic term.

$$\begin{bmatrix} \dot{T}_w \\ \dot{T}_{int} \end{bmatrix} = \begin{bmatrix} -\frac{1}{C_w}(\frac{1}{R_w}+\frac{1}{R_b}) & \frac{1}{C_w R_b} \\ \frac{1}{C_b R_b} & -\frac{1}{C_b}(\frac{1}{R_b}+\frac{1}{R_g}) \end{bmatrix} \begin{bmatrix} T_w \\ T_{int} \end{bmatrix} + \begin{bmatrix} \frac{1}{C_w R_w} & 0 & 0 \\ \frac{1}{C_b R_g} & \frac{gA}{C_b} & \frac{1}{C_b} \end{bmatrix} \begin{bmatrix} T_{ext} \\ I_{glo,Hor} \\ Q_h \end{bmatrix} \tag{1.9}$$

$$\begin{bmatrix} \dot{T}_w \\ \dot{T}_{int} \end{bmatrix} = \begin{bmatrix} -\frac{1}{C_w}(\frac{1}{R_w}+\frac{1}{R_b}) & \frac{1}{C_w R_b} \\ \frac{1}{C_b R_b} & -\frac{1}{C_b}(\frac{1}{R_b}+\frac{1}{R_g}) \end{bmatrix} \begin{bmatrix} T_w \\ T_{int} \end{bmatrix} + \begin{bmatrix} \frac{1}{C_w R_w} & 0 & 0 \\ \frac{1}{C_b R_g} & \frac{gA}{C_b} & \frac{1}{C_b} \end{bmatrix} \begin{bmatrix} T_{ext} \\ I_{glo,Hor} \\ Q_h \end{bmatrix} + \sigma d\omega \tag{1.10}$$

Let us also mention that such representation assumes that variables T_w, T_i, Q_h,$I_{glo,Hor}$ and T_{ext} are time dependent which means that the model in Equation 1.9 is written a state space model in continuous time Madsen et al. (2010). Further using a discrete form is also possible, as well as a transfer function form (see Madsen et al. (2010); Rouchier et al. (2018)), but as underlined by

Madsen et al. (2010), physical interpretation of the parameters becomes limited, measurements must then necessarily be equidistant and for the latter, noise is no longer decomposed into system and measurement noises.

As for physical interpretation of RC models, an overall thermal resistance, inverse of the heat transfer coefficient (HTC), is easily obtained and has been widely used in literature, with for example Madsen et al. (2016); Thébault and Bouchié (2018). Resistances add up when in the same branch ($R_{eq} = R_w + R_b$), inverse resistances add up for parallel branches ($1/R_{overall} = 1/R_{eq} + 1/R_g$).

Whether each parameter can be separately interpreted at building scale is yet to be determined, let alone in a non intrusive framework.

1.5 Algorithmic and numerical tools for solving the inverse problem

In the overall process of estimating the thermal performance of a building in a non intrusive framework (see Figure 1.1), let us now suppose that data has been collected and that a model structure $\mathscr{M}(.)$ has been chosen. Left to do is to infer the actual values θ^* of the model parameters θ from the observations y.

Aside from very simple inverse problems, the parameters cannot be analytically defined as a function of the data and the model themselves. Instead, it is necessary to find a numerical approximation $\hat{\theta}$ of the parameters. Parameter estimation, or model calibration, consists in finding the best estimate $\hat{\theta}$ through a numerical routine in the hope that $\hat{\theta}$ will be close enough to the actual values θ^* (Muñoz-Tamayo et al., 2018; Walter and Pronzato, 1994).

Very shortly, the general idea is the following: if the model appropriately describes the system, then a forward simulation of the model with the actual parameter values $\mathscr{M}(\theta^*)$ is identical to the observed data, up to the noise measurement. *Model calibration* is then just the matter of numerically estimating a set of parameters $\hat{\theta}$ which will match the observed data (Walter and Pronzato, 1994).

Strictly speaking, according to Tarantola (2005), finding *one* set of parameters for the model to fit the measured data is not rigorous. What is actually desired is to infer all plausible sets of parameters, given the current *state of information*. This state of information is the sum of physical expectations, prior information on the parameters and information gained from the measured data. Whatever the tools used for the numerical estimation of all plausible sets of parameters, solving an inverse problem is then a matter of refining this state of information in the form of uncertainty intervals or in a broader sense of probability distributions.

This section describes the two approaches to model calibration, the frequentist and the Bayesian one, and discusses tools associated to both approaches, to be applied in the next chapters of this work.

1.5.1 Frequentist approach

The frequentist approach considers that parameter estimation will asymptotically converge towards the true values θ^* as the number of observations grow (Bayarri and Berger, 2004; Betancourt, 2018). To paraphrase Gustafson (2015), this means that given a model that perfectly characterizes the system, its parameters will converge with certainty to the correct target values θ^* at a $\sqrt{n}$ rate, $\sqrt{n}$ being proportional to the width of the uncertainty band. So in theory, the larger the dataset, the narrower the estimation interval, hence the closer to the correct target values.

The assumption of asymptotic behaviour also means that refining the state of information will

be done from the data only. Physical expectations on the parameter values or prior information bear no pertinent information compared to the observed data. Model calibration in a frequentist approach consists therefore in determining the probability density of the set of parameters θ from the data y only: $p(\theta|y)$. This measure is called the likelihood.

Given the asymptotic behaviour of the parameter estimation, the likelihood is often considered to have a Gaussian shape (Mosegaard and Tarantola, 2002). The calibration procedure is then much simplified as it suffices to determine the only point around which the likelihood is maximum. Variances and covariances are then approximated thanks to the Gaussian-shape assumption. Model calibration is therefore simplified to a simple maximization numerical problem or to a minimization problem when the negative logarithmic likelihood is considered, which is numerically simpler to solve.

As stressed out by Mosegaard and Tarantola (2002); Tarantola (2005), the estimate θ_{ML} of the point of maximum likelihood is not the point where the probability density $f(.)$ is maximal. Instead, maximum likelihood estimation θ_{ML} is the maximum of $F(\theta) = \frac{f(\theta)}{v(\theta)}$ with $v(.)$ a volumetric element in the parameter space (Tarantola, 2005). This definition is independent of the probability density and thus independent of a change of variable, like a change in the order of magnitude.

Among many existing methods, see reviews by Scales et al. (1992), Walter and Pronzato (1994) or Tarantola (2005), steepest-descent methods are proven to be numerically effective (Nocedal and Wright, 2006). Starting from θ_k at iteration k and given a certain radius α around θ_k, the idea is to find the direction within said radius that gives the largest change in the likelihood. Finding the maximum of $F(.)$ therefore means finding the direction around θ_k that induces the largest growth in $F(\theta)$. The algorithm continues until no radius however small produces any improvement on $F(\theta)$, which is where θ_{ML} is found. To avoid flat regions upon initialisation randomly far from the optimum, it is best to minimize $-log(F(.))$ than to maximize $F(.)$ as it transforms a Gaussian-shaped objective into a nicer paraboloid shaped objective (Mosegaard and Tarantola, 2002).

Descent methods are proven to be effective because they use the derivatives of around θ_k to find the steepest direction. The first and second derivatives may be given analytically as in the Newton methods, or may be numerically estimated as in the quasi-Newton methods. The BFGS algorithm, standing for Broyden, Fletcher, Goldfarb, and Shanno its discoverers, is a popular quasi-Newton method Nocedal and Wright (2006) when the derivates to the second order of the objective function $F(.)$ are not known. The BFGS algorithm has in this context superlinear convergence rate, i.e. faster than a linear convergence, which is a desirable property for finding in a reasonable number of iterations an acceptable estimation of the optimum. The BFGS algorithm is implemented in Matlab, R and Python. For the latter, library CTSM-R (Juhl et al., 2013) use a BFGS algorithm as well as library PySIP for RC models structures estimation (Raillon et al., 2019) that uses the SciPy library (Virtanen et al., 2020) for its BFGS algorithm in the *optimize.minimize* function.

As this work will henceforth use PySIP for RC model calibration, let us briefly describe the outputs of a model calibration and draw attention to the those that need a careful review before interpretation. PySIP produces 3 elements after calibration: a summary (Table 1.2), a covariance matrix (Table 1.3) and the SciPy output of the *optimize.minimize* function (Table 1.4).

The summary, as in Table 1.2, recalls the maximum likelihood estimator (column θ) for each parameter, associated to its standard deviation (column $\sigma(\theta)$). A p-value is given too and represents the significance of the estimated value compared to a null hypothesis where the parameter would equal 0. If the p-value is lower than a certain significance value, usually 0.05, then the estimation is considered significantly different than 0. In the example given in 1.2, the p-values are all extremely close to 0, except for parameter A_w which has a p-value around 0.0085. The estimation is still significant, but the standard deviation is indeed quite large compared to

the estimate, which explains why the significance is less strong than for the other parameters. Finally, $|g(\eta)|$ are the absolute values of the Jacobian at θ_{ML} and are algorithm related. $|dpen(\theta)|$ is related to the closeness of the estimated parameter to the optional bounds: if the value is high, the algorithms tends to search very close to the fixed bounds, which means that there is either a problem with the model or with the bounds.

The covariance matrix, as in Table 1.3, relates to the covariance between two parameters. Covariances vary on the interval $[-1, 1]$. Large covariances between θ_i and θ_j mean that around θ_{ML}, a small variation in parameter θ_i is also a small variation in θ_j without significant variation in the likelihood. The closer their covariance to 1 or to -1, the larger the correlation. On the contrary, covariances around 0 mean that around θ_{ML}, each parameter may vary and independently provoke variations in the likelihood. Chapter 2 will clarify the relationship between covariance and unicity of the estimation θ_{ML}.

Finally, Table 1.4 shows the SciPy output for the minimization itself. Among the various informations, one should verify that the optimization is successful (through *success: True* and through the message *'Optimization terminated successfully.'*), which is also visible in the number of iterations *nit*, which if lower than say 15 would be a suspicious output. Noteworthy, output *fun* gives the value of the negative log-likelihood.

As a few words of conclusion, the frequentist approach presents the undeniable advantage of being fast and already implemented in numerous available numerical tools.

It however relies on the major assumption that the number of observations suffice for the asymptotic behaviour to hold and therefore on the assumption that the parameters have a Gaussian-shaped distribution. Simplified models that have intrinsic error and low informative data might violate the asymptotic behaviour and the Gaussian-shaped distribution would not hold any more. This statement calls for an actual estimation of the posterior distribution by a Bayesian approach.

	θ	$\sigma(\theta)$	p-value	$\lvert g(\eta)\rvert$	$\lvert dpen(\theta)\rvert$
R_o	2.306e-3	9.929e-5	0.000e+00	2.819e-6	1.881e-17
R_i	5.921e-4	8.725e-6	0.000e+00	1.655e-5	2.853e-18
C_w	2.474e+7	2.047e+6	0.000e+00	3.881e-6	1.634e-17
C_i	1.671e+6	2.394e+4	0.000e+00	1.605e-5	3.581e-19
A_w	3.594e+00	1.357e+00	8.456e-3	8.389e-8	7.742e-18
$sigw_w$	2.465e-3	2.312e-4	0.000e+00	2.296e-6	1.646e-17
sig_v	4.268e-2	2.239e-3	0.000e+00	3.917e-6	5.489e-18
x_{0w}	2.610e+1	7.300e-2	0.000e+00	1.368e-5	1.467e-17

Table 1.2 – Illustration of a PySIP output: summary of frequentist model calibration

	R_i	C_w	C_i	A_w	$sigw_w$	sig_v	$x0_w$
R_o	- 0.1080	- 0.3861	- 0.1325	- 0.5512	0.0804	- 0.0221	0.0625
R_i		0.1123	0.0782	- 0.0102	0.1475	- 0.0980	- 0.0098
C_w			0.3520	0.5161	- 0.2751	0.0342	- 0.0169
C_i				0.1895	- 0.3360	0.2304	0.0246
A_w					- 0.1988	0.0515	- 0.1276
$sigw_w$						- 0.1082	0.0473
sig_v							0.0145

Table 1.3 – Illustration of a PySIP output: covariance of the parameters estimation

fun:	-468.7084807976759
hess_inv:	array([[1.85436099e-03, -6.85175864e-05, -1.37612189e-03, ...]])
jac:	array([-2.81890171e-06, -1.65452956e-05, -3.88149529e-06, ...])
message:	'Optimization terminated successfully.'
nfev:	44
nit:	33
njev:	44
status:	0
success:	True
x:	array([0.83542128, 1.77846407, 0.90575485, 2.81607462, 1.27926348, ...])

Table 1.4 – Illustration of a PySIP output: output of the SciPy minimize function

1.5.2 Bayesian approach

Tarantola (2005) states that the most natural solution to an inverse problem is to look at the overall *state of information* of the parameters θ: the distribution of the parameter θ depends on the data y and on any prior knowledge $p(\theta)$ on the plausible values for θ. The question is to determine the distribution of θ given the collected data y, i.e. to determine its posterior distribution $p(\theta|y)$. As in most inverse problems, the posterior distribution is not analytically computable and it needs to be estimated from a numerical algorithm. This section will show the principles of such an estimation, good practices and tools to perform it.

1.5.2.1 Principles of the Bayesian inference

The philosophy behind Bayesian inference is to consider the unknown as a probability distribution and to estimate it given some prior knowledge and some collected data. As Kaipio and Somersalo (2006) present it, the idea is that the probability distribution of a parameter represents the degree of information one has about the parameter. Taking as an example the estimation of the heat transfer coefficient of a building, the degree of information one might have is both expert knowledge on the building envelope (through construction date or visual inspection) and information gained on the energy behaviour of the building through appropriate data collection. Shortly said, there is no *best-fitting value* for parameter θ but a probability density given all available information.

In this sense, Bayesian inference is not simply finding the most probable parameters to explain some data, which would be determining the "peak" as in a frequentist approach. It is rather the entire distribution that is of interest (Kaipio and Fox, 2011; MacKay, 2005).

Let us define $p(\theta)$ the prior distribution (i.e. the prior expert knowledge) and $p(\theta|y)$ the posterior distribution, both are linked by a functional relationship, the Bayes theorem of Equation 1.11. The Bayes theorem states that the posterior distribution is proportional (meaning of character $\propto$) to the prior and to the likelihood $p(y|\theta)$.

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \propto p(y|\theta)p(\theta) \tag{1.11}$$

As mentioned earlier (Tarantola, 2005), although a functional relationship exists, the analytic solution does not exist in the framework of thermal performance estimation. The posterior distribution needs to be estimated by an ad hoc algorithm, i.e. drawing a large number of parameter values, samples, and determine their posterior probability density. Tarantola (2005) discusses available methods to do so.

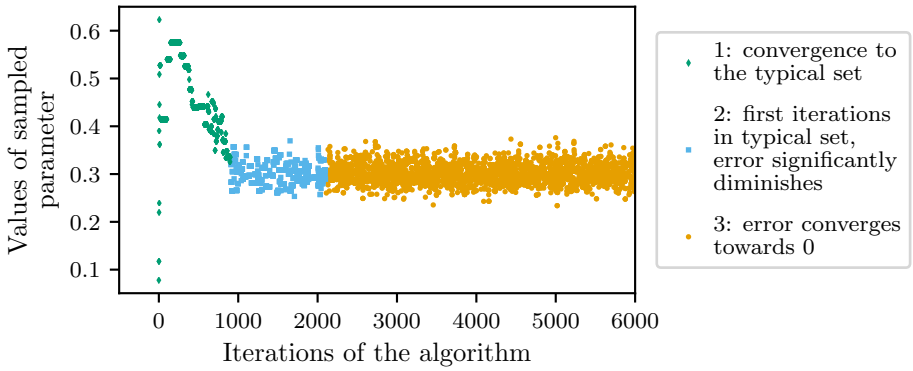
The most thorough method would be to extensively explore the parameter space by a Monte-Carlo method, which would yield posterior probabilities over the entire space. Such an extensive exploration is however not efficient, even less so as the number of parameters grows: for each additional dimension of the space, the number of points necessary to cover the space grows exponentially (see the illustration of the curse of dimensionality for parameter space exploration in Saltelli and Annoni (2010)). If covering the space is difficult, then finding the region of higher probability, quite small in comparison to the entire space, is like finding a needle on a haystack (Kaipio and Fox, 2011; Tarantola, 2005).

To remedy the curse of dimensionality, the idea is to efficiently explore the parameter space by *almost randomly* drawing samples that target the posterior distribution only: pseudo-randomly walk towards the regions of higher probability as well as not leaving such higher probability regions, while assuring the random, or pseudo-random, character of the samples.

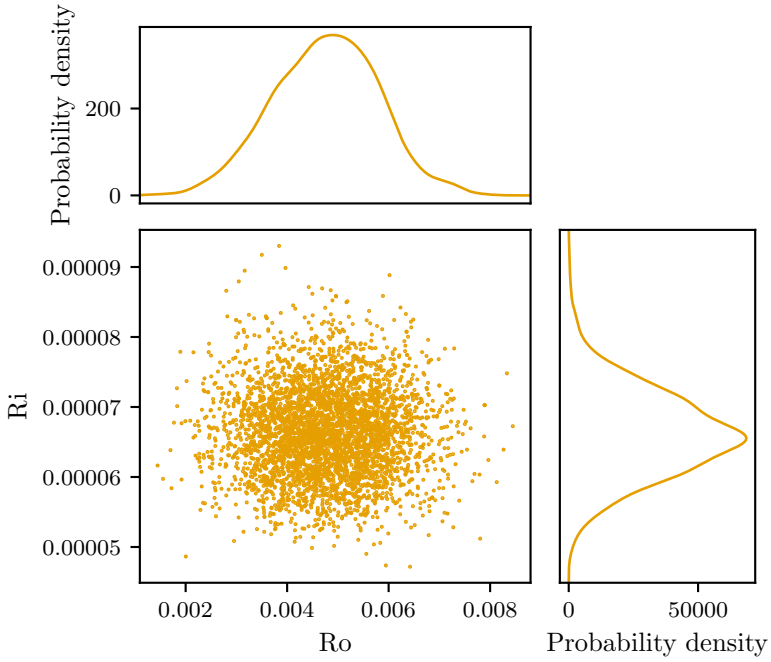
The principle goes as in Figure 1.4a. Whatever the algorithm chosen, a pseudo-random walk would first go through a phase of space exploration (green dots phase). Then, once the higher posterior probability is found (blue dots phase), the algorithm explores this particular area (orange dots phase), with more or less efficiency, depending on the chosen algorithm. From the last phase

of posterior distribution sampling, in a multi-dimensional space as in Figure 1.4b, the sampler converges to a *joint posterior distribution*: the distribution is multi-dimensional and covers all parameters. Statistics from a single particular parameter is obtained from the *marginal distribution* which are easily inferred from the multi-dimensional distribution, called the joint distribution (Figure 1.4b).

The Metropolis and Hamiltonian are popular algorithms to approximate a joint posterior distribution, with variable efficiencies (Betancourt, 2017; Tarantola, 2005).



(a) Random walk in a single parameter space



(b) Marginal distributions are inferred from the joint posterior distribution

Figure 1.4 – Illustration of posterior distribution sampling

From Tarantola (2005) and Betancourt (2017), the Metropolis-Hastings algorithm is a Markov Chain Monte-Carlo sampler. It is a random sampler (the Monte-Carlo part of the method) and has no memory from one iteration to the next (Markov Chain part) (Hastings, 1970; Metropolis et al., 1953). The principle is the following: from iteration k, instead of jumping randomly to position θ_{k+1} at iteration $k+1$, the jump is conditioned such as to go towards the higher probability area, also called *typical set*. The jump is indeed randomly accepted, with an acceptance rate depending on the probability density at the proposal jump. In the end, the history of all Markov Chain iterations will hence converge in time towards the posterior distribution, i.e. sample in detail the

typical set.

Figure 1.4a shows how the Metropolis Hastings sampler ideally performs the approximation of the posterior distribution (see also Betancourt (2017)). There is first a space exploration phase which ends in the convergence towards the typical set. At this stage, the history of the chain is useless. Statistics from it are not representative of the posterior distribution, the chain is said to be strongly biased. Phase two starts once the typical set is found, the chain stays in the area and starts to actually sample the posterior distribution. If stopped too soon in phase 2, there is however still a chance that the chain has not sampled properly the entire typical set and that the error be significant. Part three is the asymptotic behaviour of the chain: with sufficient number of iterations, the chain will converge towards the posterior distribution. The error of the statistics from such a chain asymptotically converge to 0.

Although more recent adaptations of the original algorithm have improved its efficiency (Haario et al., 2001), the Metroplis-Hastings algorithm suffers from the asymptotic nature of its convergence (Tarantola, 2005). One has no means of knowing when to stop the chain, i.e. when to consider that the posterior distribution has been successfully sampled. If stopped to soon, there is a risk that the walk has missed an important region in the parameter space. At the same time, it would be time consuming or at the least computationally costly to let a Markov Chain run for a long time. This drawback is particularly concerning when the posterior distribution is sampled from a large dimensional space (Betancourt, 2017). That comes from the acceptance of a conditional jump of the Markov Chain being made difficult exactly because there are too many dimensions that drive from the typical set instead of remaining in it.

The Hamiltonian Monte-Carlo provides a more powerful way to explore the typical set by taking advantage of its geometry. The idea behind this algorithm is that the typical set exploration would be much faster if instead of randomly explore the space around iteration k and then see whether the proposed jump is accepted, rather only explore propositional jumps that are in the typical set (Betancourt, 2017).

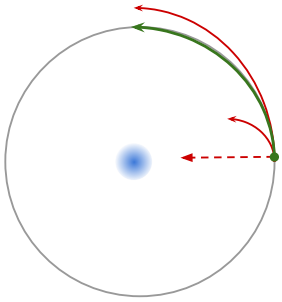
Betancourt (2017) makes a very clear explanation of the differential geometry concepts behind the theory of Hamiltonian Monte-Carlo through a gravitational metaphor as illustrated in Figure 1.5.

Instead of randomly making a step in the parameter space around θ_k, the step is driven by a purposely chosen useful metric on the state of information at iteration k. Available information is scarce, but includes the gradient of the probability density around θ_k. However, simply following the gradient would pull directly towards the higher probability density peak. In the gravitational metaphor, the sampler would directly "fall" on the heavy celestial body (in Figure 1.5a the red dotted arrow). There would be no sampling of the space around the body.

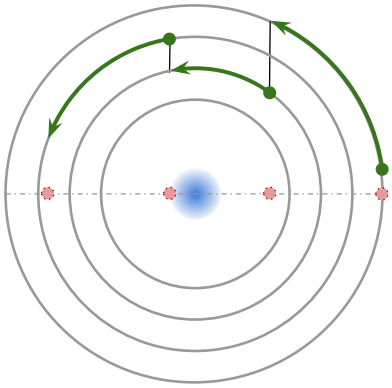
Alternatively, the only direction that would not either drift away or towards the highest probability peak is to follow an orbit around the body (grey circle in Figure 1.5a). This "orbit" keeps equal energy during its course, which means that it neither drifts away nor drifts towards the "celestial body".

But if it remains at equal energy at all time, only a part of the typical set space is sampled. So the algorithm changes once every few iterations the energy level. One could say that the sampler changes the "orbit" as illustrated in Figure 1.5b.

In the end, the history of iterations produces a chain that is a sample of the posterior distribution. Without going further into the details of the differential geometry theory behind it, it can be understood that its capacity to drive the next iteration into an informative part of the parameter space, makes the Hamiltonian sampler to produce in a reasonable number of iterations an approximation of the posterior distribution. The Hamiltonian Monte-Carlo is therefore an interesting choice when a Bayesian approach is chosen for model calibration. A Hamiltonian Monte-Carlo is implemented in python libraries pymc3 (Salvatier et al., 2016) and PySIP (Raillon et al., 2019). The latter is used in its version 0.7 for producing the results of this thesis.



(a) Hamiltonian Monte-Carlo metaphor of gravity: for the sampler to efficiently cover all the space of interest (horizontal line at sphere height), it should not follow gravitational force towards higher probability (dotted or full red), neither drift away from it , but rather follow an orbit around it (green arrow). Each "orbit" remains at equal energy (Betancourt, 2017).



(b) Hamiltonian Monte-Carlo leaps of energy: sampler leaps on energy level E_k and follows a determinist path along E_k, projects back on the horizontal line (target), then leaps again to energy level E_{k+1}, projects back again, leaps again to another E_{k+2} etc... In the end, it produces a random chain on the target distribution (Betancourt, 2017).

Figure 1.5 – Illustration of the Hamiltonian Monte-Carlo sampler

1.5.2.2 Good practice for sampling posterior distributions

Similarly to a frequentist approach, a Bayesian approach comes with necessary verifications before inferring any information from the posterior distribution samples. This section reviews some of the verifications in the light of the use of python libraries PySIP (Raillon et al., 2019) and ArviZ (Kumar et al., 2019).

First, whatever the sampler chosen for approximating the posterior distribution, the first iterations must be discarded as they will bias the results. Such a feature is included in library PySIP, where the first 1000 iterations are discarded by default.

Then, good practice demands that at least 4 chains be sampled independently, as shown in Figure 1.6. Each chain is initialized in different points of the parameter space. If all 4 approximations are significantly similar, one may consider that the posterior distribution has been successfully approximated. Again, this feature is included in PySIP, which samples 4 chains by default. The similarity of all chains may be measured by a Gelman-Rubin statistic, which is automatically calculated in the PySIP output summary. The closer the metric to 1, the more similar the chains. One may consider as satisfactory a Gelman-Rubin metric below 1.05.

In parallel, one needs to consider possible auto-correlation in each chain. With the samplers described in the previous section, the samples are produced by a Markov Chain. This means that there is at least a 1 iteration lag autocorrelation, and sometimes more. Auto-correlation is an issue

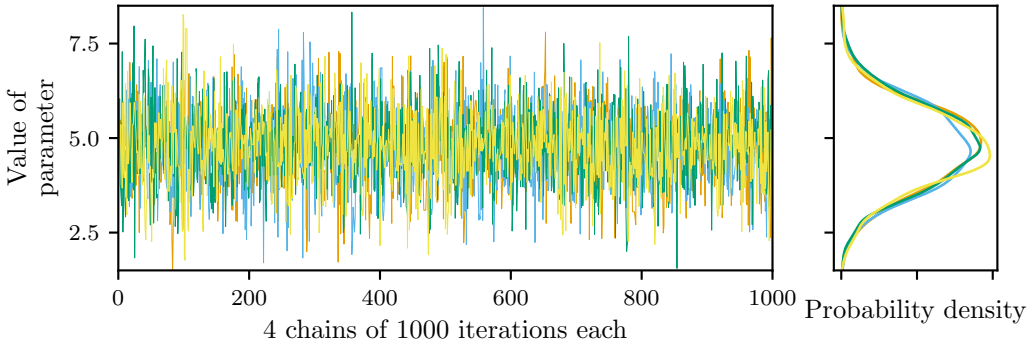


Figure 1.6 – Good practice in Bayesian calibration: 4 independent chains are drawn for repeatability check (1000 iterations left after discarding the first 1000). Gelman Rubin statistic: 1.0021

as a representative sampling needs independent draws.

A way to verify low auto-correlation is to calculate the average effective sample size (aESS). The aESS represents the number of samples that bear the same amount of information as the total number of samples considered (Vehtari et al., 2019). Given say 1000 draws, a significantly lower aESS would mean that there are only few samples that are independent, i.e. there were large auto-correlations in the chain. The aESS may also be larger than the number of actual draws in the chain, which is stricto sensu not possible but comes from the fact that it is an approximated metric. It can be implied in that case that there is insignificant auto-correlation.

To push the ESS verification further, it is also possible with ArviZ to plot the ESS throughout quantiles of the posterior distribution (Vehtari et al., 2019). The effective sample size is then calculated for small quantiles of the posterior distribution. Ideally, all quantiles, tails included, should have significant ESS. Poorly sampled quantiles will therefore be visible as low ESS. Figure 1.7b illustrates the estimated ESS for 20 quantiles of the same illustrative posterior distribution. All values are above the 400 base value for significance.

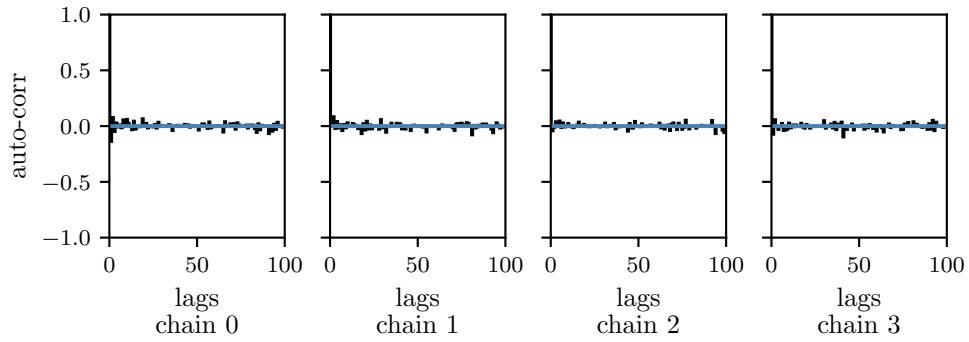
Finally, Betancourt (2016) raises the issue of low performing Hamiltonian chains: if the chain does not sample on enough energy levels ("orbits" in the gravitational metaphor in Figure 1.5), it means that the exploration is slow and that there are certainly large auto-correlations in the chain. Such a chain very probably produces biased estimates.

Without going further into detail, Betancourt (2016) presents a graphical verification method and an ad hoc indicator for poor energy sampling. As shown in Figure 1.8, the graphical verification consists in comparing the *energy distribution given q* ($\Pi_{E|q}$) to the *marginal energy distribution* Π_E. Variable q here is the variable in the target distribution, i.e. distribution to sample. In other words, given a value of the target distribution, a proper coverage of the space would give the same energy distribution than the totality of the energy levels covered. If different, it means that the target distribution has not been properly sampled and that additional samples are necessary. The ArviZ library proposes this graphical comparison in its base plot functions (as in Figure 1.8).

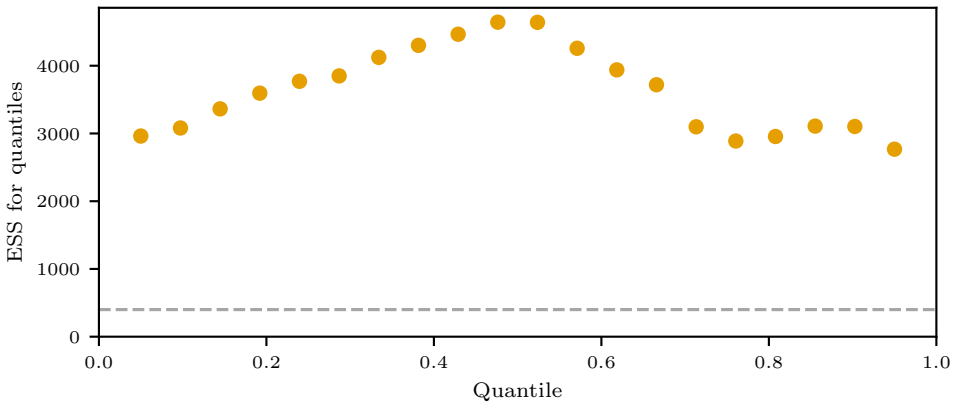
Betancourt (2016) then suggests a numerical criterion that would also represent poor exploration of the energy levels. Based on the idea of a Bayesian Fraction of Missing Information (BFMI), Betancourt (2016) gives an estimate shown in Equation 1.12. $\widehat{BFMI} \to 1$ means excellent coverage of the energy levels, $\widehat{BFMI} \to 0$ means on the contrary poor sampling.

$$BFMI \approx \widehat{BFMI} = \frac{\sum_{n=1}^{N}(E_n - E_{n-1})^2}{\sum_{n=0}^{N}(E_n - \bar{E})^2} \tag{1.12}$$

Finally, it goes almost without saying, good practice relies also on appropriate tuning of



(a) Verification of the auto-correlation of the Markov chain. Out of 4 chains of 1000 samples, aESS estimated around 4477 *meaning insignificant auto-correlation.*



(b) Small quantiles - ESS. Poorly sampled quantiles would yield low ESS. A grey dotted line serves as minimum significant value.

Figure 1.7 – Good practice in Bayesian calibration: illustration of low auto-correlation of the chains and sufficient effective sample size by quantiles

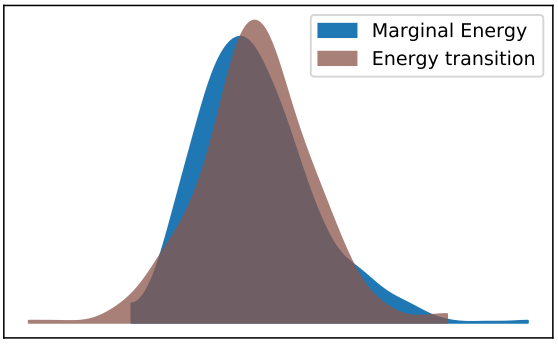


Figure 1.8 – Good practice in Bayesian inference: energy transition given q and marginal energy distributions should be close. The associated criterion BFMI should be close to 1 (here BFMI = 0.98*) (Betancourt, 2016)*

the algorithms, in particular when the aforementioned verifications uncover suspicious space coverage. This point is however not the core of the present work and it is trusted that the default parametrisation of the algorithms in the used libraries is sufficient. Chain verification is still necessary.

1.5.2.3 Inferring results from samples of a posterior distribution

Once the basic verifications have been performed on the Markov chains, one has reasonable certainty that the model space has been correctly sampled. Results may now be inferred from the samples. Figure 1.9 shows for an illustrative purpose a marginal posterior distribution. The samples drawn allow to build an estimation of the posterior density function.

Although the entire distribution is of interest, it can also shortly be described by its modal (value with highest density), mean or median value associated to a *credible* interval (Kruschke et al., 2012).

When the posterior density is normally distributed, the mean, the median and the modal values will be almost identical, as can be seen in Figure 1.9. Only when the posterior distribution is skewed will there be a difference. The choice resolves in the objective of each study: the mean would minimize the expected squared error, the median the expected absolute error whereas the mode is the most probable value (Makowski et al., 2019).

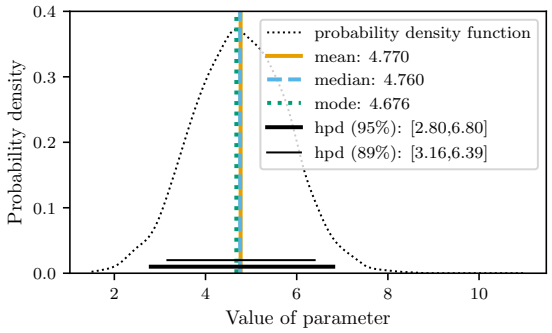


Figure 1.9 – Bayesian inference: presenting and reading results

In association, a credible interval needs to be specified. It is philosophically different from the frequentist confidence interval as it does not rely on a Gaussian assumption. As the posterior density is known, it can be defined as the region with highest probability density (hpd), regardless of the distribution's tails or on the contrary defined as the regions that leaves equal tails on both sides. Interestingly, these definitions are not per se centered around the modal value, unless the posterior is perfectly normally distributed. The hpd definition is however better suited for skewed distributions (Kruschke et al., 2012). The hpd region is shown in Figure 1.9 as the 89% and 95% hpd regions. Seeing that the *95* value from the frequentist approach is purely arbitrary, some authors argue that another number such as prime number *89* is as satisfactory as *95*.

1.6 Conclusions and orientations of the following work

Estimation of the actual thermal performance of a building envelope from on site measurement has been a focus over the past decades. Through a literature review on methods at envelope scale, this chapter has first revealed that these existing methods, although very data dependent, may achieve thermal characterisation with more or less accuracy. Controlled conditions, in which the

heating power is scheduled and provides a rich signal, seem to have the most accurate results. In non intrusive conditions, there is no possibility to use such heating power schedules.

Data driven estimation of the thermal estimation from poorly informative data constitutes therefore an ill-posed inverse problem. This means that the data might be insufficient to provide a solution, i.e. estimates of the quantities of interest, and that the solution might be highly sensitive to errors in the data. As such, the model chosen for data exploitation might be decisive for providing estimates with bound uncertainties. Among all available models, physics-driven rather than data-driven models should be chosen, as their parameters are meant to be physical. As such, stochastic RC models have the advantage of being physics-driven while not having too many parameters, which will avoid overfitting. In addition, RC models in existing methods seem to achieve satisfactory thermal characterisation in relative short duration of experiment. They indeed advantageously benefit from the dynamic nature of the data as they model both thermal resistances and capacities. Stochastic RC models have however, to the extent of our knowledge, not been applied to exploit data from non intrusive measurements.

With a given set of measurements and a model, the numerical estimation itself of the stochastic model's parameters can be performed in a frequentist or a Bayesian approach. The frequentist approach is faster in providing estimates, but relies on the assumption that the estimated parameters are normally distributed. On the contrary, the Bayesian approach actually samples the posterior distribution of the parameters, as a combination of prior expert information and data-inferred information but at the cost of much longer estimations.

In a nutshell, chapter 1 has proposed models and tools to solve the ill-posed problem of thermal characterisation of a building envelope from poorly informative data. The overall objective remains however to physically interpret the estimates. Inference from poorly informative data provides however no guarantee that the estimates exist nor that they are unique, let alone interpretable. This issue, called identifiability, is therefore thoroughly developed in chapter 2, where structural conditions for RC modelling as well as data-related conditions are presented.

2

Identifiability and interpretability

Identifiability, which is the property of unicity and existence of a solution to an inverse problem, is necessary upon estimation of the thermal performance of a building from any data. Identifiability is first a structural property, in the sense that the structure of the model must mathematically allow for a unique solution. Identifiability is then also a matter of sufficiently informative data. This chapter presents methods to check for structural identifiability of RC models as well as a good practice workflow for practical identifiability. Finally, the chapter will clarify why physical interpretability of parameters cannot be directly derived from identifiability.

Contents

2.1 Introduction

Thermal performance estimation from non intrusive data is a particularly ill posed inverse problem. A non intrusive framework induces lower quality and lower quantity on the data as well as possible bias in their values. This means not only that the collected data makes the numerical estimation of the quantity of interest difficult, but also that the existence of a unique solution is questioned.

The concept of identifiability of a set of parameters precisely relates to this issue. Roughly speaking, identifiability means that it is possible to determine parameter estimates (Glover and Willems, 1974). Conditions for parameter identifiability of RC models in a non intrusive framework need therefore to be investigated.

Starting with Reiersol (1950), the concept of identifiability of a parameter has been formally defined: a parameter is identifiable if it can uniquely be determined from the observations. Walter and Pronzato (1994) define identifiability of parameter $\hat{\theta}$ as a sort of proof by contradiction, as shown in Equation 2.1. The idea goes as follows: let us suppose there were a different set θ yielding a model behaviour $\mathscr{M}(\theta)$ identical to $\mathscr{M}(\hat{\theta})$. If it is then found that θ must actually equal $\hat{\theta}$ i.e. there is no other possibility for such supposition to be true than for both sets θ and $\hat{\theta}$ to be equal, then it means that θ is unique, thus identifiable.

$$\begin{aligned} &\text{Parameter } \theta_i \in \theta \text{ is globally identifiable if for almost all } \theta^* \in \mathbb{P}, \\ &\mathscr{M}(\hat{\theta}) = \mathscr{M}(\theta^*) \quad \Rightarrow \quad \hat{\theta}_i = \theta_i^* \end{aligned} \tag{2.1}$$

This definition implies that if found identifiable, parameter estimates $\hat{\theta}$ equal the actual quantities of interest θ^* of the system under study, **if** the following assumptions hold (Walter and Pronzato, 1994) (with the notations defined in section 1.3.1):

- the system $\mathfrak{S}$ and the model $\mathscr{M}(.)$ "have identical structures (there is no characterization error)",
- the data is noise free,
- and the observed quantities u and y "can be chosen at will", in quantity and quality.

As stated in Jacquez and Greif (1985), the identifiability property of Equation 2.1 relates to the theoretical part of the issue of unicity of the solution, as it relies on noise free data. It is then a strictly mathematical problem, necessary but not sufficient.

In building physics applications, it is part of a larger problem. Jacquez and Greif (1985) state that 1) the scientist first tackles the issue of determining the model structure best describing the system and 2) that parameters may be identifiable but that they may be poorly numerically estimable from the data.

Later on, this nuance in the identifiability property has been semantically developed by Raue et al. (2009) among others, as structural (or a priori) identifiability and practical identifiability (or a posteriori). Raue et al. (2009) underline that a structurally identifiable parameter may not be practically identifiable. Raue et al. (2009) state that practical non identifiability "arises frequently if amount and quality of experimental data is insufficient". Poorly informative data will therefore will significantly influence the practical identifiability of the parameters of interest in this work.

Identifiability, whether it be structural or practical, appears to be necessarily verification when solving an inverse problem, all the more with poorly informative data.

This section therefore proposes to first go in depth into the structural identifiability concept in Section 2.2 in order to establish a list of structurally identifiable RC models, adequate for thermal characterisation and which will be used later on. The practical identifiability concept is then

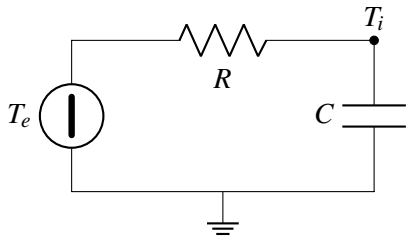
discussed in Section 2.3, with a focus on the differences brought by the frequentist and Bayesian approaches. Section 2.4 finally discusses how poorly informative data collection is a particular threat to identifiability and therefore to physical interpretation of the parameter estimates. Section 2.4 also proposes a good practice workflow, particularly necessary for calibration in a poorly informative context.

2.2 Structural identifiability : a necessary condition

2.2.1 What is structural identifiability?

Structural identifiability is a purely mathematical property of the model structure and is an absolute necessary condition before applying data to a model Walter and Pronzato (1994). Without structural identifiability, parameter estimation from data becomes intractable as an infinity of parameters produce the same output from a given input.

Figure 2.1b illustrates how an infinity of parameters may yield identical likelihoods. A T_iR model, also called $1C1R$ in free floating conditions described in Model 2.1, from 12 h night data (i.e. no solar input), is evaluated and its negative log-likelihood shown. An infinity of $R \times C$ values yields identical log-likelihoods, visible as unbound identically coloured likelihood regions.



$$\begin{cases} \dot{T}_i &= -\frac{1}{CR}T_i + \frac{1}{CR}T_e \\ y &= T_i \end{cases}$$

Model 2.1 – First order RC model T_iR in free floating temperatures conditions

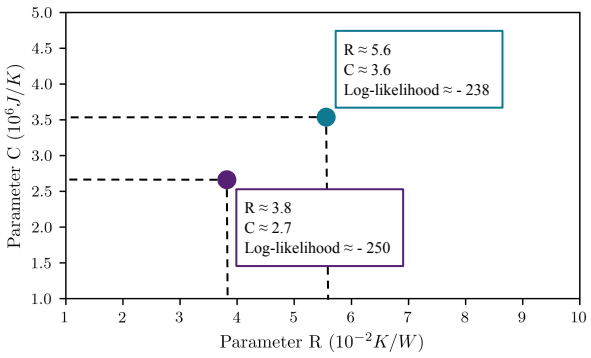
More formally, structural identifiability has first been introduced by Bellman and Åström (1970) and has since then been widely adressed (Grewal and Glover, 1976; Ljung and Glad, 1994; Vajda, 1983; Walter and Pronzato, 1994). For clarity, let us from now on use the notations from Thomassin (2005) and Walter and Pronzato (1994).

In agreement with the definitions in section 1.3.1, a model structure $\mathscr{M}(.)$ is constituted of a set of equations describing a system $\mathfrak{S}$, here the building envelope. This set of equations relies on a set of parameters θ to the time-dependent states $x(t)$. Some, or all, states are observed, they are noted $y(t)$. Model structure $\mathscr{M}(.)$ can then be written as in Equation 2.2:

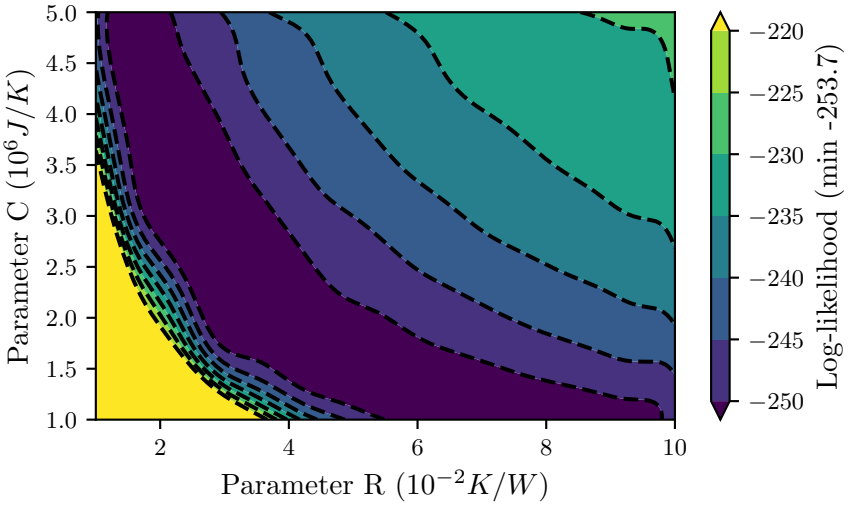
$$\begin{cases} \dot{x}(t) = f(x,\theta,t,u), \\ y(t) = g(x,\theta,t,u) \end{cases} \tag{2.2}$$

From this model structural definition, it is possible to define the input-output relation application F_θ, shown in Equation 2.3 :

$$\begin{cases} y(t) &= F_\theta(u(t),x_0) \end{cases} \tag{2.3}$$



(a) Principle of sampling the parameter space: each point represents a set (R,C) which enables to evaluate the negative log-likelihood of the RC model thus created. Different points yield variable likelihoods, rendered as a color.



(b) Illustration of structural non identifiability of a first order RC model T_iR in free floating temperatures: the darker the color the better the fit to the data. However, an infinity of (R,C) sets yield the exact same negative log-likelihoods. It is impossible to determine the actual values of the parameters as an infinity are equally probable.

Figure 2.1 – Illustration of the effect of structural non identifiability on the likelihood

Structural identifiability can then be formally defined: there exists only a single and unique set of parameters θ to link a certain output from a given input, for any θ and for any input-output set. In other terms, if two sets of parameters in the same model structure yield the same input-output relation $F(u(t),x_0)$, then both sets must actually be equal. Equation 2.4 now states this definition of structural identifiability:

$$\forall\theta, \qquad F_\theta(u(t),x_0) = F_{\theta^*}(u(t),x_0) \quad \Rightarrow \quad \theta = \theta^* \tag{2.4}$$

Let us stress out that structural identifiability of a model means identifiability of its parameters. This definition further implies that when calibrating a structurally identifiable model, if a set of parameters is found to fit the input and output data, it must be the unique solution to the inverse problem.

On the contrary, with a structurally *non* identifiable model, an infinity of parameter sets yield the exact same input-output relation. In practice, when calibrating a model from some input and output data, this means that an infinity of parameter sets will fit the data. With a non identifiable model structure, the parameters cannot be uniquely identified and the problem becomes numerically intractable (Bellman and Åström, 1970). At best the calibration algorithm fails and cannot converge. Worst case scenario, the algorithm provides a result that has actually no meaning. For this reason, structural identifiability needs to be checked.

2.2.2 Principle of a unique input-output expression: the exhaustive summary

Deriving structural identifiability directly from the input-output expression $F_\theta(u(t),x_0)$ is burdensome because it is defined with dependence on input, states and time variables.

The general idea to deal with this intractability is to form an expression equivalent to the input-output relation such as Equation 2.5 is true, containing thus the same information as given in the model structure $\mathscr{M}(.)$ but eliminating time and inputs from the equations. This equivalent expression, called $\mathsf{s}(\theta)$, is called an exhaustive summary (Walter and Lecourtier, 1982).

$$\forall\theta,\hat{\theta}, \quad F_\theta(u(t),x_0) = F_{\hat{\theta}}(u(t),x_0) \Leftrightarrow \mathsf{s}(\theta) = \mathsf{s}(\hat{\theta}) \tag{2.5}$$

Using the exhaustive summary facilitates the study of the structural identifiability and most of the methods use one (Walter and Lecourtier, 1982). As both formulations are equivalent, proving unicity of the parameters for the first implies unicity of the parameters of the model.

$$\forall\theta, \mathsf{s}(\theta) = \mathsf{s}(\hat{\theta}) \quad \Rightarrow \quad \theta = \hat{\theta} \tag{2.6}$$

The general principle of using an exhaustive summary is therefore common to all methods to verify structural identifiability. Each method has in addition a different way to expressing it, which will now be developed.

Structural identifiability has widely been discussed in literature and several methods are indeed available. Some are applicable on linear models only and others indifferently on linear or non linear models.

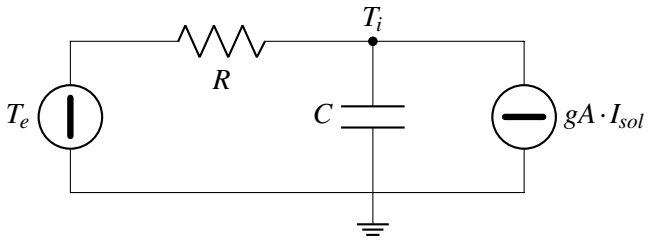
Let us from now on consider a state space model structure, as in Equation 2.7, where:

- X represents the states of the model,
- Y the observed states,
- matrices A, B, C and D contain the parameters θ describing the system. They are time-invariant.

$$\begin{cases} \dot{X}(t) &= A(\theta)\cdot X(t) + B(\theta)\cdot U(t) \\ Y(t) &= C(\theta)\cdot X(t) + D(\theta)\cdot U(t) \\ X(0) &= X_0 \end{cases} \tag{2.7}$$

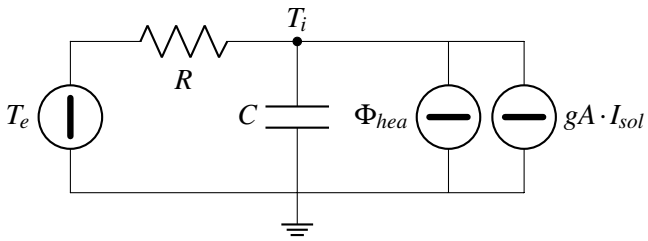
2.2.3 Some existing methods for verifying structural identifiability

This section will first lay out the common principle of all methods and detail three methods on basis of the notation from the general state space model from Equation 2.7. Then, all methods are applied to two simple RC models: 2.2 and 2.3. The code written to apply each method to both models can be found in Annex B.



$$\begin{cases} \left[\dot{T}_i\right] &= \left[-\frac{1}{CR}\right]\left[T_i\right] + \left[\frac{1}{CR} \quad \frac{gA}{C}\right]\begin{bmatrix} T_e \\ I_{\text{sol}} \end{bmatrix} \\ y &= \left[T_i\right] \end{cases}$$

Model 2.2 – First order RC model in free floating temperatures condition: T_iRA



$$\begin{cases} \left[\dot{T}_i\right] &= \left[-\frac{1}{CR}\right]\left[T_i\right] + \left[\frac{1}{CR} \quad \frac{gA}{C} \quad \frac{1}{C}\right]\begin{bmatrix} T_e \\ I_{\text{sol}} \\ \Phi_{hea} \end{bmatrix} \\ y &= \left[T_i\right] \end{cases}$$

Model 2.3 – First order RC model with measured heating input in controlled temperature condition: T_iRA

For the reader, other clear and intuitive applications in the field of population dynamics are done by Perasso (2009), in the field of automation by Thomassin (2005) and in the field of biological systems by Raue et al. (2014).

2.2.3.1 Laplace transform approach for linear models

Introduced by Bellman and Åström (1970) and also developed in Walter and Pronzato (1994) and shortly in Grewal and Glover (1976), the Laplace transform approach is based on an expression of the state space model in the Laplace domain, thus acting as the exhaustive summary.

From the general state space formulation from Equation 2.7, the proof goes as follows:

$$\begin{aligned} \dot{X}(t) &= A\cdot X(t) + B\cdot U(t) \\ sX(s) &= A\cdot X(s) + B\cdot U(s) \\ (sI-A)\cdot X(s) &= B\cdot U(s) \\ X(s) &= (sI-A)^{-1}B\cdot U(s) \end{aligned} \tag{2.8}$$

In parallel, the state space model (Equation 2.7) has some observable outputs Y:

$$Y(s) = C\cdot X(s) + D\cdot U(s) \tag{2.9}$$

which implies that $Y = C((sI-A)^{-1}BU) + DU$ and thus follows the transfer function, acting as exhaustive summary:

$$H(s,\theta) = C(sI-A)^{-1}B + D$$

Back to the definition of structural identifiability in Equation 2.4, the model is structurally identifiable if and only if:

$$H(s,\theta) = H(s,\hat{\theta}) \Rightarrow \theta = \hat{\theta} \tag{2.10}$$

Applied to Model 2.3, it follows:

$$H(s,\theta) = \begin{bmatrix} \frac{1}{RCs+1} & \frac{RgA}{RCs+1} & \frac{R}{RCs+1} \end{bmatrix} \tag{2.11}$$

and in its canonical form (i.e. the highest order term in the denominator is set to 1, the denominator is then said to be monic):

$$H(s,\theta) = \begin{bmatrix} \frac{\frac{1}{RC}}{s+\frac{1}{RC}} & \frac{\frac{gA}{C}}{s+\frac{1}{RC}} & \frac{\frac{1}{C}}{s+\frac{1}{RC}} \end{bmatrix} \tag{2.12}$$

The transfer function must be used in its canonical form as it is a unique representation of any rational function. It then allows to solve $H(s,\theta) = H(s,\hat{\theta})$ because a unique representation means that each coefficient of both numerator and denominator of the left hand side function $H(s,\theta)$ may be set as equal to the mirrored coefficients of the right hand side function $H(s,\hat{\theta})$. Therefore comes the following:

$$\begin{aligned} H(s,\theta) = H(s,\hat{\theta}) &\iff \begin{bmatrix} \frac{\frac{1}{RC}}{s+\frac{1}{RC}} & \frac{\frac{gA}{C}}{s+\frac{1}{RC}} & \frac{\frac{1}{C}}{s+\frac{1}{RC}} \end{bmatrix} = \begin{bmatrix} \frac{\frac{1}{\hat{R}\hat{C}}}{s+\frac{1}{\hat{R}\hat{C}}} & \frac{\frac{\widehat{gA}}{\hat{C}}}{s+\frac{1}{\hat{R}\hat{C}}} & \frac{\frac{1}{\hat{C}}}{s+\frac{1}{\hat{R}\hat{C}}} \end{bmatrix} \\ &\iff \begin{cases} \frac{1}{RC} = \frac{1}{\hat{R}\hat{C}} \\ \frac{gA}{C} = \frac{\widehat{gA}}{\hat{C}} \\ \frac{1}{C} = \frac{1}{\hat{C}} \end{cases} \end{aligned} \tag{2.13}$$

From there it is trivial that:

$$\begin{cases} \frac{1}{R} &= \frac{1}{\hat{R}} \\ gA &= \widehat{gA} \\ \frac{1}{C} &= \frac{1}{\hat{C}} \end{cases}$$

Hence the structural identifiability of Model 2.3.

Similarly, the transfer function of the same model in free floating conditions, Model 2.2, is:

$$H_{2.2}(s,\theta) = \begin{bmatrix} \frac{\frac{1}{RC}}{s+\frac{1}{RC}} & \frac{\frac{gA}{RC}}{s+\frac{1}{RC}} \end{bmatrix} \tag{2.14}$$

from which:

$$H_{2.2}(s,\theta) = H_{2.2}(s,\hat{\theta}) \iff \begin{cases} \frac{1}{RC} &= \frac{1}{\hat{R}\hat{C}} \\ \frac{gA}{C} &= \frac{\widehat{gA}}{\hat{C}} \end{cases} \tag{2.15}$$

which has an infinity of solutions, unless one of the coefficients is known. Model 2.2 *TiRA* in free floating temperature conditions is structurally non identifiable.

2.2.3.2 Input-output in a Taylor series expansion for linear models

The Taylor series expansion approach applied to continuous time non linear models, was developed by Pohjanpalo (1978). In discrete time and for linear state space models, it is equivalent to using the so-called Markov parameters, as detailed in Grewal and Glover (1976). The reader will find recent and easy to read applications of the Markov parameters approach in Van Doren et al. (2008) or in Agbi et al. (2012).

The approach relies on an exhaustive summary in the form of a Taylor series expansion of the model around $\hat{\theta}$. As just mentioned, for state space models in discrete time, it is equivalent to writing the model with Markov parameters. Such an expansion is by definition infinite, but Grewal and Glover (1976) has shown that it suffices to study the first m Markov coefficients to prove the structural identifiability, with m strictly greater than twice the dimension of the state space vector. The Markov parameter of order k is defined as following :

$$\begin{aligned} &\text{for } k \in [1,m], \quad s(k,\theta) = CA^{k-1}B \\ &\text{with} \quad m > 2 \cdot dim(X) \end{aligned}$$

All m Markov parameters are stacked into a vector S_m :

$$S_m(\theta) = \begin{bmatrix} s(1,\theta) & s(2,\theta) & \dots & s(m,\theta) \end{bmatrix} \tag{2.16}$$

From there, the Jacobian of the Markov parameters array S_m is calculated in θ (with p the dimension of θ) :

$$J(m) = \frac{dS_m(\theta)}{d\theta} = \begin{bmatrix} \frac{dS_m(\theta)}{d\theta_1} \\ \frac{dS_m(\theta)}{d\theta_2} \\ \dots \\ \frac{dS_m(\theta)}{d\theta_p} \end{bmatrix} = \begin{bmatrix} \frac{ds(1,\theta)}{d\theta_1} & \frac{ds(2,\theta)}{d\theta_1} & \dots & \frac{ds(1,\theta)}{d\theta_1} \\ \frac{ds(1,\theta)}{d\theta_2} & \frac{ds(2,\theta)}{d\theta_2} & \dots & \frac{ds(m,\theta)}{d\theta_2} \\ \dots & \dots & \dots & \dots \\ \frac{ds(1,\theta)}{d\theta_p} & \frac{ds(2,\theta)}{d\theta_p} & \dots & \frac{ds(m,\theta)}{d\theta_p} \end{bmatrix} \tag{2.17}$$

Then, if $rank(J(m))$ is bigger or equal to p the dimension of θ than the model is structurally identifiable.

$$rank(J(m)) \geq dim(\theta) \quad \Rightarrow \quad \text{State space model structurally identifiable}$$

Applied to Model 2.2 (T_iRA in free floating temperature conditions), which dimension of the state vector is $n = 1$, it suffices to:

- calculate the first $2 \times n + 1 = 2 \times 1 + 1 = 3$ Markov parameters;
- determine $S_3(\theta)$;
- determine the Jacobian $J(3)$.

The model is structurally identifiable if and only if the rank of $J(3)$ is greater or equal to the dimension of the parameter vector θ if $\forall \theta, rank(J(3)) \geq 3$.

Let us now proceed to the application and calculate the first 3 Markov parameters and stack them into $S_3(\theta)$:

$$S_3(\theta) = \begin{bmatrix} \frac{1}{RC} & \frac{gA}{C} & -\frac{1}{R^2C^2} & -\frac{gA}{RC^2} & \frac{1}{R^3C^3} & \frac{gA}{R^2C^3} \end{bmatrix} \tag{2.18}$$

$$J(3) = \begin{bmatrix} \frac{\mathrm{d}S_3(\theta)}{\mathrm{d}C} \\ \frac{\mathrm{d}S_3(\theta)}{\mathrm{d}R} \\ \frac{\mathrm{d}S_3(\theta)}{\mathrm{d}gA} \end{bmatrix} = \begin{bmatrix} -\frac{1}{RC^2} & -\frac{gA}{C^2} & \frac{2}{C^3R^2} & \frac{2gA}{C^3R} & \frac{-3}{C^4R^3} & \frac{-3gA}{C^4R^2} \\ -\frac{1}{CR^2} & 0 & \frac{2}{C^2R^3} & \frac{gA}{C^2R^2} & -\frac{3}{C^3R^4} & -\frac{2gA}{C^3R^3} \\ 0 & \frac{1}{C} & 0 & -\frac{1}{C^2R} & 0 & \frac{1}{C^3R^2} \end{bmatrix} \tag{2.19}$$

The Jacobian gives $rank(J(3)) = 2 < 3$ hence, again, Model 2.2, in free floating temperatures conditions, is structurally non identifiable.

As for Model 2.3, with controllable indoor temperature conditions, the proof goes as follows:

$$S_3(\theta) = \begin{bmatrix} \frac{1}{RC} & \frac{gA}{C} & \frac{1}{C} & -\frac{1}{R^2C^2} & -\frac{gA}{RC^2} & -\frac{1}{RC^2} & \frac{1}{R^3C^3} & \frac{gA}{R^2C^3} & \frac{1}{R^2C^3} \end{bmatrix} \tag{2.20}$$

$$J(3) = \begin{bmatrix} \frac{\mathrm{d}S_3(\theta)}{\mathrm{d}C} \\ \frac{\mathrm{d}S_3(\theta)}{\mathrm{d}R} \\ \frac{\mathrm{d}S_3(\theta)}{\mathrm{d}gA} \end{bmatrix} = \begin{bmatrix} -\frac{1}{RC^2} & -\frac{gA}{C^2} & -\frac{1}{C^2} & \frac{2}{R^2C^3} & \frac{2gA}{RC^3} & \frac{2}{RC^3} & -\frac{3}{R^3C^4} & -\frac{3gA}{R^2C^4} & -\frac{3}{R^2C^4} \\ -\frac{1}{R^2C} & 0 & 0 & \frac{2}{R^3C^2} & \frac{gA}{R^2C^2} & \frac{1}{R^2C^2} & -\frac{3}{R^4C^3} & -\frac{2gA}{R^3C^3} & -\frac{2}{R^3C^3} \\ 0 & \frac{1}{C} & 0 & 0 & -\frac{1}{RC^2} & 0 & 0 & \frac{1}{R^2C^3} & 0 \end{bmatrix} \tag{2.21}$$

where $rank(J(3)) = 3 = dim(\theta)$ hence structural identifiability of Model 2.3.

2.2.3.3 Exhaustive summary expressed in a differential algebra for linear and non-linear models

As also underlined by Thomassin (2005), Perasso (2009) or Raue et al. (2014), differential algebra is an advanced field of mathematics, barely accessible to applied scientists and practitioners. This is why this section will not undergo details of the methods, but just draw its main principles.

The main idea is that the input output relation, equipped with the addition, the multiplication and derivation forms what is called a differential algebra, which has interesting properties for the field of linear dynamic models (see Ritt (1950)). In particular, a pseudo-algorithm allows

to uniquely express the input-output relation in this particular algebra thus yielding an exhaustive summary. In practice, it is a sort of non-linear polynom, function of the inputs and their derivatives.

Several authors have provided algorithms or applied this idea to structural identifiability analysis see among others Ljung and Glad (1994), Sedoglavic (2002), Saccomani et al. (2003), Audoly et al. (2001), Wu et al. (2008), Xia and Moog (2003), Anguelova (2007) or Karlsson et al. (2012).

In particular, two open source algorithms will be presented here:

- the DAISY algorithm Bellu et al. (2007) : written in the REDUCE environment, it determines global structural identifiability of linear and non linear models.
- the observabilityTest for Maple Sedoglavic (2001) : it is a probabilistic algorithm that shows local structural identifiability, but is much more efficient than the DAISY algorithm for large models.

To apply the DAISY algorithm, the Models 2.2 and 2.3 need to be written in REDUCE for the DAISY package to process them. As suggested by the package documentation, one may either write each command separately directly in the REDUCE software or write all commands in a separate text file and simply call the file to be executed in the REDUCE environment. The second option is applied here and details of how to build the text file are given in Annex B.2.3.

After running the text file describing the model under study, the algorithm yields the following results : a characteristic set is calculated (lines 5 and 6), the algorithm then sets arbitrary values to parameters C, R and gA, solves the system of equations and concludes line 16 that the model is non identifiable. From line 14 can be understood that fixing one of the parameters will yield an identifiable model.

```
MODEL EQUATION(S)                                                    1
c_ := {df(ti,t)=(text - ti + ga*isol*r)/(c*r),y=ti}                   2
                                                                     3
CHARACTERISTIC SET                                                   4
aa_(1) := df(y,t)*c*r - isol*ga*r - text + y                         5
aa_(2) := - ti + y                                                   6
                                                                     7
MODEL ALGEBRAICALLY OBSERVABLE                                       8
                                                                     9
PARAMETER VALUES                                                     10
b2_ := {c=2,r=3,ga=5}                                                11
                                                                     12
MODEL PARAMETER SOLUTION(S)                                          13
g_ := {{c=(2*ga)/5,r=15/ga}}                                         14
                                                                     15
MODEL NON IDENTIFIABLE                                               16
```

As for Model 2.3, the algorithm finds it globally identifiable :

```
MODEL EQUATION(S)                                                    1
c_ := {df(ti,t)=(text - ti + qheat*r + a*isol*r)/(c*r),y=ti}          2
                                                                     3
CHARACTERISTIC SET                                                   4
aa_(1) := df(y,t)*c*r - isol*a*r - qheat*r - text + y                5
aa_(2) := - ti + y                                                   6
                                                                     7
MODEL ALGEBRAICALLY OBSERVABLE                                       8
```

```
                                                                    9
PARAMETER VALUES                                                    10
b2_ := {c=2,r=3,a=5}                                                11
                                                                    12
MODEL PARAMETER SOLUTION(S)                                         13
g_ := {{a=5,c=2,r=3}}                                               14
                                                                    15
MODEL GLOBALLY IDENTIFIABLE                                         16
```

The observabilityTest on the other hand is based on the Maple software for formal computation. The observability library is added to the current Maple library by running the following line at the beginning of each worksheet.

```
libname := libname, "path_to_dir/ObservabilityTest/maple.lib"
```

The use of it is then quite straightforward as shown below with Model 2.2:

```
F:=[-1/(C*R)*T+1/(C*R)*Text+gA/C*Isol];                             1
X:=[T];                                                             2
G:=[T];                                                             3
Theta:=[C,R,gA];                                                    4
U:=[Text,Isol];                                                     5
                                                                    6
observabilityTest(F, X, G, Theta, U)                                7
```

which yields :

```
observabilityTest: Modular computation with version 0.0             1
observabilityTest: Some informations about the system               2
observabilityTest: Nb Inputs       U  : 2                           3
observabilityTest: Nb Outputs      Y  : 1                           4
observabilityTest: Nb Variables    X  : 1                           5
observabilityTest: Nb Parameters Theta: 3                           6
observabilityTest: The computation are done modulo the following    7
    prime number 10000000000037
observabilityTest: System treatment                                 8
observabilityTest: End of system treatment 0.89e-1                  9
observabilityTest: Begining of integration process                  10
observabilityTest: For time estimation of this process, set         11
   current infolevel to 2
observabilityTest: End of integration                               12
observabilityTest: Evaluation of output system                      13
observabilityTest: End of evaluation of output system               14
ObservabilityAnalysis: The transcendence degree of k(U,Y) --> k(    15
   U,Y,X,Theta) is 1
ObservabilityAnalysis: [C R gA] are not observable.                 16
observabilityTest: Total used time .154                             17
[1, [C, R, gA], [T], 3]                                             18
```

The last and second to last lines are of interest. First, the parameters C, R and gA are not observable, meaning not identifiable. Secondly, the last line is a list of four elements : the second element shows the non observable parameters, the third the observable ones. Noteworthy is that whereas the parameters are not observable, the temperatures states are.

Applied to Model 2.3, the algorithm yields the following results : all parameters and states are observable hence structural identifiability.

```
ObservabilityAnalysis: The transcendence degree of k(U,Y) --> k(U,Y,X,Theta) is 0
observabilityTest: Total used time 0.19e-1
[0, [], [C, R, gA, T], 0]
```

Last but not least, such methods for verifying the structural identifiability can, in some very particular cases, be sensitive to the initial conditions, as has been adressed by Villaverde and Banga (2017). The reason behind it is that both methods use an augmented observability matrix and calculate its rank (augmented in the sense that the matrix assembles state variables and parameters). But it is possible that a certain combination of initial conditions influences durably the state variables (induces e.g. a state to be constantly null for example). In such case, when the rank calculation is done symbolically, the result might be misleading. Villaverde and Banga (2017) illustrates this issue with examples of biochemical networks and it is expected that this issue will almost never be encountered in a building heat transfer application, as initial conditions in such systems are rarely so particular. It should however be kept in mind as a possible exception when the calibration process proves to be numerically problematic.

2.2.3.4 Conclusion: choice of a structural identifiability method

This section has reviewed four different methods to assess the structural identifiability of a state space model.

As an illustration, all presented methods have been applied to models 2.2 and 2.3 and it has been found, regardless of the method, that the former, Model 2.2, is not structurally identifiable but that the latter, Model 2.3 is.

This difference illustrates that such first order RC model with indoor air temperature as an output needs a heating input to be structurally identifiable. Further higher order RC models will therefore always include a heating input.

Let us here stress out that providing the heating input in Model 2.3 is modelled on the hypothesis that the heating is perfectly measured and perfectly adequate for model calibration. This assumption will be later discussed.

Finally, for further tests of the structural identifiability of different RC models, the DAISY algorithm will be used as it is an easy-to-use method thanks to its implementation in the REDUCE environment.

2.2.4 Why some models for forward problems are unfit for inverse modelling

Literature has shown the wide use of RC models for forward as well as inverse modelling, see Section 1.4.2. To estimate the thermal performance of a building, one could expect detailed physics-driven RC model to perform better than simpler models, as they are more precise in their description of the heat transfers at building scale. Although it is impossible to be exhaustive, this section reviews a few of such models and shows that they are actually ***not*** structurally identifiable.

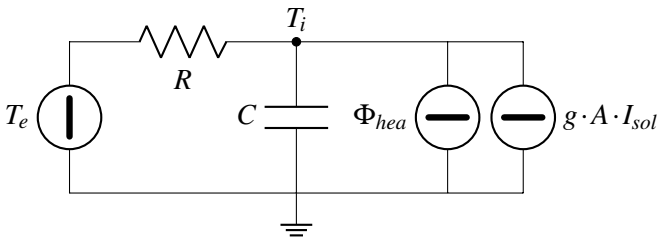
Details of the (non) structural identifiability proofs are not rehearsed here but are given in Appendix B.

Particular case of the effective solar aperture coefficient

The effective solar aperture coefficient, usually noted gA or simply A modulates the global solar irradiation input $I_{glo,sol}$ in the RC model. It can apply to only one or all nodes, when seen fit from a physical perspective. This coefficient could be seen as a product of

- A_{node} the aperture surface in m^2 and;
- a coefficient $g \in [0,1]$.

One may then consider the product $g \cdot A$ with both g and A unknown, as in model 2.4. This would however yield structural non identifiability (see annex B.3.1 for proof). However, when the product is considered as a single unknown, the model becomes identifiable.



$$\begin{cases} \left[\dot{T}_i\right] = \left[-\frac{1}{CR}\right]\left[T_i\right] + \begin{bmatrix} \frac{1}{CR} & \frac{g \cdot A}{C} & \frac{1}{C} \end{bmatrix} \begin{bmatrix} T_e \\ I_{\text{sol}} \\ \Phi_{hea} \end{bmatrix} \\ y = \left[T_i\right] \end{cases}$$

Model 2.4 – First order model $T_i RC$ $g \cdot A$ in controllable indoor temperature conditions

Parameters g or A_{node} are therefore individually not identifiable, but the product is. Let us note that it does not concern all products in a state space representation. Although parameters R and C often appear as a product, yet not always thanks to the heating input coefficient, both are individually structurally identifiable. However, product $g \cdot A$ never appears separately in the state space representation, which might be the reason why they are not separately identifiable.

Incidently, in free floating conditions such as described by model 2.2 and illustrated in Figure 2.1b, parameters R and C are separately non identifiable for the same reason. Indeed, a model with indoor temperature as output and no other input than outdoor temperature and solar irradiation displays parameters R and C always as a product.

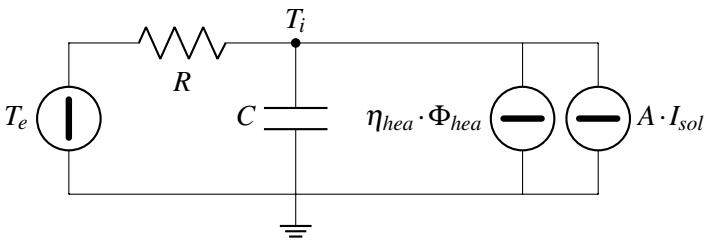
Back to a conclusion regarding the solar irradiation coefficient, gA_{node} is seen as a single parameter from now on, and noted simply A.

Taking into account a heating system efficiency coefficient

With a physics-driven model and in-situ data, it would be essential to account for the heating system efficiency. Aside from particular cases, the effective heating power delivered in the building would not be known. Rather, the measurement would either consist in gross aggregated electricity consumption or gas consumption. In both cases, there would be a multiplicative coefficient to modulate the energy power input.

Such modulation could be written as in first order model 2.5, where the heating power $\eta_{hea} \times \Phi_{hea}$ is the effective heat flux on the single node. Parameter η_{hea} is unknown and input variable Φ_{hea} is measured.

However, such model is structurally not identifiable (see code in annex B.3.2). It brings an additional unknown in comparison to the almost identical Model 2.3 which turns it into non identifiability. Second or larger order RC models with an η_{hea} term on the heating power input also yield non identifiable structures. Section 2.3 will further discuss how a Bayesian approach could obviate this obstacle.



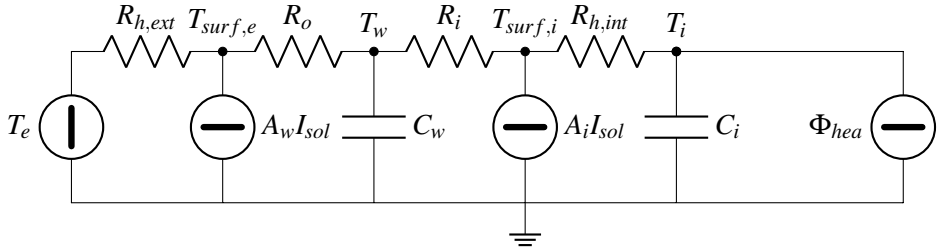
$$\begin{cases} \left[\dot{T_i}\right] &= \left[-\frac{1}{CR}\right]\left[T_i\right] + \left[\frac{1}{CR} \quad \frac{A}{C} \quad \frac{\eta_{hea}}{C}\right]\begin{bmatrix} T_e \\ I_{\text{sol}} \\ \Phi_{hea} \end{bmatrix} \\ y &= \left[T_i\right] \end{cases}$$

Model 2.5 – First order model T_iRC A in controllable indoor temperature conditions with heating system efficiency coefficient η_{hea}

Surface thermal resistances

Last example of structural non-identifiability are models that account specifically for aggregated surface convection and radiation thermal resistances, such as first order model 2.6.

Such model is also structurally non identifiable (see annex B.3.3 for the code and the model equations). In particular, the serial resistances $R_{h,ext}$ and R_w even with a solar irradiation input on $T_{surf,e}$ are not identifiable. If a surface temperature were known, or a heat flux measurement were available, it could be accounted for in the model which would become identifiable. This would however affect the former simplicity of the measurement protocol, so far with just an indoor temperature sensor.



Model 2.6 – Second order model $T_wT_iR_oR_iA_eA_i$ in controllable indoor temperature conditions with surface convecto-radiative thermal resistances $R_{h,ext}$ and $R_{h,int}$

2.2.5 Application: a set of structurally identifiable state space models

The previous section ruled out some RC models that were found to be structurally non identifiable. These type of RC models are indeed rather fit for forward modelling. To form a set of identifiable models, surface thermal resistances will be excluded, as well as heating system efficiency coefficients and particular cases of parameter products.

RC models without solar irradiation input (such as model 2.1) are not discussed in the following work as they are only representative of heat transfers at building scale at night time. Such models imply to proceed to calibrations from discontinuous datasets. This brings upfront

other specific numerical issues such as algorithm initialization and inconsistent state space initial conditions. They are therefore purposely left aside in this work.

Let us now form a set of structurally identifiable models, fit for inverse problems while hopefully physically interpretable because based on aggregate physics. All models are presented in Table 2.1 along with their inputs and outputs. The details of all models are sorted according to their order (number of thermal capacities) and follow in the next three subsections. The models' state space representation is given in its grey-box form, with a stochastic term $\sigma d\omega$, in conformity with section 1.4.2. The measurement error is noted ε, regardless of its dimension.

Let us also note that all models include a solar irradiation variable I_{sol}. Although debatable, it is from now on considered that this variable is the global horizontal solar irradiation (in W/m^2). Indeed, the non intrusive aspect of the considered experiments lead to choosing relatively easy to measure variables, which is the case of global horizontal irradiation. It could also be argued that a measure of horizontal irradiation could be corrected by trigonometry or by B-splines (Rasmussen et al., 2020) to better reflect on solar gains on each façade. Such correction has been considered outside the scope of this work but definitely is a valuable prospect. In addition, the solar aperture coefficients are not the objective of this work. Were it the goal, an in depth study of solar irradiation as input variable should be conducted and a different choice would most probably be made.

This list should not be considered as an exhaustive list of all first, second and third order RC models, but rather as a list of models to be possibly used in the following chapters.

	1^{st} order	2^{nd} order				3^{d} order	
	T_i R A_i	T_wT_i R_oR_i A_wA_i	T_wT_i $R_oR_i\ c_v$ A_wA_i	T_wT_i $R_oR_i\ R_p$ A_wA_i	T_wT_i $R_oR_i\ R_b$ A_wA_i	$T_{we}T_wT_i$ $R_oR_wR_i$ $A_eA_wA_i$	$T_wT_i\ T_b$ $R_wR_i\ R_{bi}R_{bb}$ A_wA_i
Model reference	2.3	2.7	2.8	2.9	2.10	2.11	2.12
Measured input	T_{ext} $\Phi_{heating}$ I_{sol}^{global}		T_{ext} $\Phi_{heating}$ I_{sol}^{global} Φ_{vent}	T_{ext} $\Phi_{heating}$ I_{sol}^{global}	T_{ext} $\Phi_{heating}$ I_{sol}^{global} $T_{crawl,space}$	T_{ext} $\Phi_{heating}$ I_{sol}^{global}	T_{ext} $\Phi_{heating}$ I_{sol}^{global} $T_{crawl,space}$
Measured output	T_{int}						

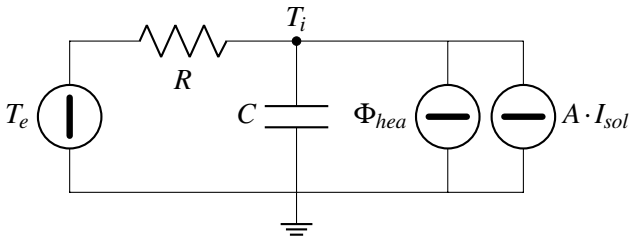
Table 2.1 – A set of structurally identifiable stochastic RC models

Note: the models are illustrated in the formalism of Bacher and Madsen (2011) with however european symbols of current and voltage sources.

Note 2: upon naming the RC models, the name lists:

- *$T_1T_2...T_m$ the m nodes with thermal capacities C_1, C_2, ... C_m,*
- *then the n resistances in the model $R_1R_2...R_n$,*
- *then the k solar aperture parameters $A_1A_2...A_p$,*
- *and finally the optional particularities of the model such as coefficient c_v.*

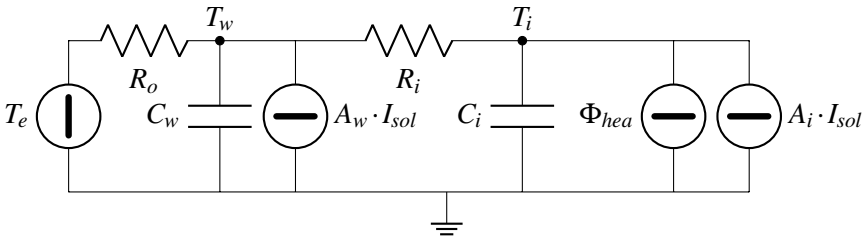
First order RC model



$$\begin{cases} \left[\dot{T}_i\right] &= \left[-\frac{1}{CR}\right]\left[T_i\right] + \left[\frac{1}{CR} \quad \frac{A}{C} \quad \frac{1}{C}\right]\begin{bmatrix} T_e \\ I_{\text{sol}} \\ \Phi_{hea} \end{bmatrix} + \sigma d\omega \\ y &= \left[T_i\right] + \varepsilon \end{cases}$$

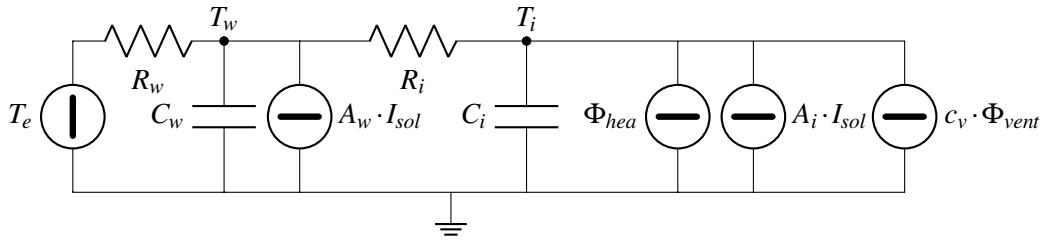
Model 2.3 - First order model T_i R A

Second order RC models



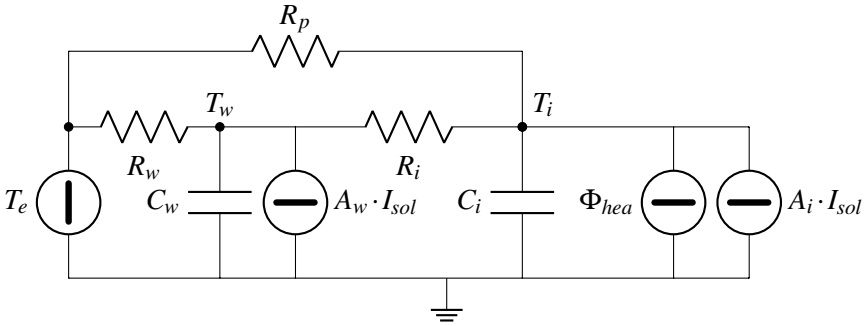
$$\begin{cases} \begin{bmatrix} \dot{T}_w \\ \dot{T}_i \end{bmatrix} &= \begin{bmatrix} -\frac{1}{C_w}\left(\frac{1}{R_o}+\frac{1}{R_i}\right) & \frac{1}{C_w R_i} \\ \frac{1}{C_i R_i} & -\frac{1}{C_i R_i} \end{bmatrix}\begin{bmatrix} T_w \\ T_i \end{bmatrix} + \begin{bmatrix} \frac{1}{C_w R_o} & \frac{A_w}{C_w} & 0 \\ 0 & \frac{A_i}{C_i} & \frac{1}{C_i} \end{bmatrix}\begin{bmatrix} T_e \\ I_{\text{sol}} \\ \Phi_{hea} \end{bmatrix} + \sigma d\omega \\ y &= \begin{bmatrix} 0 & 1 \end{bmatrix}\begin{bmatrix} T_w \\ T_i \end{bmatrix} + \varepsilon \end{cases}$$

Model 2.7 – Second order model $T_w T_i$ $R_w R_i$ $A_w A_i$



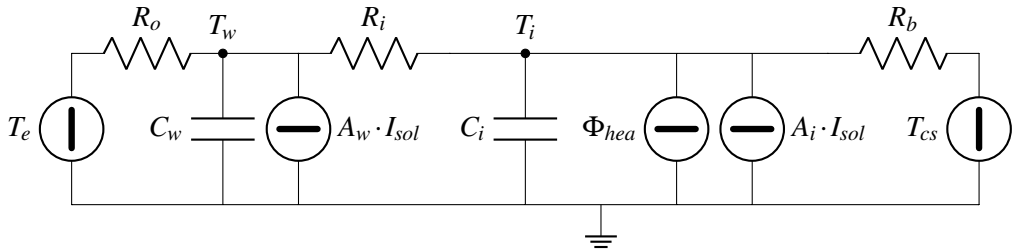
$$\begin{cases} \begin{bmatrix} \dot{T}_w \\ \dot{T}_i \end{bmatrix} = \begin{bmatrix} -\frac{1}{C_w}(\frac{1}{R_o}+\frac{1}{R_i}) & \frac{1}{C_w R_i} \\ \frac{1}{C_i R_i} & -\frac{1}{C_i R_i} \end{bmatrix} \begin{bmatrix} T_w \\ T_i \end{bmatrix} + \begin{bmatrix} \frac{1}{C_w R_o} & \frac{A_w}{C_w} & 0 & 0 \\ 0 & \frac{A_i}{C_i} & \frac{1}{C_i} & \frac{c_v}{C_i} \end{bmatrix} \begin{bmatrix} T_e \\ I_{\text{sol}} \\ \Phi_{hea} \\ \Phi_{vent} \end{bmatrix} + \sigma d\omega \\ y = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} T_w \\ T_i \end{bmatrix} + \varepsilon \end{cases}$$

Model 2.8 – Second order model $T_w T_i$ $R_w R_i$ $A_w A_i$ c_v with additional ventilation/infiltration flux



$$\begin{cases} \begin{bmatrix} \dot{T}_w \\ \dot{T}_i \end{bmatrix} = \begin{bmatrix} -\frac{1}{C_w}(\frac{1}{R_o}+\frac{1}{R_i}) & \frac{1}{C_w R_i} \\ \frac{1}{C_i R_i} & -\frac{1}{C_i}(\frac{1}{R_i}+\frac{1}{R_p}) \end{bmatrix} \begin{bmatrix} T_w \\ T_i \end{bmatrix} + \begin{bmatrix} \frac{1}{C_w R_o} & \frac{A_w}{C_w} & 0 \\ \frac{1}{C_i R_p} & \frac{A_i}{C_i} & \frac{1}{C_i} \end{bmatrix} \begin{bmatrix} T_e \\ I_{\text{sol}} \\ \Phi_{hea} \end{bmatrix} + \sigma d\omega \\ y = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} T_w \\ T_i \end{bmatrix} + \varepsilon \end{cases}$$

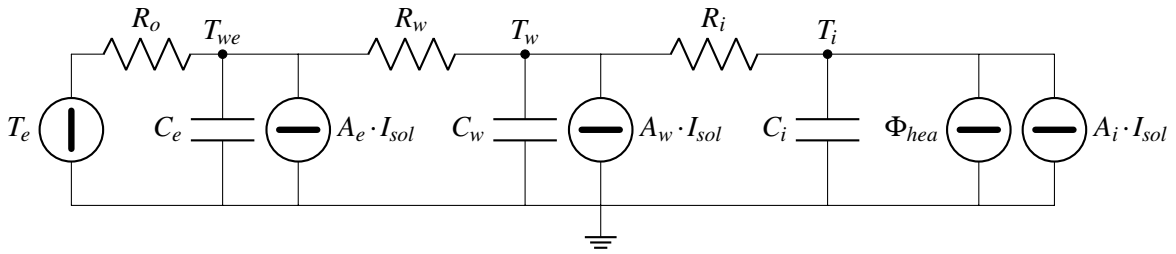
Model 2.9 – Second order model $T_w T_i$ $R_o R_i$ $A_w A_i$ with additional R_p for heat transfers with no thermal mass (equivalent to model 2.8 when Φ_{vent} is function of $\Delta T = T_e - T_i$)



$$\begin{cases} \begin{bmatrix} \dot{T}_w \\ \dot{T}_i \end{bmatrix} = \begin{bmatrix} -\frac{1}{C_w}(\frac{1}{R_o}+\frac{1}{R_i}) & \frac{1}{C_w R_i} \\ \frac{1}{C_i R_i} & -\frac{1}{C_i}(\frac{1}{R_i}+\frac{1}{R_b}) \end{bmatrix} \begin{bmatrix} T_w \\ T_i \end{bmatrix} + \begin{bmatrix} \frac{1}{C_w R_o} & \frac{A_w}{C_w} & 0 & 0 \\ 0 & \frac{A_i}{C_i} & \frac{1}{C_i} & \frac{1}{C_i R_b} \end{bmatrix} \begin{bmatrix} T_e \\ I_{\text{sol}} \\ \Phi_{hea} \\ T_{cs} \end{bmatrix} + \sigma d\omega \\ y = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} T_w \\ T_i \end{bmatrix} + \varepsilon \end{cases}$$

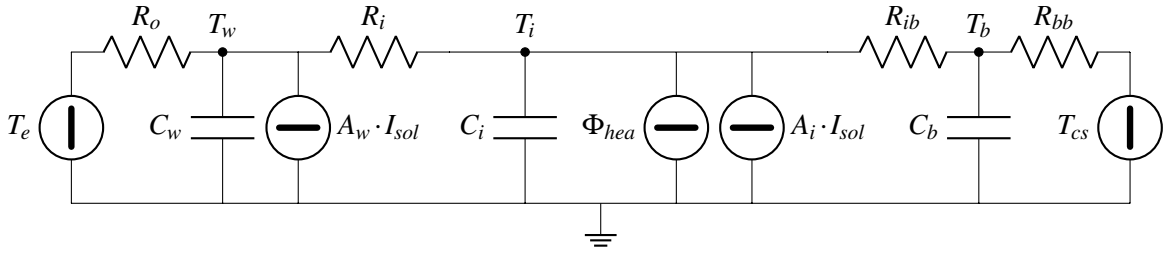
Model 2.10 – Second order model $T_w T_i$ $R_o R_i R_b$ $A_w A_i$

Third order RC models



$$\begin{cases} \begin{bmatrix} \dot{T_{we}} \\ \dot{T_w} \\ \dot{T_i} \end{bmatrix} = \begin{bmatrix} -\frac{1}{C_e}(\frac{1}{R_o}+\frac{1}{R_w}) & \frac{1}{C_e R_w} & 0 \\ \frac{1}{C_w R_w} & -\frac{1}{C_w}(\frac{1}{R_w}+\frac{1}{R_i}) & \frac{1}{C_w R_i} \\ 0 & \frac{1}{C_i R_i} & -\frac{1}{C_i R_i} \end{bmatrix} \begin{bmatrix} T_{we} \\ T_w \\ T_i \end{bmatrix} + \begin{bmatrix} \frac{1}{C_e R_o} & \frac{A_e}{C_e} & 0 \\ 0 & \frac{A_w}{C_w} & 0 \\ 0 & \frac{A_i}{C_i} & \frac{1}{C_i} \end{bmatrix} \begin{bmatrix} T_e \\ I_{\text{sol}} \\ \Phi_{hea} \end{bmatrix} + \sigma d\omega \\ y = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} T_{we} \\ T_w \\ T_i \end{bmatrix} + \varepsilon \end{cases}$$

Model 2.11 – Third order model $T_{we}T_wT_i$ $R_oR_wR_i$ $A_eA_wA_i$



$$\begin{cases} \begin{bmatrix} \dot{T_w} \\ \dot{T_i} \\ \dot{T_b} \end{bmatrix} = \begin{bmatrix} -\frac{1}{C_w}(\frac{1}{R_i}+\frac{1}{R_o}) & \frac{1}{C_w R_i} & 0 \\ \frac{1}{C_i R_i} & -\frac{1}{C_i}(\frac{1}{R_i}+\frac{1}{R_{ib}}) & \frac{1}{C_i R_{ib}} \\ 0 & \frac{1}{C_b R_{ib}} & -\frac{1}{C_b}(\frac{1}{R_{ib}}+\frac{1}{R_{bb}}) \end{bmatrix} \begin{bmatrix} T_w \\ T_i \\ T_b \end{bmatrix} \\ \qquad + \begin{bmatrix} \frac{1}{C_w R_o} & \frac{A_w}{C_w} & 0 & 0 \\ 0 & \frac{A_i}{C_i} & \frac{1}{C_i} & 0 \\ 0 & 0 & 0 & \frac{1}{C_b R_{bb}} \end{bmatrix} \begin{bmatrix} T_e \\ I_{\text{sol}} \\ \Phi_{hea} \\ T_{cs} \end{bmatrix} + \sigma d\omega \\ y = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} T_w \\ T_i \\ T_b \end{bmatrix} + \varepsilon \end{cases}$$

Model 2.12 – Third order model $T_wT_iT_b$ R_oR_i $R_{bi}R_{bb}$ A_wA_i, with an additional capacity C_b representing the ground floor, with measured air temperature in the crawl space T_{cs}

2.3 Practical identifiability

Exclusively working with structurally identifiable models has been shown to be necessary and existing tools allow a proper verification of such a property before exploiting any data. However, as expressed in the introduction 2.1, unicity of each parameter estimation might only be achieved with sufficiently informative data, hence the *practical* aspect of identifiability.

In the particular case of non intrusive data, the level of data information is expected to be lower than in controlled measurement conditions. Similarly to structural identifiability, practical identifiability should be verified upon each calibration. This section will review existing tools for such verification in the light of both frequentist and Bayesian approaches.

2.3.1 Grasping the necessity of practical identifiability

The building system is excited by boundary variables and by its systems in continuous time. The thermal response of the building too is in continuous time. Then, some physical quantities related to the boundary variables and the systems (the input) and the thermal response (the output) may be measured. The issue of practical identifiability is therefore function of :

- sufficient excitation of the system by the boundary and system variables,
- and adequate measurement of meaningful inputs and outputs.

Regarding the first aspect, the goal would be to have sufficient variability in the input variables in order to produce significant response in the measured output variables. Response to input variability is naturally function of the physical properties of the system. Therefore, a parameter is practically identifiable if the input data produce sufficient variability in the output in which that parameter is expressed. A lemma is that the variability needs to be distinguishable from the measurement noise. This issue is related to the frequencies and the amplitudes of the inputs, in relation to the system and to the low order model.

The second aspect relates to the determinant choice of the input and output variables to measure. Exhaustive measurements of the whole system are impossible, choosing to observe some particular input and output variable constitutes a first source of information loss. Secondly, there is also information loss inherent to the measurement sampling, spatial and temporal (from continuous to discrete time).

In addition, upon considering uncontrolled measurements, input variables are not optimized and may be correlated. Output data is spatially particularly scarce as to not burden possible occupancy and it might therefore not carry a lot of information either.

In the case of uncontrolled measurements, the data becomes poorly informative which is a threat to unicity of parameter estimation. Verification of unicity of the estimation after calibration becomes as important as a priori verification of the structural identifiability.

2.3.2 Assessing practical identifiability

The previous section underlined that practical identifiability needs to be verified. Existing tools depend on the approach chosen for solving the inverse problems. Tools for assessing the practical identifiability in first the frequentist approach and then the Bayesian approach is now discussed.

2.3.2.1 Practical identifiability in a frequentist approach

The frequentist approach considers that as the number of observations increases, any prior information on the parameters becomes negligible compared to that collected in the data Raillon and Ghiaus (2018), see also section 1.5.1.

From Raue et al. (2009), Raue et al. (2013) and also Kreutz (2019), in a frequentist approach, parameters $\hat{\theta}$ are estimated by minimization of an objective function, for example a least square method or by maximising the likelihood $\mathscr{L}$ for a Maximum Likelihood Estimation (MLE):

$$\hat{\theta}_{MLE} = argmax_{\theta}\mathscr{L}(\theta) \tag{2.22}$$

As asymptotic conditions are reached with a sufficient number of observations, it is possible to calculate asymptotic confidence intervals of the parameter estimation, assuming that the error is Gaussian:

$$\sigma_i = \hat{\theta}_i \pm \sqrt{\Delta_{\alpha,df} \cdot C_{ii}} \tag{2.23}$$

with $\Delta_{\alpha,df} = \chi^2(\alpha, df)$ the α quantile of the $\chi^2(.)$ distribution with df degrees of freedom ($df = 1$ for pointwise confidence intervals and $df = \theta$ for simultaneous confidence intervals) and with $C = 2H^{-1}$, H the Hessian of the objective function at $\hat{\theta}$.

In this frame, practical non identifiability will present as a flatness in the likelihood, which should be seen as an infinite confidence interval. However, asymptotic confidence intervals calculation cannot yield an infinite standard deviation and can therefore not be a sign of practical non identifiability. Likelihood based confidence intervals 2.24 should therefore been preferred (Raue et al., 2009):

$$\{\theta | log(\mathscr{L}(\theta)) - log(\mathscr{L}(\hat{\theta})) < \Delta_{\alpha,df}/2\} \tag{2.24}$$

Then, practical non identifiability is defined (Raue et al., 2009) as an infinitely extended region towards increasing and/or decreasing values of θ, although there is a unique minimum $\hat{\theta}$ for this parameter:

$$\exists \alpha | (\forall \theta < \hat{\theta}) \vee (\forall \theta > \hat{\theta}), log(\mathscr{L}(\theta)) - log(\mathscr{L}(\hat{\theta})) < \Delta_{\alpha,df} \tag{2.25}$$

The study of the identifiability of a parameter may then be done through the study of the profile likelihood function described in Equation 2.26:

$$PL_j(p) = max_{\theta|\theta_j=p} log\mathscr{L}(y|\theta) \tag{2.26}$$

The profile likelihood is the log-likelihood function for values of θ_j taken at p when all other parameters are re-optimized. A non identifiable parameter will, according to definition 2.25, be seen as an infinitely extended curve towards increasing or decreasing values of θ_j, which log-likelihood remains under $\Delta_{\alpha,df}$. Different profile likelihood regions are illustrated in Figure 2.2.

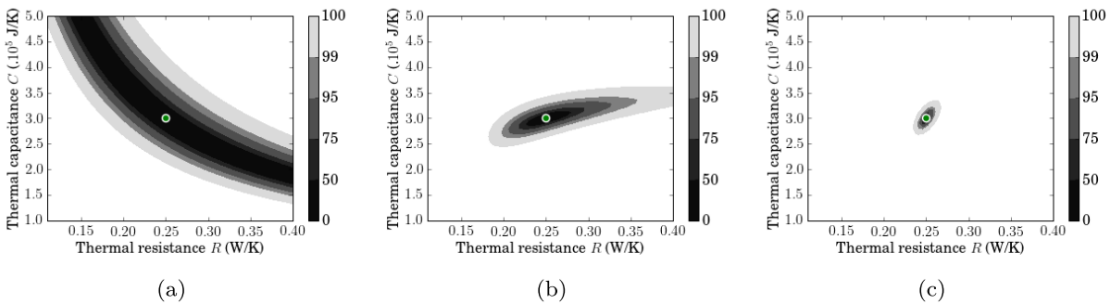


Figure 2.2 – Illustration of pairwise profile likelihood regions of a first order RC model (from Rouchier (2017)): (a) no heat input, (b) poor heat input, (c) very informative heat input. Narrow dark regions imply good identifiability. Large and infinite dark regions respectively imply unsatisfactory and non identifiability.

Applied in Maiwald et al. (2011), Kreutz et al. (2013), Raue et al. (2014), Deconinck and Roels (2017) and Brastein et al. (2019) among many others, this method can however be computationally burdensome. Kreutz (2018) proposes a locally equivalent and more efficient method : the Identifiability-Test by Radial Penalization (ITRP). The principle is to perform two separate calibrations:

- a first fit, giving $\hat{\theta}$ a first estimate,
- and make a second fit with a quadratic penalization function in the form of a circular manifold around $\hat{\theta}$ with radius R, thus pulling the new best estimate θ^* away from $\hat{\theta}$.

The model is practically non identifiable if the newly computed penalized cost function yields the same value than at the first fit. If both objective function values are identical, there is a non unique solution to the problem, hence non identifiability. The author shows that this method is 100 times faster than the profile likelihood method, although it comes with a few restrictions : the fitted model needs to be deterministic and the optimization routine used needs to be reliable.

2.3.2.2 Practical identifiability in a Bayesian framework a.k.a. "There is no Bayesian free lunch"

Fundamentally, Bayesian identifiability is not strictly speaking an issue as any parameter with a proper (integrable) posterior distribution is estimable Xie and Carlin (2006). Although a Bayesian analysis with proper prior knowledge is always feasible, if there is no learning from the data, i.e. from the likelihood, prior and posterior distribution will be identical Poirier (1998). In this sense, identifiability in a Bayesian framework is close to the notion of identifiability in a frequentist approach as it is a matter of learning from the likelihood Xie and Carlin (2006).

Figure 2.3 illustrates that principle. The prior, when it is well defined, brings information on the parameter of interest however valuable the collected data. When proceeding to calibration, the information gained from the data moves the distribution. Unless the prior reflected by chance perfectly the possible total knowledge there is to acquire on a parameter, the posterior will be more or less far from the prior.

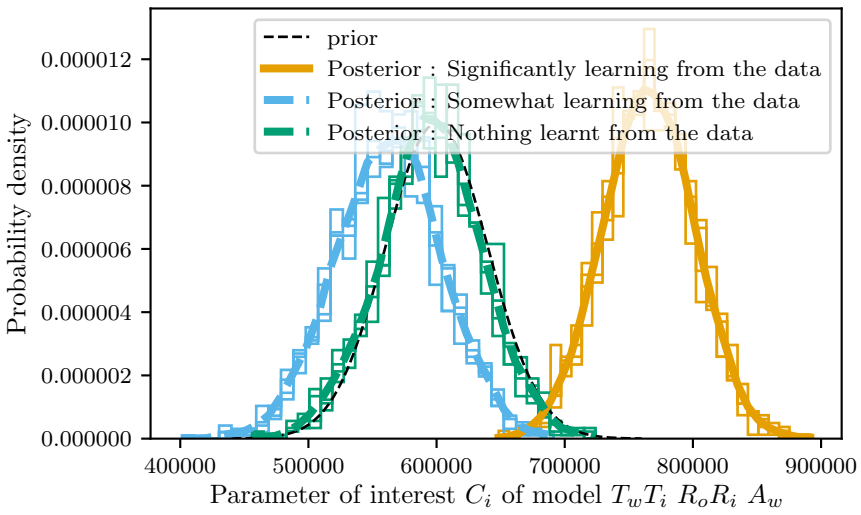


Figure 2.3 – Illustration of Bayesian learning: the prior brings in any case some information. The data then may move the knowledge of a parameter of interest, depending on richness of data

Considering $\theta = (\theta_1, \theta_2)$ a set of parameters, Dawid (1979) proposes a formal definition : the subset θ_2 is not identified by the data if the observation of the data does not increase our prior knowledge about θ_2 given θ_1 :

$$p(\theta_2|\theta_1, y) = p(\theta_2|\theta_1) \tag{2.27}$$

From 2.27, Sahu and Gelfand (1999) discuss this definition and underline that it does not imply that there is no Bayesian learning (i.e. it does not imply that $p(\theta_2|y) = p(\theta_2)$), but just that there is no conditional learning. The authors then propose a looser yet less formal definition : if "for at least some parameters, the data provides little information" this subset of parameters are "weakly identified". Provided there were a suitable metric between both distributions, this definition would read as 2.28 :

$$p(\theta_2|\theta_1, y) \approx p(\theta_2|\theta_1) \tag{2.28}$$

To this purpose, Xie and Carlin (2006) propose a metric 2.29 based on the Kullback-Leibler (KL) divergence, a quantity largely used for measuring the difference between two distributions. $D_{\theta_1,y}$ measures how much is left to learn given data y and is defined as :

$$D_{\theta_1,y} = KL(p(\theta_2|\theta_1), p(\theta_2|y)) = \int_{-\infty}^{\infty} p(\theta_2|\theta_1) log \frac{p(\theta_2|\theta_1)}{p(\theta_2|y)} d\theta_2 \tag{2.29}$$

As such a metric is almost never analytically computable, Xie and Carlin (2006) propose a methodology to estimate this metric with an MCMC approach, however requiring a second complete sampling of the posterior with the identifiable parameters fixed. Another solution would be to estimate the divergence by a k-nearest-neighbor distance method from samples of any distribution Hartland (2018); Wang et al. (2009). Figure 2.4 displays three posterior distributions of a poorly identifiable parameter with their associated prior distribution and their KL divergence. Divergences close to 0 show similarity between two distributions. Null divergence between prior and posterior shows therefore poor learning from the data.

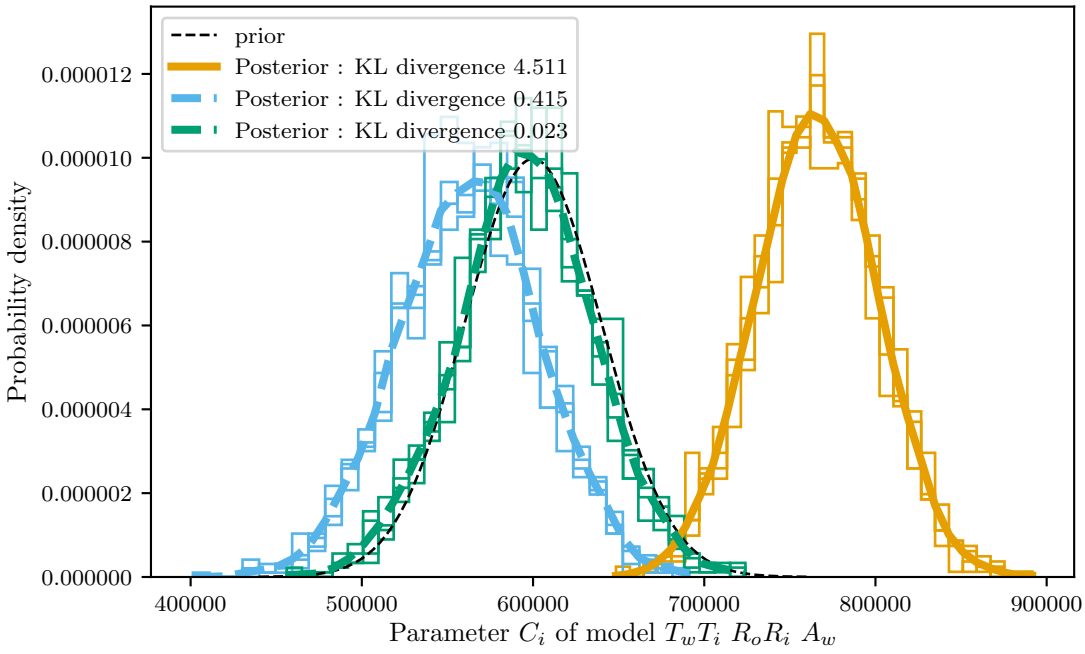


Figure 2.4 – Illustration of the Kullback Leibler divergence for three posterior distributions : each is compared to the prior in black dotted line. The less identifiable, the closer the distributions, the lower the KL divergence

In a nutshell, practical non-identifiability in a Bayesian framework is not per se an issue and can be quantified by a KL based metric. However, Raue et al. (2013) put forward that a possible

consequence of practical non identifiability is the impossibility to correctly sample the posterior distribution : the Markov Chain might be extremely slow in exploring the posterior and/or result in incorrect or even improper posterior distribution (Hobert and Casella (1996), Sahu and Gelfand (1999), Bayarri and Berger (2004), Gustafson (2015)). Bayarri and Berger (2004) add that if the only way for reaching a proper posterior is to set a proper prior, the posterior does not carry any significant meaning, which case will happen in case of non identifiability. These are the reasons why *"there is no Bayesian free lunch"*: a Bayesian approach is not incompatible with non identifiability, but it comes with the cost of burdensome or impossible computation.

Bayarri and Berger (2004) point out that in this particular case, frequentists methods about identifiability may become useful for understanding where the identifiability issue lies and for redefining afterwards more accurate priors. Raue et al. (2013) therefore call for an initial practical identifiability assessment through profile likelihood before sampling the posterior.

This calls for further work from a Bayesian angle on particular cases of structurally non identifiable models on which there is solid prior knowledge about the parameters. For example, model 2.5 accounts for a heating system efficiency coefficient and has been found to be structurally non identifiable. However, practitioners might have a significant expertise on actual values of efficiency for a given heating system, with reasonable uncertainty.

In that case, it could be possible to determine posterior distributions of the model parameters, provided that η_{hea} were defined with a prior distribution relevant to the expert knowledge. No information can be learnt from the data on parameter η_{hea} which means that it remains practically unidentifiable (prior and posterior should be identical). The posterior distributions of the other parameters would on the other hand reflect on the additional uncertainty with wider distributions than if η_{hea} were assumed constant. Far from being a drawback, wider uncertainties actually better reflect on the true overall uncertainty, given the available data and call for specific additional measurements, if one's goal is to improve the estimation.

2.4 Threats to physical interpretation and calibration good practice

Similarly to structural identifiability, practical identifiability has now been proven to be a major risk to parameter inference if it is not achieved. Practical identifiability needs therefore to be checked at each calibration run. Yet, in a poorly informative framework, there are numerous reasons to doubt that physical interpretation is feasible even when practical identifiability is proven. It becomes essential to understand how such bias happens as the objective of this work is the interpretation of the estimated thermal resistances.

Subsection 2.4.1 discusses the threats to physical interpretation of parameter estimates that have been identified from literature. Then, subsection 2.4.2 establishes some requirements regarding data collection in the frame of thermal performance estimation. Finally, subsection 2.4.3.2 proposes a workflow of good practice, valid for any model calibration from poorly informative data.

2.4.1 Threats to physical interpretation from poorly informative data

In Gustafson (2015), the author identifies various types of threats to the validity of a data set:

- "unobserved confounding variables": solving the inverse problem relied so far on the assumption that the observable model output y was solely explained by the measured input u. If however some physical quantities in the system $u_{unmeasured}$ have been overlooked and

not measured and yet had an influence on the output, it would corrupt the solution found.

- "poor measurement": the collected measurements of variable u_i may be not representative of the actual physical quantity (for example an indoor temperature when there is strong temperature gradient in the building, or wind velocity measurement which can be very tricky).
- "selection bias": this may happen when the assumption that the time- and space-finite collected data is representative of the system does not hold. For example, a priori, collecting data during 24h in mid-season might be not representative of the entire year and hence might bias the estimation result.

To unvalid data can be added a notion from Walter and Pronzato (1997): the model characterisation error. Indeed, upon defining structural identifiability, the authors incidently specify that the "process and model [must] have identifical structure (no characterization error)". This means that parameter estimates can be uniquely identified to the true values on the critical assumption that the model perfectly describes the system ("process" in Walter and Pronzato (1997)). However, as established in section 1.4.2, RC models are simplified. The bias induced by these simplifications on identifiability and interpretability has so far never been quantified and it cannot be excluded that they be significant.

All in all, Gustafson (2015) concludes that these threats are "a sort of corruption of the ideal data we wish we had". Indeed, physical interpretation would be straightforward from the parameter estimates had there been ideal data and ideal models.

From the angle of thermal performance estimation and under the light of Gustafson (2015), the conundrum translates as the following statement. A physics-driven RC model has been chosen, it is structurally identifiable but because it is physics-driven, it only suits ideal data.

Yet only corrupted data is available. The worst case scenario is when the nature of the corruption has not, or cannot be identified. Parameter inference carries a bias which will be difficult to uncover. Were the nature of the corruption known (for example the existence of heating system efficiency), it could be added to the model through a stochastic term although there is possible non identifiability.

In addition, the naive physics-driven nature of RC models actually also question the existence of ideal data for such models to deliver satisfactory estimates.

As a response to these threats, Gustafson (2015) concludes that "we are forced down the path of investigating what corrupted data can actually tell us about the quantities of scientific interest [...] in the hope that at least sometimes corrupted data can tell us something useful".

2.4.2 Enhancing information in data from uncontrolled measurements

The previous section explained how practical identifiability could be achieved, but corrupted data and inadequate RC model are a threat to identifiability to the actual value of interest, i.e. a threat to interpretation of the estimated parameters, without any warning sign being raised.

This section will therefore focus on the particular case of thermal performance estimation learnt from RC models in a non intrusive experiment. A few minimal requirements can be easily adopted in both the data collection and the choice of model.

As highlighted by Figure 2.5, the pitfalls that will be examined are:

- adequate measurements of the system (section 2.4.2.1)
- adequate temporal sampling (section 2.4.2.2)
- adequate model choice (immediatly following section 2.4.3.1)

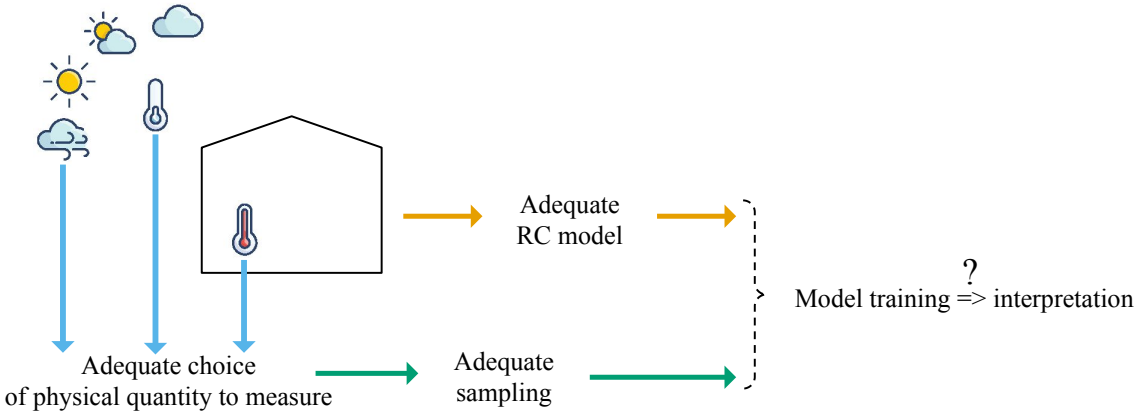


Figure 2.5 – Tackling threats to physical interpretation: adequate data collection and verification of model validity

2.4.2.1 Adequacy and availability: a compromise on input and output variables

Whereas controlled measurement conditions decreases the data corruption in the sense of Gustafson (2015), there is yet valuable information in data from uncontrolled non intrusive conditions. As illustrated by Figure 2.6, without being intrusive, there is most probably a set of possible set of physical quantities to measure that could provide satisfactory information for accurate parameter estimation. In addition, with a given set of measured variables, there are most probably settings that provide more satisfactory information, or at least that pulls away from data corruption.

The choice of physical quantities to measure resumes in outdoor (weather) and indoor conditions measurements.

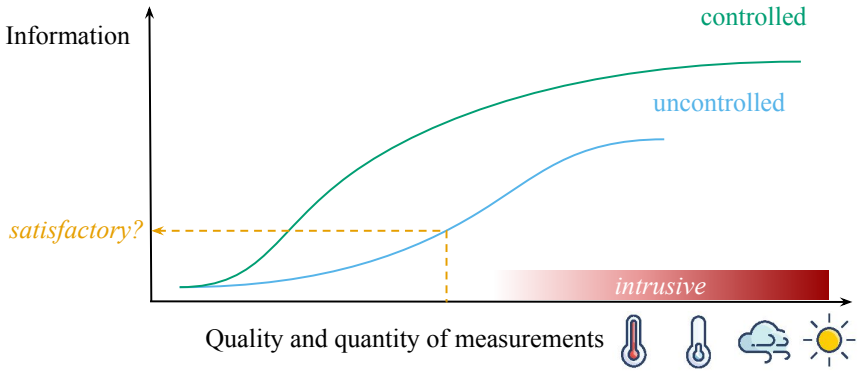


*Figure 2.6 – The quantity/quality versus information compromise: **controlled experiments** such as enriched heating signals naturally contain higher information value as they are designed for better identification. In **uncontrolled experiments** however, is there a minimal and sufficient set of variables that enables satisfactory practical identifiability and physical interpretation?*

Sensitivity of a physical quantity chosen as output variable to a desired parameter of interest

The choice of variables to measure should actually also be driven by the physical quantity of interest, in this work the thermal performance of the building envelope. Sampling here is assumed ideal.

The set of output is sufficient if, for given boundary conditions, the set is sensitive to the

parameter(s) of interest (p.o.i.) and distinguishable from the measurement noise. Applied to the building scale, it means that the measured output variable should be sensitive to the envelope thermal property of interest.

Figure 2.7 illustrates the findings, on this issue from Josse (2017), where the author performed a Morris sensitivity analysis of 30 building thermal properties on indoor temperature. Measured indoor air temperature, a viable output choice, is sensitive to a change in insulation thickness but not found sensitive to a thermal bridge. In this case, whatever the RC model used for data exploitation, it will be impossible to distinguish the effect of the p.o.i. in the output chosen: it is not practically identifiable. This means that if thermal bridges are the parameter of interest, indoor air temperature alone is inadequate. For the p.o.i. to be identifiable, a suitable output needs to be found. This issue has been widely discussed in the similar topic of numerical building models calibration, as in Strachan (1993).

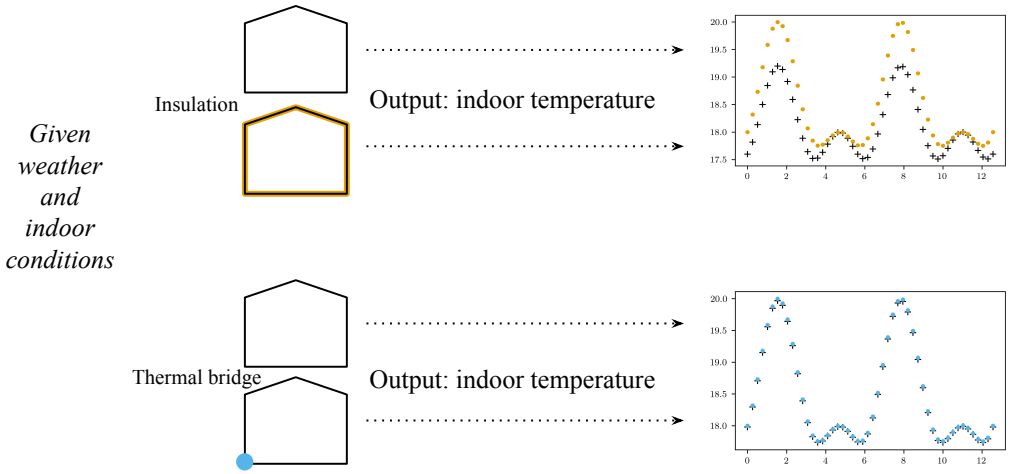


Figure 2.7 – The measured output variables should be sensitive to the parameter of interest. For example, measurement of the indoor air temperature is significantly sensitive to a change in the envelope insulation, but probably not significantly to a thermal bridge.

In parallel, Van Schijndel (2009) adds that the way the RC model (or any model to solve the inverse problem) accounts for the p.o.i. needs to also provide a sensitive predictive model output. This condition actually falls under a more numerical and algorithmic issue. If the p.o.i. has no influence on the output, whatever the value taken by the p.o.i., all will be considered equally likely (Van Schijndel, 2009). This rather numerical issue would normally be rarely encountered as the model is supposed to almost perfectly characterise the system: if the p.o.i. yields significant variability in the measurement, the p.o.i. will also yield significant variability in the predictive output.

Measuring the information breadth in data

Weather conditions cannot be controlled but are at the same time expected to play a major role in the identifiability of RC models in a non intrusive measurements approach. Van Schijndel (2009) state however that the measured input variables should at least deliver sufficient power into the system and propose to assess it through a Crest Factor 2.30. The lower the Crest Factor (minimum being 1), the larger the total energy delivery, thus larger excitation of the system and better signal to noise ratio.

$$C = \frac{max(u(t))}{\sqrt{mean(u^2(t))}} \tag{2.30}$$

For instance, upon deciding if it is interesting to use separately solar diffuse irradiation or just global irradiation, the calculation of the Crest Factor of these three variables on 5 days of winter weather 2.2 shows that diffuse solar irradiation carries intrinsically lower energy than global irradiation. Global horizontal irradiation has less energy than vertical south global solar irradiation. For the purpose of thermal characterisation, vertical south global irradiation may be not representative of the solar heat gains in the energy balance.

	Crest Factor
Diffuse solar irradiation	4.93
Global horizontal irradiation	3.24
Global vertical south irradiation	2.72

Table 2.2 – Crest Factor of solar diffuse, horizontal direct and south vertical direct irradiation on 5 days of winter data

However, this metric is independent of both the system and the calibrated RC model, whereas the interest in a weather variable lies in its dynamic relationship to both. In addition, as underlined in Godfrey et al. (2005), such metric is not particularly relevant for model identification.

Any linear RC model representing the thermal behaviour of the building may be transformed in the Laplace domain and studied as a transfer function. Frequency and amplitude of the output may be studied with respect to frequencies and amplitudes of the inputs, thus offering a more complete view of minimum information needed in the data in relationship with the building model. On this principle, Naveros and Ghiaus (2015) proposed at wall scale to establish the minimal amplitude of each input variable at characteristic frequencies of the wall that can be measurable in the output. This is done by analysing in the Bode plot the response of the significant frequencies of the wall, as shown in Figure 2.8.

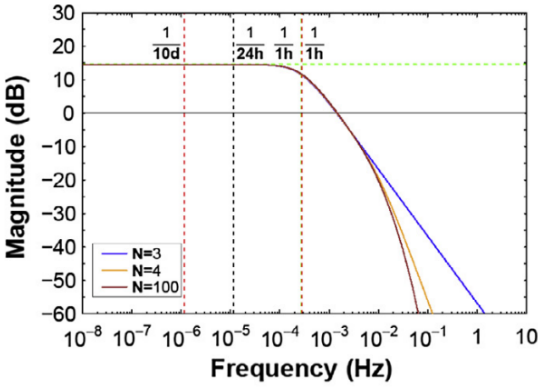


Figure 2.8 – Illustration of a Bode plot from Naveros and Ghiaus (2015): the response of the inside heat flux depends on the frequency of the outdoor temperature condition. Given the heat flux measurement uncertainty, high frequencies can only be measured from very large (and very quick then) amplitudes in the outdoor temperature, because the corresponding magnitude is large.

Response of a measured output to high or to low frequency phenomena will be different. Considering in addition the measurement uncertainty, each relevant frequency has a detectable amplitude threshold. On the example of a gypsum wall displayed in Figure 2.8, relevant

frequencies of the wall are at $^1/_{24h}$ (for hourly measurements during 10 days, hence the $^1/_{1h}$ and $^1/_{10d}$ lines). Magnitude at frequency $^1/_{24h}$ is 15 dB, which implies an amplitude gain of $10^{15/20} = 5.62$. The authors estimate at 2.50 W the accuracy of the measurement, which means that the minimal measurable variation of outdoor temperature at a frequency of $^1/_{24h}$ h is $^{2.50}/_{5.62} = 0.44\ ^\circ C$.

Applied at building scale, if the RC model were to be a fitted $T_wT_i\ R_oR_i\ A_w$ model (model 2.7), the measured output would be the indoor temperature and the inputs outdoor temperature, solar global irradiation and heating power. Table 2.3 shows the minimal variability needed for the input to be measured in the output, distinctively from the noise. The minimal values are easily achieved with ordinary weather and heating conditions. This means that at building scale, the inputs variability provide at least a measurable effect on the chosen output.

Output	Input	Frequency h^{-1}	Magnitude dB	Amplitude gain $10^{M/20}$	Minimal variation
T_i (°C)	T_e (°C)	1/24	- 11.9	0.25	0.4 °C
T_i (°C)	I_{sol} (W/m^2)	1/24	- 34.6	0.019	5.4 W/m^2
T_i (°C)	$P_{heating}$ (W)	1/24	- 35.0	0.018	5.6 W
T_i (°C)	$P_{heating}$ (W)	1/12	- 40.3	0.009	10.3 W

Table 2.3 – Minimal measurable variation of input as in Naveros and Ghiaus (2015) for model $T_wT_i\ R_oR_i\ A_w$ (2.7)

This method does not however provide a metric of the actual information in the data nor does it guarantee identifiability of the low order RC models.

How to avoid missing a significant input variable

Again provided that the input is ideally sampled, input measurement is adequate if there is no other significantly influential variable on the output (Gustafson, 2015; Madsen et al., 2015; Madsen and Holst, 1995). This issue is actually related to the choice of a proper model and how it takes the different input variables into account.

On this matter, Kristensen et al. (2004); Madsen and Holst (1995) both define and use a procedure for model selection. Checking for adequate input modelling and measurement may be assessed through an analysis of the so-called residuals, the difference between the prediction of the calibrated model and the measurements. Normally, if the model has no characterisation error and the measurement error is normal, then the residuals show white noise properties. Shortcomings of the model, all the more so with a missing measurement of a significant boundary condition, will be detected as non white residuals. A thorough analysis of the correlations between boundary conditions and the residuals can uncover such shortcomings. The model selection process itself will be later discussed in section 2.4.3.1.

Making the best of non intrusive heating control

Weather conditions are not controllable, but the heating may be. In a non destructive but intrusive approach, the usual solution to tackle optimal informativeness is through a design of experiment of the heating signal (Bacher and Madsen, 2011; Fedorov, 2010; Rouchier, 2017). In non intrusive conditions however, the heating signal needs to provide occupant friendly temperatures and in the prospect of possibly using smart meter data, design of experiment is not an option.

In occupied conditions, the indoor temperature does indeed not freely vary in function of outdoor conditions and a heating power, but it is rather programmed to a specific temperature set point, and the heating power adjusts as to follow the set point program. Let us by the way emphasize that temperature scheduling when the indoor temperature is the model output variable is far from optimal. The system would be indeed better excited by an independent heating signal.

All the same, in occupied buildings, temperature set points may take different forms. Their dynamics are concededly not ideal nor even adequate for RC models calibration, but could yet be informative. For example, Senave et al. (2019) test different indoor temperature scenarii and their influence on physical interpretation of ARX models:

- constant or varying temperature
- frequency of temperature set point switch
- amplitude of temperature set point switch
- week-ends different or identical to week days

Similarly, for RC models, these aspects are expected to play a role on practical identifiability. According to Wei et al. (2014), heating schedules are driven by occupancy, day and night time and by room. As for the temperature set point, it has been found to be related to the room type, the installed systems and thermal sensation.

Constant temperature set point schedule is not fit for RC model exploitation, as the data lacks dynamics, as previously mentioned in section 1.2. Occupant-friendly indoor temperature schedules however also include varying temperatures, which produces a slight dynamic in the indoor conditions and is profitable to practical identifiability.

Conclusion

This section has reviewed methods that aim at determining whether the choices of input and output measurements are sufficient. Strictly speaking, there is indeed no tool to measure the information in a given dataset which would translate as a practical identifiability quantitative or qualitative indicator. It is however possible to determine the extent to which a measured output is responsive to stress in the inputs through a frequency analysis as in Naveros and Ghiaus (2015).

In any case, missing information can be made clear if a statistical model validation is performed. The latter is developed in section 2.4.3.1.

2.4.2.2 Adequate temporal sampling of the measured input and output

Literature background on temporal sampling

From a continuous to a discretized time expression, RC models may suffer from aliasing. Aliasing happens when two different signals become indistinguishable when sampled. In indoor temperature measurements, faulty sampling will result in misreading or ignoring significant variations of temperature at high frequencies (Madsen et al., 2015).

Overlooking high frequency temperature variations may then result in practical non identifiability. RC model calibration should therefore necessarily use datasets with a minimal sampling frequency of twice the highest frequencies, as stated by the Shannon theorem (Madsen, 2008). Madsen et al. (2015) suggest that the sampling time should ideally be kept under the hour.

This issue also relates to that of the representative characteristic times of the indoor air temperature behaviour. Sicard et al. (1985) performed a spectral analysis of the response of indoor air temperature to external stress on a numerical study case : outdoor temperature, heat flux indoor and heat flux on indoor walls (see Fig 2). They showed that the indoor air temperature shows a strong mode at characteristic time 55 h. The value itself is related to the simulated building, but the order of magnitude is significant. In addition, they show how indoor air temperature has several significant modes around the characteristic time 0.1 h, corresponding to local heating

of the air and of the surface of neighbouring walls. Although particularly large in magnitude compared to the others modes, they show that the 55 h mode is not sufficient to make accurate prediction. A value of 55 h is naturally extremely case dependent, but gives a general idea of the order of magnitude that caneb expected for characteristic time. Predictions were more accurate when taking the two modes with the lowest frequency and in addition a higher frequency mode around 10 h^{-1} (time 0.1 h i.e. 6 min). These high frequency phenomena suggest then, according to Shannon's theorem, that the sampling time step should be maximum 3 minutes. Longer time steps will fail to provide informative data for 2nd and higher order dynamic models.

Influence of temporal granularity on parameter calibration: measuring how much is learnt

The conclusions of Sicard et al. (1985) were based on prediction performance but suggest much lower temporal sampling frequencies than the "under the hour" in Madsen et al. (2015).

This section proposes therefore to numerically assess the influence of temporal sampling on parameter estimation. To do so, the collected data from an actual experiment in a house of the platform INCAS is used as training set for RC model calibration.

The data, measured at a 1 minute time step, has been collected in a two storeys unoccupied *I-BB* house in Le Bourget du Lac (Savoie, France). The *I-BB* house is part of an experimental platform of 4 different houses, called the INCAS houses. The building envelope of this specific house is made of traditional shuttered concrete walls. For this experiment, there were no thermal insulation on the vertical walls, but heavy insulation below the ground floor concrete slab and in the attics.

During the experiment, the heating power is controlled as to follow a *Pseudo-Random Binary Signal* (PRBS) (Godfrey, 1979), as shown in Figure 2.9. This type of signal takes two possible values, 100 W or 5600 W and ressembles a square wave signal with multiple frequencies. The PRBS signal is designed to cover both high and low frequencies to which a building might be sensitive.

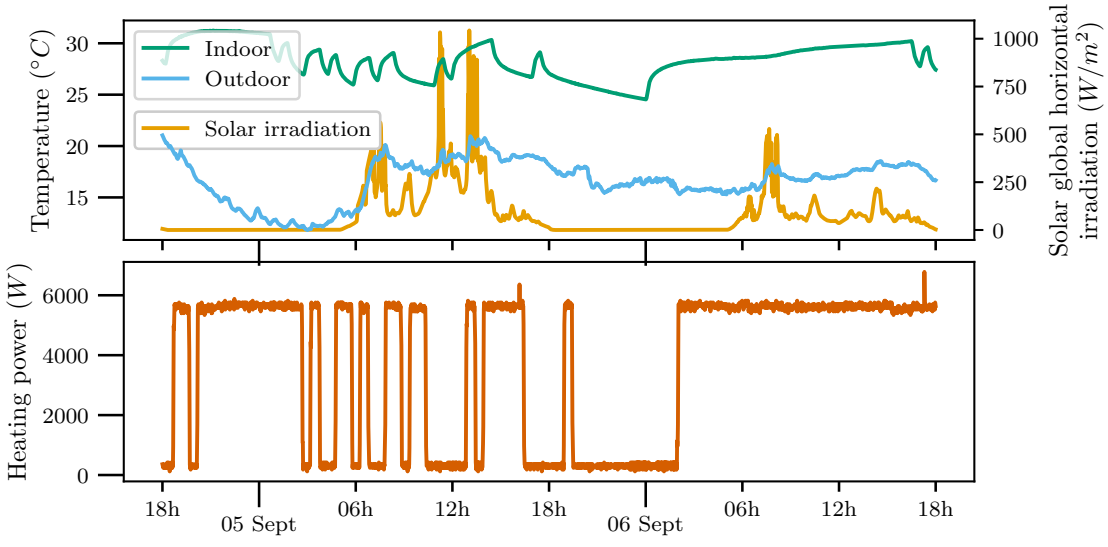


Figure 2.9 – Boundary conditions in the I-BB INCAS house during the pseudo-random signal test

As shown in Fig. 2.9, 48 h of data is used for the calibration. The outdoor temperature varies between 13 °*C* and 23 °*C* as is expected at the beginning of September in Le Bourget du Lac. The blind shutters were kept closed during the experiment as to limit the influence of the solar irradiation.

The data collected in the *I-BB* house serves now as training data for the second order model 2.7 (model $T_w T_i\ R_o R_i\ A_w$). The 1 minute time step measurements allow to resample the data at will. Time steps from 8 minutes up to 120 minutes are performed. Each subsequent dataset is used as training data for calibration in a Bayesian approach.

In addition, it is proposed to use the Kullback-Leibler (K-L) divergence to measure the information gained by the RC model from each differently sampled dataset.

Adequate temporal sampling for RC model calibration

Fig. 2.10 shows the posterior distributions of four parameters of the second order model $R_o R_i$ $C_w C_i\ A_w$. There are roughly two cases:

- the parameter has rather similar posterior distributions, regardless of the time granularity. It is the case of parameters R_o;
- posterior distributions seem to significantly depend on the time granularity, as in particular for parameter C_i and to a lesser extent R_i and C_w.

Parameter R_o presents distinct prior and posterior distributions. All posterior distributions look relatively close. Confirmation is brought by the KL divergences, estimated at around 1 except for the 3 minutes and 120 minutes time steps. This would indicate that, to some extent, the estimation of R_o is insensitive to sampling with time steps shorter than 120 minutes. Shorter time steps however result in lower uncertainty in the estimation.

Contrastingly, estimations of parameter C_i or C_w show a different behaviour. Estimations from time step sampling larger than 15 minutes have significantly lower KL divergences than short time step samplings. Posterior distributions tend to show much narrower uncertainties with the 8, 11 and 15 minutes time step samplings, which contrasts with the 90 and 120 minutes sampling, closer to the prior. This all suggests that parameters C_w and C_i barely learns with low frequency sampled data.

Parameter R_i too shows a growing KL divergence with shorter time step samplings. The uncertainties are significantly narrower with the 8, 11 and 15 minutes time step samplings. There has been however in each estimation learning from the data. The amount of information gained seems just larger with a short time step sampling. Noteworthy, divergences from 90 and 120 minutes are very large, their posterior distributions are quite spread, and show higher values than the shorter time steps results. This might be the result of aliasing or also that the data is clearly insufficient for this second order model calibration. Model selection validation would make the latter clear and must in any case be performed (Madsen et al., 2015).

Discussion and conclusion on temporal sampling

An obvious criticism of using the KL divergence for that purpose is that the metric just measures the difference between two distributions. If the prior had been correctly guessed, maybe by pure and random chance, the low KL divergence is not a sign of non identifiability. This means that no additional information was gained from the data, as the prior was already significantly informative. In this sense, the KL divergence is ***not*** an indicator of practical identifiability but merely an indicator of additional information gained from the data alone.

Interestingly, the KL divergences are not dependant on the order of magnitude of the parameters. Whether of order 10^6 or 10^{-3}, the metric seems to be consistent. This aspect is an argument in favour of the use of the K-L divergence.

All in all, estimating the KL divergence could be used as a warning sign, provided the model had passed basic model selection. Indeed, a practically non identifiable parameter will result in a

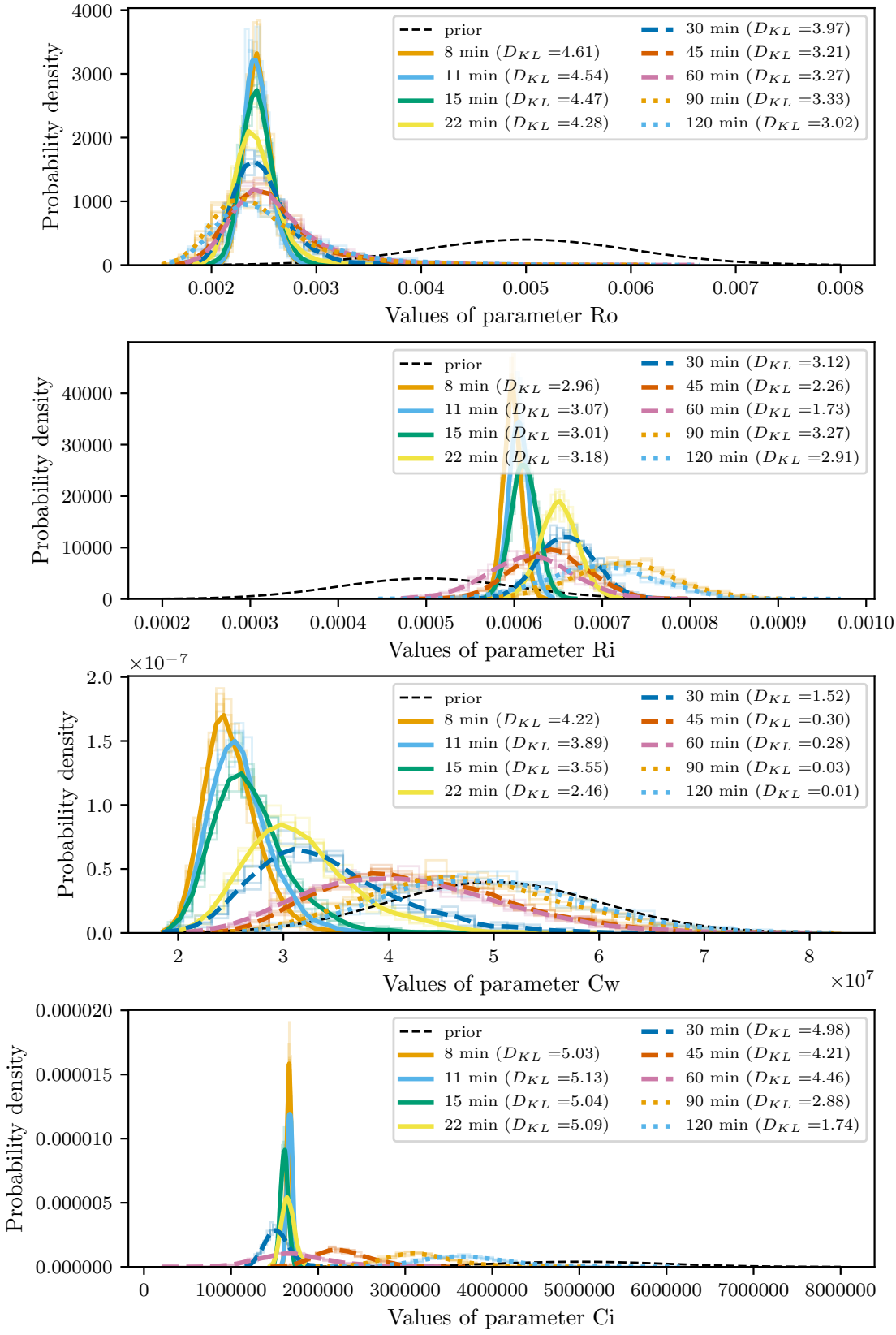


Figure 2.10 – Posterior distributions of parameters R_o, R_i, C_w and C_i from model 2.7, 48 hour data on different sampling time steps. The larger the KL-divergence, the more different posterior and prior distribution, the more the model has learnt from the data.

null KL divergence but a null KL divergence merely indicates no learning from the data. To waive the doubt, a second Bayesian calibration could be performed with less informative priors, in which case a practically non identifiable parameter would still yield null divergence.

Now focusing on adequate time granularity, the results of Fig. 2.10 suggested that a sampling time lower than 15 minutes is preferable to reach practical identifiability of all parameters of the $R_o R_i\, C_w C_i\, A_w$ model. This time granularity is larger than what was suggested in the work of Sicard et al. (1985)(time steps around 3 minutes). Yet it is not surprising as adequate sampling depend on the case study: if the building has higher characteristic frequencies, 8 to 15 minutes sampling might be sufficient. Another explanation is that such second order model does not catch the higher frequency physical phenomena, and does therefore not need higher temporal granularity to be successfully calibrated.

It should in any case be inferred that the sampling time is fundamentally dependent on the thermal properties of the building under study and that the risk of practical non identifiability grows with sampling of large time steps.

Let us also underline the following. When shorter time step samplings imply better learning for parameters representing high frequency phenomena (R_i or C_i), it is not detrimental to identifiability of parameters representing low frequency phenomena (R_o or to a certain extent C_w).

2.4.3 And yet not enough for interpretation: workflow for meaningful calibration

Data from non intrusive, non destructive and occupant friendly measurements cannot be considered as rich, let alone ideally informative. The previous sections however established that specific measurement conditions may provide acceptable data for RC model calibration.

The parameter estimates may converge towards the *true* values of the system provided that (Walter and Pronzato, 1994):

- the model is structurally identifiable (all parameters are theoretically identifiable),
- the data is adequately informative, which brings practical identifiability,
- the model characterizes perfectly the system.

The last point is in the application of thermal performance estimation never met. Perfect characterisation would indeed mean that the thermal model perfectly describes the entire system. Any building model, however detailed, cannot be expected to describe the entirety of the physics. In a non intrusive framework, where high order models are excluded, lower order models may overlook, simplify and/or aggregate some of the physics.

Section 2.4.3.1 first details how model discrepancy too is a threat to physical interpretation and presents means to deal with this issue: model discrepancy acknowledgment on the one hand and on the other hand model selection and validation.

These steps again are necessary and complete an overall calibration procedure of good practice for hopefully meaningful results. The procedure is laid out in section 2.4.3.2.

2.4.3.1 Addressing the issue of model discrepancy

Model discrepancy has been identified as a threat to physical interpretation, even when structural and practical identifiability are given. There are two complementary means to deal with the issue: model selection and validation on the one hand and acknowledgement of model discrepancy on the other hand.

Model selection and validation consists, from a set of admissible models, in choosing the model that best fit the available data. The underlying assumption is that between two models, there is higher chance of practical identifiability and physical interpretability with the best fitting model.

Model selection on one hand implies comparison between two models according to a certain performance metric. Model validation on the other hand consists in assessing of how well it performs without comparison to another model. Fit for a frequentist approach, the following verifications are usually performed (Madsen et al., 2015):

- **Student tests for parameter significance**: if a parameter is not significant, it should be removed from the model (or fixed at a certain value). The application is meant for a frequentist approach (t tests in a Bayesian approach are debatable (Kruschke, 2013));
- **correlation between estimated parameters**: high correlation (close to -1 or +1) is a sign of practical non identifiability as it indicates potential combination of parameters yielding identical likelihoods;
- validation of the **white noise property of the residuals**: if the residuals have no white noise property, the model should be enhanced to a higher order, see Section 2.4.2.1. Correlation of the residuals to input variables is an option to uncover what physics is missing on the model;
- **likelihood ratios** for selection among nested models (Bacher and Madsen, 2011);
- **information criteria** like Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC) for selection among **non** nested models. Madsen et al. (2015) add that for physical interpretation, the BIC criterion should be preferred. Let us here note that the BIC is not properly Bayesian as it is calculated from the maximum likelihood estimate, which can be obtained in a frequentist approach.

In a Bayesian framework, model selection on the basis of prediction accuracy can be done by use of the full information from the posterior distribution of the parameters. Such selection methods include the Bayes factor (for null hypothesis testing (Kruschke, 2018; Wakefield, 2013)), the Widely Applicable Information Criterion (WAIC) (Watanabe, 2013) or the Leave One Out (LOO) cross-validation, approximated by the Pareto smoothed importance sampling to achieve a robust and numerically feasible estimation of its value (Vehtari et al., 2017). Furthermore, Burnham and Anderson (2004) discuss how the AIC and BIC could be used as model selection metrics regardless of the frequentist or Bayesian approach as such criteria have foundations in both approaches.

Independently to any model selection procedure, model validation is done through an analysis of the prediction residuals $\hat{\varepsilon}$: $\hat{\varepsilon}(t) = y(t) - F(X,U,t)$ with F the prediction of the RC model (Kristensen et al., 2004; Madsen and Holst, 1995). The prediction can be performed on either the same set of collected data or a different set as a sort of cross-validation dataset. For the model to be considered validated, the residuals have to show white noise properties. In particular, an absence of correlation to any input or other confounding variable is expected.

A series of statistical tests are possible to assess the white noise properties of residuals. Kristensen et al. (2004) states that the most powerful methods consist in inspecting the autocorrelation function and partial autocorrelation function of the residuals as in Figure 2.11. These tests assess the white noise property of the residuals : if there is a lag dependency, the residuals do not have good prediction capability and cannot be considered as white noise.

In addition, a quantile-quantile (q-q) plot analysis of the residuals can be done to verify the normality property of the residuals (Heiberger and Holland, 2015). The idea is to plot the quantiles of the inverse cumulative distribution function of the residuals against that of a normal distribution taken as reference. If the residuals have white noise properties, they should be close to normally distributed. The qq-plot shows then similarity in both the residuals and the normal reference distribution.

Figure 2.12 shows on the contrary non normal residuals. Two models (models 2.3 and 2.7) have been calibrated on some data and the qq-plot of their residuals determined. Although the residuals are in overall close to the red zone of normality, the models both diverge on the left and

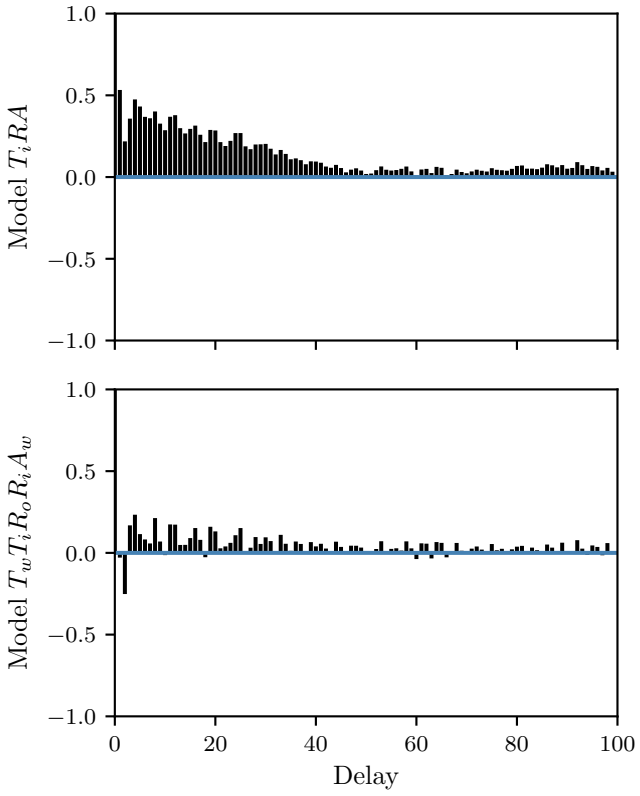


Figure 2.11 – Illustration of an auto-regression function (ACF) test on the residuals from predictions of models T_iRA (2.3) and a $T_wT_iR_oR_iA_w$ (2.7). Although not perfectly uncorrelated, the second ACF shows less autocorrelation than the first.

right side of the plot, indicating heavier tails than a normal distribution. This is due to the fact that both models have some shortcomings in their prediction capability.

Bacher and Madsen (2011) add to these methods a cumulative periodogram test to the model evaluation step to check if the dynamics at different time scales are properly modelled. An illustration of a cumulative periodogram is given in Figure 2.13.

When the residuals are found not to have white noise properties, Madsen et al. (2016) suggest to check for cross-correlation of the residuals with the input variables. If the model improperly accounts for one of the input variables, the cross-correlation of the residuals with that variable will be significant. The input variables and the residuals may also be visually compared by plotting the former along with the latter.

In any case, if the non white property is due to an unobserved confounding variable, the cross-correlation functions with the measured input will yield insignificant values whereas the residuals are non white. This would be a strong indication of a missing input in the system's environment.

Model selection and validation has then the advantage to bring upfront a model that at least describes satisfactorily the collected data. At the same time, it is misleading in the sense that it is more a matter of un-validating the other models than validating the chosen model as the one that characterizes the system. At the same time, model selection tools tend to validate model based on prediction, which is no per se a guarantee for physical interpretation, although there is a certain correlation.

About model selection, Tarantola (2006) argues that inferring the "best model" from data is philosophically incorrect: "Observations cannot produce models, they can only falsify [them]". There is then no point in thinking there is a "right" model to fit some collected data. Rather,

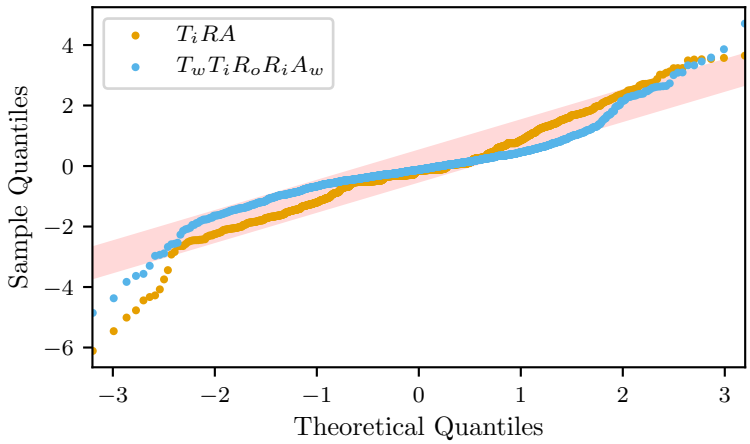


Figure 2.12 – Illustration of a quantile quantile analysis of the residuals of a T_iRA (2.3) and a $T_wT_iR_oR_iA_w$ (2.7) predictions: the red zone shows how a normal distribution would appear, such as residuals with white noise properties. Here, both models have shortcomings visible as heavy tails on the left and right hand side.

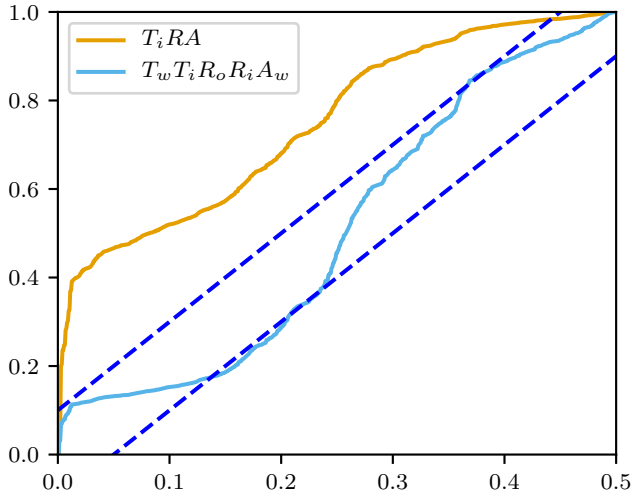


Figure 2.13 – Illustration of cumulative periodogram of the residuals of predictions with both models T_iRA (2.3) and a $T_wT_iR_oR_iA_w$ (2.7). The blue dotted lines indicate the confidence intervals for normally distributed noise. Residuals of the first order model do not remain within these bounds, meaning they are strongly non-normal. The second order model performs better, although an almost centered line would be the objective in a model selection process.

model selection and validation is a matter of selecting some models that are not unvalidated by the data, models that are "not too wrong" (Wit et al., 2012), while acknowledging that each model is intrinsically a biased representation of reality.

Model discrepancy should therefore in any case be explicitely formulated in the model structure. In Madsen et al. (2015), the Subtask 3 of the IEA EBC Annex 58 concludes that "the structure describing the uncertainty should be built into the model", therefore accounting for known shortcomings of the models used. For RC models, this can be done as an addition of a stochastic term $\sigma \cdot d\omega$ as in Equation 2.31a, where $d\omega$ is a Wiener process (Madsen et al., 2016).

Uncertainty can also be introduced as shown in Equation 2.31b with a model discrepancy time dependent term $\delta(U_i)$ (Kennedy and O'Hagan, 2001). In this latter formulation, $\eta(U_i)$ is the deterministic RC model output. The shortcomings of the deterministic form are corrected by a correlation factor ρ and by the model discrepancy term $\delta(.)$. Model discrepancy term $\delta(.)$ can take the form of a Gaussian process and can be calibrated either at the same time than the RC model, or separately in a modular calibration procedure (Arendt et al., 2012). The Kennedy and O'Hagan (2001) technique is also called "Bayesian calibration" in literature (Chong and Menberg, 2018; Heo et al., 2012).

$$dX(t) = F(X,U,t) = A \cdot X(t) \cdot dt + B \cdot U(t) \cdot dt + \sigma \cdot d\omega \tag{2.31a}$$

$$y_i = \rho \cdot \eta(U_i) + \delta(U_i) + \varepsilon_i \tag{2.31b}$$

The stochastic term has the advantage to have an effect on the model's states, whereas the approach proposed in Kennedy and O'Hagan (2001) considers the model to be deterministic, i.e. the states are unchanged by the Gaussian process. The stochastic term as recognition of model discrepancy for RC representations may be then better suited, as the physics in the RC model are simplified and aggregated. There is little chance indeed that a deterministic model be close to the actual building behaviour.

Kennedy and O'Hagan (2001) furthermore expressly underline that the interpretation of the parameter estimates as estimates of the true values is "dangerous", even with acknowledgement of model discrepancy. Brynjarsdottir and O'Hagan (2014) went further into the study of physical interpretability and found that learning about the physical meaning of parameters from inaccurate models without error model recognition is unwise. Adding a model discrepancy term as in Kennedy and O'Hagan (2001) with weak prior enhances the accuracy of predictions but may not help with learning about physical values. Limited physical learning predicted by Kennedy and O'Hagan (2001) has indeed been found true in Chong and Menberg (2018): in a Bayesian calibration, the authors test variate prior distributions and model complexities. They conclude that although the prediction performance had been found very satisfactory, physical interpretation was uncertain as it seemed dependent on the chosen priors and the model order.

In any case, recognising model discrepancy as an unknown in the model formulation takes the uncertainty into account in the calibration process and propagates that uncertainty onto the results, therefore reflecting the actual state of knowledge about the parameters.

2.4.3.2 A workflow of good practice for model calibration

Model selection and validation completes a workflow of steps to necessarily perform in order to estimate uniquely the model's parameters and, hypothetically, have physically interpretation.

The workflow, depicted in Figure 2.14, is constituted by the following steps:

- consideration of a quantity of interest (q.o.i), such as a thermal property of the building envelope (color gold in the figure),

- adequate data collection (in green) through sufficient input and output variables observation and through adequate sampling. Is driven by:
  - choice of q.o.i
  - practical non identifiability of 1 or more parameter(s)
- modelling choices (in light blue) driven by:
  - choice of q.o.i.
  - choice of model discrepancy recognition
  - verification of the structural identifiability
  - practical non identifiability: need of simplification
  - model statistically unvalidated: need of simplification
- adequate choice of model (in purple) driven by:
  - verified practical identifiability
  - model validation
  - best performing model selected among several tested
- in the end, estimated model parameters should lead to an estimate of the q.o.i.

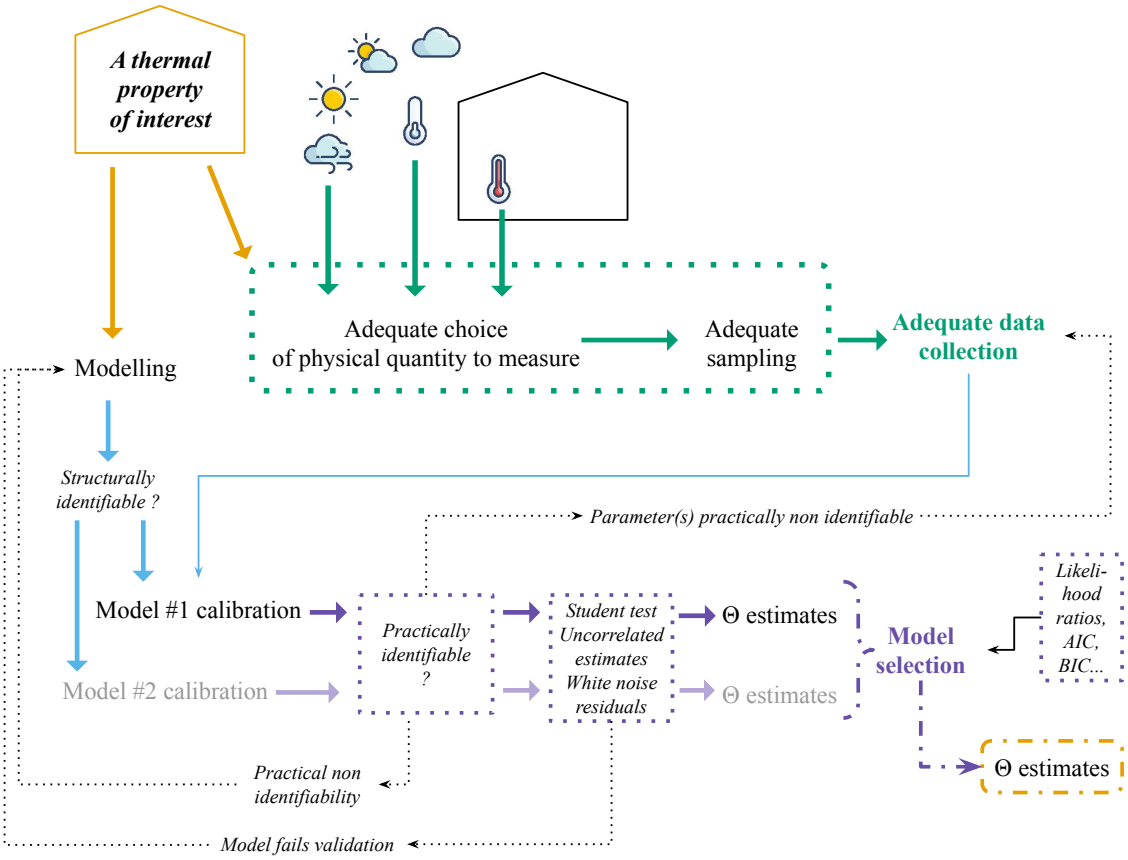


Figure 2.14 – Workflow of necessary steps for a meaningful calibration: the steps are not sufficient to guarantee physical interpretation of the results, but are yet necessary verifications to perform.

Whether the estimates of the model parameters equal the quantity of interest in the system remains hypothetical because of the threats identified so far. It means that the workflow is concededly insufficient for blind interpretation, yet necessary. If one of the preceding conditions are indeed not met, unicity or interpretability is necessarily compromised.

To bring a perspective on this workflow, literature suggests that the relationship between model validation, identifiability and interpretation can be tricky. Indeed, it is suggested in Deconinck and Roels (2017) that some models that were strictly speaking not validated performed well on physical interpretation of some of their parameters. While using stochastic RC models at wall scale, i.e. *with* a model discrepancy stochastic term, Deconinck and Roels (2017) find that some combinations of model parameters had satisfactory physical interpretability, but that the parameters individually were not statistically validated, or were even found practically poorly identifiable. The authors suggest that the unsignificant parameters rather be seen as "nuisance parameters with no physical meaning", but necessary for estimating overall thermal properties. These results illustrate therefore the not so straightforward relationship between identifiability and interpretability.

2.5 Conclusion and work prospects

This chapter has dealt with the question of proving existence and unicity of a solution to the inverse problem of thermal performance estimation in the particular perspective of poorly informative data. If the data were ideal and if the chosen model perfectly characterized the system, a simple structural identifiability verification would conclude for unicity and existence of the solution.

Instead, the issues raised in this chapter are manifold : non intrusive data leads to poor and possibly biased information. In addition, low order RC models aggregate the physics. From there arose the additional necessity to verify practical identifiability as well, i.e. numerical unicity of the calibration solution and existence of the estimates. Yet, it still cannot be considered to be a guarantee that estimated parameters are identifiable to the physical properties.

Some leads were found aiming at bridging the gap between identifiability and interpretability. A set of adequate RC models has been proposed. Collection of data can be adequately performed even in the frame of non intrusive experiments with adequate choice of input and output variables and adequate sampling. Finally, the major pitfall of model discrepancy is reduced through model selection and statistical validation.

This chapter also uncovered promising prospects that have not been studied in this work, which are listed non exhaustively in the following paragraphs.

Adequate spatial sampling of the input and output variables

Weather variables such as wind speed are not homogeneous at building scale. Indoors, air temperature can also significantly vary from one room to the other as well as in each room individually. However, spatially exhaustive measurements are usually not an option, even less so in the case of a non intrusive experiment.

Spatial sampling relates to the issue of having to choose a specific place for each measured variable, with the knowledge that there are discrepancies at building scale, but ignoring at the same time the order of magnitude of these discrepancies.

Senave et al. (2020b) study the influence of the choice of indoor air temperature measurement on the estimation of an overall Heat Transfer Coefficient by an ARX model. The authors find that a single sensor in the living room supposed to represent an homogeneous temperature across the

building leads to errors up to 30 % of the actual HTC value. Deviations up to 12 % and 6 % are observed when several sensors are used and an adequate weighing rule is used.

This issue joins a larger issue of dealing with systematic error in measurements. The systematic error is the intrinsic error of the sensor that brings a constant deviation in the measurement with respect to the actual physical quantity (Thébault, 2017). For RC models, Thébault and Bouchié (2018) propose to propagate such error after calibration by a Monte-Carlo sampling, for the reason that it cannot be accounted for in the calibration process, at least not in an identifiable model formulation. Calibration and Monte-Carlo uncertainty propagation is computationally burdensome to perform and makes it more difficult to disseminate for practitioners.

Adequate spatial sampling can therefore first be dealt with through representative measurements in situ and secondly be accounted for in the model formulation. Both aspects deserve further work, for which a Bayesian approach, even in a structurally non identifiable set up, would possibly prove interesting.

Uncertainty propagation in general with a Bayesian approach: example of η_{hea}

More generally, this chapter showed how a Bayesian approach could still be effective to calibrate structurally non identifiable models, which can otherwise not be dealt with.

Representative priors may overcome the high computational costs and possible failure of the calibration algorithms strongly suggested by the literature. As such, a Bayesian approach provides a tool for an integrated uncertainty analysis, which results in a calibration outcome with wider posterior probability distributions, hence with better representation of the actual uncertainties of the result.

For example, as mentioned in section 2.2.4, any RC model that would include a coefficient for the efficiency of its heating system is structurally non identifiable. The heating efficiency coefficient still cannot be estimated because of the non structural identifiability, but all parameters will most probably have wider probability distributions. The goal of having at least representative posterior distributions is at least obtained with less effort than adding a Monte-Carlo uncertainty propagation.

Let us underline that not all non structurally identifiable models have equal interest. From the examples of section 2.2.4, structurally non identifiable model with additional surface thermal resistances (model 2.6) is less interesting, because the error from this physical aspect of heat transfer on the calibration result is barely significant. The case of heating system efficiency is on the other hand brought forward as it is a major pitfall to overcome for better exploitation of in situ measurements.

Influence of the indoor temperature set point schedule on parameter estimation

Non intrusive measurement conditions limit the possibility to optimize the information in the collected data. Concededly, adequate sampling and adequate sensors can enhance the amount of information in the data. It is also clear that constant indoor temperature leads to very poor practical identifiability. There could then be further research done on acceptable temperature schedules for occupancy that enhance practical identifiability of RC models.

There remain indeed questions on the frequential interference of occupancy related heating schedules and weather variables, on acceptable temperature amplitudes as well as on the duration and the number of higher temperature periods, in the spirit of the literature exposed in section 2.4.2.1.

Questions to tackle

From the findings and literature review in this chapter, it becomes clear that poorly informative data and model discrepancy are major threats to physical interpretation. At the same time, there exist tests and good practice procedures to verify structural and practical identifiability as well as model satisfactory adequacy.

The relationship between practical identifiability and physical interpretation is however not as straightforward. There does not seem to be any sufficient indicator for physical interpretation to be inferred, in particular from poorly informative data.

This calls for further work on the interpretability of RC models in a non intrusive framework. The next chapter proposes therefore a numerical platform to test the interpretability of a calibration result. A comprehensive numerical building energy model serves as reference and its thermal performance is naturally known. The building energy model can then be simulated and the output serves as training data for a model, here applied to RC models.

The numerical platform then aims at answering the following questions: if there is a change in the building properties or in the building boundary conditions, are the calibration results still trustworthy? How does the physical interpretation of the calibrated RC model then vary?

3

Numerical model assessment methodology for physical interpretability

This chapter proposes a methodology to assess the physical interpretability of parameter estimates. The methodology is meant to be applied to the estimation of thermal performance by stochastic RC models. It is based on synthetic data from a detailed reference model, which target theoretical thermal performance is accurately known. An original indicator for model assessment is proposed to account for physical interpretability to the target theoretical value. To gain better understanding of the relationship between simulation conditions and building configuration to interpretability, the reference models undergoes modifications according to a design of experiment as to enable global sensitivity analysis.

Contents

3.1 Introduction

Chapter 1 established how stochastic RC models are promising models to learn from poorly informative data. Chapter 2 showed how to carefully choose the RC model to calibrate and exposed a good practice workflow for calibration from a given dataset. Concurrently, chapter 2 also laid out how model discrepancy and poorly informative data may corrupt the calibration process and bias the results.

In the perspective of using stochastic RC models to infer thermal performance estimation of a building from nonintrusive measurements, it would be desirable to assess the extent to which RC models are physically interpretable. In literature, interpretability of a model is done case by case on the basis of expert knowledge of the tested building, such as in Bauwens (2015); Chong and Menberg (2018) or on a numerical simulated building energy model as in Deconinck and Roels (2017); Senave et al. (2020a). Extrapolating the validity of the models to other test cases is risky, all the more so as the data used in non intrusive measurements poorly informative and possibly biased. In addition, exploiting actual on-site measurements brings other issues which make inference of interpretability tricky: effect of initial conditions and thermal states of the building, including thermal inertia, duration possibly not as long as necessary, limited measurements although necessary for better knowledge of the building behaviour.

This chapter proposes therefore a numerical model assessment framework aiming at:

- creating a realistic building simulation environment to synthetically generate datasets for models to be trained,
- proposing an accurate target value of thermal performance of the building envelope and its components,
- assessing the physical interpretability of any model calibrated by a robust indicator,
- comparing in a common framework the interpretability of different models,
- providing the possibility to assess the interpretability of a model in variable configurations, as to verify repeatability of a calibration and to a certain extent generalisation of the results.

Section 3.2 first exposes the principles of the framework for model assessment while detailing how the proposed framework alleviates the issues attached to the realism of numerical simulations. Section 3.2 also details the case study simulated in the framework and in the end used to generate synthetic data. Section 3.3 describes how to assess a model with respect to the physical interpretability of its thermal performance estimation. Finally, section 3.4 presents how to efficiently quantify variability of the thermal performance estimation when there is variability in the building energy reference model.

3.2 A numerical assessment framework for physical interpretability

The objective of a numerical framework for model assessment is to act as a sort of test-bed: the framework should recreate a non intrusive experiment as close to reality as possible to provide realistic synthetic datasets that serve as training data for the model under study.

To be of use, benchmark studies should avoid common pitfalls that prevent the generalization of the results (Kreutz, 2019). The recommendations in Kreutz (2019) are concededly intended for optimization algorithms but are still valid for the present application, with in particular the following pitfalls:

- relying on an unrealistic setup: ideal measurements (no systematic error, no measurement error),
- use the same model for data generation as for calibration (cf. the "inverse crime" from Kaipio and Somersalo (2007)),
- not generalisable to other building configurations and typologies.

This section describes in 3.2.1 the general principles of such framework and how it alleviates the aforementioned pitfalls. This section also details in 3.2.2 the choices made to ensure realism of the simulations. Finally, the case study used as building energy model is described in 3.2.3.

3.2.1 Proposition for a numerical assessment framework

The proposed model assessment framework is shown in Figure 3.1. A comprehensive highly detailed building is modelled in the EnergyPlus software (see Section 3.2.2). It is defined with a number of physical properties (scalars), the quantity of interest, that the model under study estimates. From a given weather dataset, EnergyPlus performs a dynamic thermal simulation from which are extracted the output necessary to the calibration of the model under study, stochastic RC models in the present application.

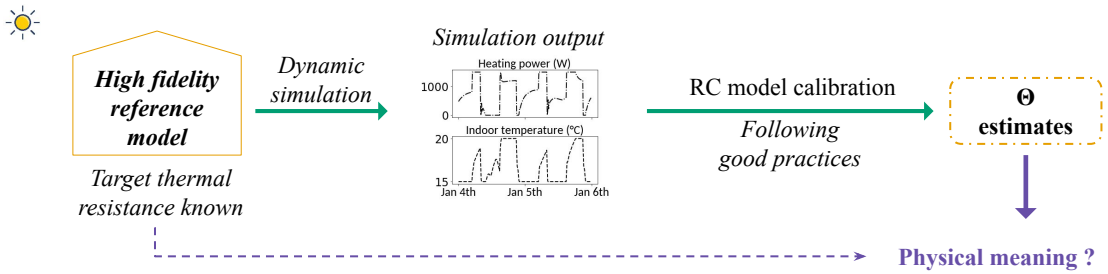


Figure 3.1 – Principle of a numerical framework for model assessment: a numerical building energy comprehensive model generates synthetic data that serves as training data for an RC stochastic model. The calibration outcome is compared to the theoretical properties of the building of the numerical model.

Model calibration delivers either sampled posterior distributions of the parameters (Bayesian approach) or parameter estimations with Gaussian assumption and flat prior (frequentist approach). Either way, the results can then be compared to the physical properties of the comprehensive model from where interpretability can be assessed. An assessment framework is detailed in section 3.3 and the calculation method for the target theoretical value is detailed in section 3.2.3.2.

The basic procedure described in Figure 3.1 assesses the interpretability of one model, but only on the basis of synthetic data from a single building energy model. Let us here underline that the model "tested" should actually be the model that best fits the generated data in the sense of the validation workflow in section 2.4.3.2. Indeed, it makes no sense to calibrate and assess a model that would not pass the validation workflow. Instead, for a given synthetic dataset, an appropriate model is selected and then assessed. This does in any case not provide generalisable outcome as it remains too case-specific.

The assessment therefore rather relies on a larger framework, where boundary conditions or physical properties of the envelope vary. The idea is illustrated in Figure 3.2. Each stochastic RC model is assessed for its physical interpretability from a coherent set of hundreds of datasets. The variations focus then either on the weather conditions or on the building envelope. The former allows a controlled repeatability assessment and the latter assesses in particular the individual

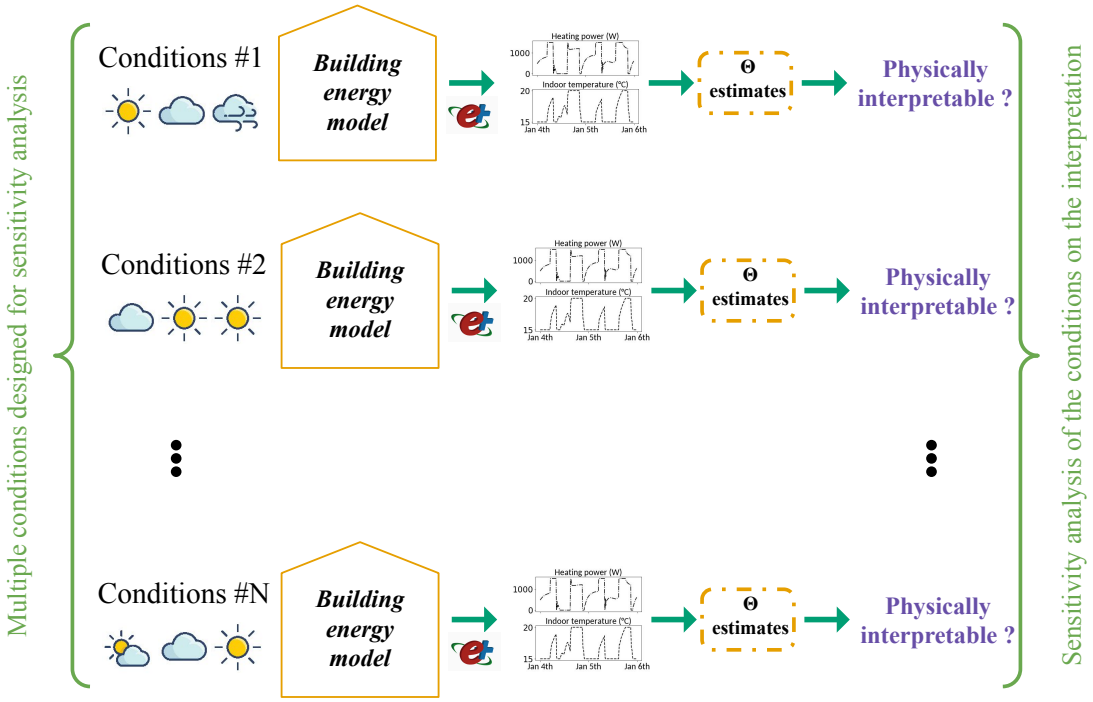


Figure 3.2 – Building energy model simulations in variable conditions: the overall framework assesses to what extent a stochastic RC model is robust and coherent to variations of the numerical building energy model. Variable simulation conditions can be weather related for repeatability assessment (as illustrated here) or can be envelope related for physical meaning of the parameters of the RC model.

physical interpretability of the model parameters. The model assessment framework has therefore better generalisation ability.

3.2.2 A comprehensive building energy model as reference

The model assessment framework relies on a comprehensive building energy computer model, from now on called *reference model*. The reference model needs to be as detailed as possible in order to depict realistic thermal dynamics.

Relying on a computer model is concededly arguable: it introduces itself a characterization error as it cannot produce data as realistic as actual data, because of discretization of the heat equation and simplifications of some phenomena. The validity of this work relies on the hypothesis that the reference model in EnergyPlus is yet detailed enough. It is indeed assumed that although not perfectly realistic, a reference model can be sufficiently detailed to study the aggregation of physical phenomena in RC models and to uncover biased results. Additionally, it does not constitute an "inverse crime", to quote Kaipio and Somersalo (2007), as the reference model is more detailed in its description of the building physics than the RC models tested.

The reference model is simulated with the simulation software EnergyPlus. The advantage of using such software is that it has been proven to be reliable for accurate thermal simulation (Crawley et al., 2001; NREL, 2016) and it has already been used for the purpose of generating synthetic data in scientific literature including recently and among others (Andrade-Cabrera et al., 2017; Foucquier et al., 2013a; Goffart et al., 2017b; Li et al., 2019). In addition, as mentioned by Tian et al. (2018), the EnergyPlus software has the advantage of using as input file a ASCII file, easily edited with many programming languages.

Various choices can be made for the reference model to account for heat and mass transfer

modelling, solar irradiation, etc. The choices made for the reference model are therefore driven by the purpose of this study, i.e. thermal behaviour, but also on the need of reasonable simulation duration as hundreds of simulations are planned for model assessment.

In the future prospect of exploiting datasets from measurements under occupancy, the indoor temperature setpoint schedules are designed to mimic occupant-like schedules.

To achieve satisfactorily thermal behaviour while keeping the simulation time reasonable, the reference model is then simulated with modelling choices detailed in the next sections:

- modelling of heat and mass transfers in 3.2.2.1,
- modelling of ventilation and infiltration in 3.2.2.2,
- solar irradiation and solar gains in 3.2.2.3,
- indoor temperature schedules in 3.2.2.4,
- simulation period and time steps in 3.2.2.5,
- addition of noise on simulation output 3.2.2.6.

3.2.2.1 Heat and mass transfer in walls

EnergyPlus offers different algorithms to implement heat and/or moisture transfers in the building:

- **Conduction Transfer Function**: this algorithm is a heat only algorithm, it does not account for moisture storage and diffusion,
- **Conduction Finite Difference**: a 1 dimensional finite difference algorithm for again heat only,
- **Moisture Penetration Depth Conduction Transfer Function** and **Combined Heat And Moisture Finite Element**: both algorithms account for heat and moisture transfers.

This work focuses on heat transfer properties of the building envelope. Using heat and moisture transfer algorithms will only slow the simulation time without adding significant improvements in the framework.

The Conduction Transfer Function algorithm is fast as it relies on a state space representation with the finite difference wall temperatures as variables (BigLadderSoftware, 2020; Seem, 1989). From the state space representation, it is possible to formulate the model output as a direct function of the input, without calculating storage and temperatures at the discretization nodes in the wall. This algorithm however relies on the assumption that the material properties remain constant.

The Conduction Finite Difference Solution algorithm has been developed to alleviate the shortcomings of the ConductionTransferFunction algorithm regarding constant material properties and calculation of temperatures inside of the walls. The heat transfer equation is then solved by a finite difference methods relying on a Crank-Nicholson or a fully implicit scheme. The algorithm uses spatial and temporal optimized discretization for each layer of material, which allows temperature calculation across the wall.

Tabares-Velasco and Griffith (2012) compared both algorithms and underline that the ConductionFiniteDifference algorithm is particularly useful for phase change material modelling. The authors also found that simulations in both algorithms achieved in overall very close.

As a conclusion, the choice of the Conduction Transfer Function module will allow faster simulations while maintaining satisfactory accuracy.

3.2.2.2 Ventilation and infiltration

The EnergyPlus software proposes two ways to account for heat transfers through ventilation and infiltration: the **DesignFlowRate** module and the **Airflow Network** model. Coupled computational fluid dynamics and building simulation is also an option as suggested by literature, but associated computational costs are prohibitive for the present application.

The Airflow Network model, based on pressure and airflow calculations with temperature and humidity calculations, has however been developed to simulate with accuracy air distribution systems and its performance, such as supply and return duct leaks, multi-zone airflows driven by outdoor wind and mechanical ventilation.

Although such detail is without question physically more accurate, it is also much more computation consuming. The DesignFlowRate module, concededly very simple, has been rather chosen for the reference model. Infiltration and ventilation flow rates are accounted for by the same module, but in separate inputs such as to enable different values.

The DesignFlowRate ventilation module Big Ladder Software (2016a) of EnergyPlus relies on equation 3.1 to calculate the airflow rate at each time step:

$$Q(t) = V_{design} \cdot F_{schedule} \cdot (A + B \cdot |T_{zone} - T_{odb}| + C \cdot WindSpeed + D \cdot WindSpeed^2) \tag{3.1}$$

where V_{design} the air flow rate (m^3/s)

F_{design} an optional schedule that can vary over time,

A, B, C and D coefficients between 0 and 1,

T_{zone} the zone indoor air temperature ($^\circ C$),

T_{odb} the outdoor dry bulb temperature ($^\circ C$).

If B, C and D were left to the default 0 value and A set to 1, the ventilation rate would be constant (except naturally if a schedule F is programmed to make it variable). The heat transfers is then only a function of the temperature difference between indoors and outdoors (see equation 1.8 in section 1.4.2). This default set up is however quite unrealistic and too close to a ventilation branch that could be implemented in a RC model, hereby creating an inverse crime. There is indeed no interest for a RC model to estimate ventilation losses generated with an identical model.

For a less caricatural ventilation model, the A, B, C and D are fixed identically as the default BLAST (EnergyPlus predecessor) (Big Ladder Software, 2016a):

$$A = 0.606, B = 0.03636, C = 0.1177, D = 0$$

3.2.2.3 Solar gains

Solar irradiation plays a large role in the building energy balance. In particular as the window blinds are maintained open in the simulations, solar irradiation entering the building through the windows come from multiple sources: direct beams with time-dependent values for each wall, diffuse irradiation by environment, reflections of direct irradiation on environment as well as indoor diffuse reflections. To account for such details, the **Full Interior And Exterior With Reflections** EnergyPlus module is used.

Let us underline that the window blinds are kept open at all times to maintain the solar gains in the heat balance. Solar gains are indeed expected to play a role on the repeatability of thermal performance estimation for future occupant-friendly experiments. In the prospect to further exploit on-site data from occupied buildings, it is therefore important to assess how solar irradiation variability influences the accuracy of the results.

3.2.2.4 Indoor temperature schedule

The indoor temperature set point schedule is designed to mimic occupant-friendly conditions to meet the objective of studying how poorly informative data influences interpretability. As suggested by section 1.2, RC models cannot adequately learn from data in which the indoor temperature is maintained constant. A realistic temperature setback is therefore scheduled and follows a usual occupant related schedule 2.4.2.1.

The indoor temperature schedule is set to reach 21 $°C$ in the morning and in the evening for workdays, and all day long during week-ends and on Wednesdays. The rest of the time, the temperature is scheduled to keep at 17 $°C$. Figure 3.3 illustrates a week of simulated indoor temperature with such schedule.

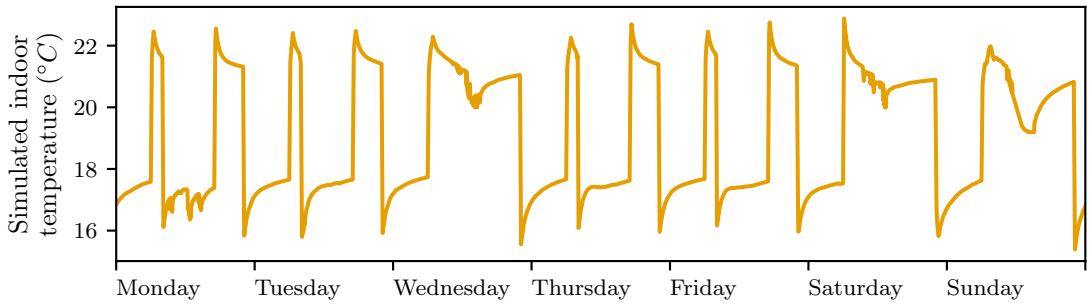


Figure 3.3 – Simulated indoor air temperature follows an occupant-friendly pattern while assuring practical identifiability (Note: the thermostat is regulated on the operative indoor temperature, which explains why the pattern is not strictly speaking rectangular shaped.)

3.2.2.5 Simulation run period and time step

As stated in chapter 2, RC models based on temperature outputs are non identifiable outside of the heating season. Indeed, a non existent heating input makes RC models at any order structurally non identifiable. The simulation run period therefore extends at maximum from November 1st to March 31st.

Furthermore, identifiability and interpretability are physically expected to be larger in the coldest months, from December to February. This is due to the larger temperature difference between indoors and outdoors and lower solar gains. This implies that the energy balance during winter days is more sensitive to the thermal performance of the building envelope.

It is however important to start the simulation period earlier as to avoid any impact of the warm up runs performed by EnergyPlus. Indeed, the software runs multiple times the same first day, beginning at an indoor temperature of 23 $°C$ until convergence of the indoor conditions is met Big Ladder Software (2016c). The warm up allows credible initial conditions for the rest of the simulation. But, as the first weather day of the simulation period is random, it might put a particular weight on a possibly unusual cold or warm day, thus misleading the first hours or days of simulation, depending on the thermal mass of the building. It is then safe to start early in the winter period and then discard the first 15 days of simulation.

As for the simulation time step, Section 2.4.2.2 established that for practical identifiability reasons, the time step should not be larger than around 15 minutes, and preferably lower in case this reference model has lower characteristic times. In addition, the EnergyPlus engineering reference recommendations suggests that the time step should be chosen depending on that of the weather data file Big Ladder Software (2016c). If the purpose is to make a simulation with an HVAC equipped model, the time step should be for example a 6^{th} of the weather data time step, i.e. every 10 minutes or less.

At the same time, smaller time steps implies continuous interpolation of the weather data if it is hourly sampled Big Ladder Software (2016b), which can be problematic. Indeed, if continuous interpolation of a smooth variable such as outdoor temperature is acceptable, interpolations for solar irradiation may be hazardous and unrealistic. Figure 3.4 illustrates the variability of global solar irradiation when sampled every minute on a partially cloudy day. Solar irradiation that would have been sampled each hour and then interpolated would yield very uncertain data.

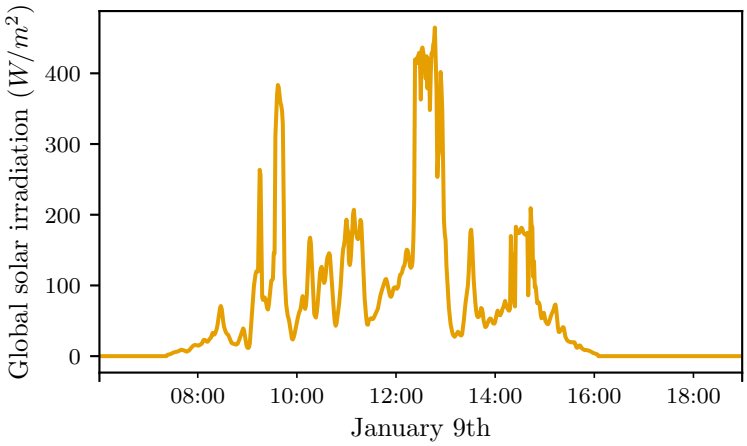


Figure 3.4 – Illustration of the global solar irradiation measured every minute on a partially cloudy day: the variability is significant and cannot be accurately interpolated from hourly samples.

As a conclusion, the simulation time step is 10 minutes for the usual hourly sampled weather data and 1 minute for the minute sampled weather data files.

3.2.2.6 Added noise to the simulation output

One of the identified pitfalls for generalisable model assessment framework is the fact that the simulation output are "ideal" measurements: there are no systematic error and no measurement error.

Accounting for the systematic error in RC modelling has been briefly mentioned in Chapter 2 as a work prospect. Such error is then deliberately not introduced in the present model assessment framework as to keep the focus on the aforementioned objectives. The present model assessment framework could however serve as test bed to evaluate RC models with controllable systematic errors.

Measurement random error are a non negligible part of the issue of solving inverse problems (Maillet et al., 2011). However, EnergyPlus outputs are deterministic. Aiming at assessing interpretability of stochastic RC models in realistic conditions and in agreement to Leroy (2010), Sengupta et al. (2015) and Stoffel et al. (2000), the present framework adds white noise to the different outputs:

- temperatures: addition of a normal noise $\mathcal{N}(0,\ 0.2\ °C)$,
- heating power: addition of a normal noise $\mathcal{N}(0,\ 20.0\ W)$,
- solar irradiation: addition of a normal noise $\mathcal{N}(0,\ 5.0\ W/m^2)$.

3.2.3 Case study

The choice for a case study to serve as reference model is driven by realism and at the same time not too complex geometry to keep the focus on heat transfers.

Preliminary applications of the model assessment framework Juricic et al. (2019, 2018a,b) were based on a BESTEST600 case, which is a simple parallelepiped with large south facing windows. Its thermal inertia is very low and it emphasizes solar gains which was not realistic and led to very case-specific conclusions.

In an effort to use a realistic case study, a multi-zone building energy model inspired by an actual building design has been chosen. The case study has a single storey. With two storeys or

more, the issue of temperature discrepancy across the building would have been inevitable. A one-storey building will be less significantly prone to temperature stratification than higher case studies.

This section details the envelope composition of the case study (in 3.2.3.1) as well as its thermal performance (in 3.2.3.2).

3.2.3.1 Details of the case study

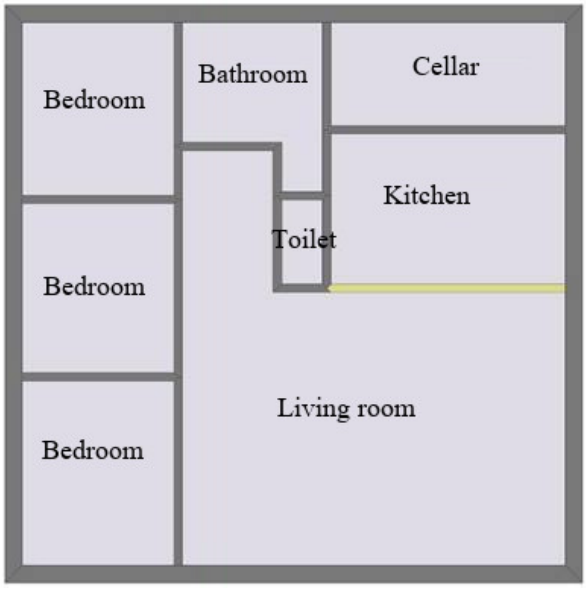


Figure 3.5 – Floor plan of the case study used as multi-zone reference model in the application of the model assessment framework

The reference model is a multi-zone building of a one-storey house. Its floor plan is shown in Figure 3.5, where each room is modelled as an independent zone in EnergyPlus. It has unheated crawlspace and attics. The heated space is about 100 m^2 and is equipped with convective heaters.

The envelope has three different type of composition, from inside to outside:

- vertical walls : 1 cm gypsum board, layer of thermal insulation material (thickness variable in applications) and 20 cm brick structure,
- attics: 1 cm gypsum board and layer of thermal insulation material (thickness variable),
- attics towards exterior: simple tiles,
- ground floor towards crawl space: concrete slab and thermal insulation material (thickness variable).

3.2.3.2 Thermal resistance properties of the case study

If actual experimental data were used, the interpretability assessment would have been less accurate. Indeed, whether based on calculations from a standard or from an experiment, the target value for thermal performance may be inaccurate and uncertain.

The advantage of basing the model assessment framework on synthetic data is that it allows to have an accurate theoretical thermal performance target value, not only of the whole envelope but also of some components of the envelope. To obviate the dynamic nature of the simulated output, the reference values are calculated from an ad hoc simulation of the reference model. This ad hoc simulation is designed to create fictive steady-state conditions with:

- a synthetic ad hoc weather data file : constant outdoor temperatures and no solar irradiation. Extraterrestrial and ground temperatures are however left as is,
- heating schedule set to maintain a constant indoor temperature in all zones.

The simulation is performed for 3 months. The energy balance of the steady state conditions thus created is described by the following linear relationship:

$$\Phi_{heating} = \frac{1}{R^*} \cdot (T_{int} - T_{ext}) + \varepsilon \tag{3.2}$$

The energy balance is incidently expressed with a thermal resistance and the target values will also be expressed with thermal resistances, in agreement with the dimension of the parameter estimates of interest of the trained RC models.

Equation 3.2 can in turn be decomposed into transmission heat transfers and transfers through infiltration and ventilation. The transmission heat transfers can also be decomposed into direct transfers to the outdoors and transfers through unheated neighbouring space when such losses are significant, hence Equation 3.3:

$$\begin{aligned} \Phi_{heating} &= (\tfrac{1}{R^*_{tr}} + \tfrac{1}{R^*_v}) \cdot (T_{int} - T_{ext}) + \varepsilon \\ &= (\tfrac{1}{R^*_{dir}} + \tfrac{1}{R^*_u} \cdot (\tfrac{T_{int} - T_u}{T_{int} - T_{ext}})) \cdot (T_{int} - T_{ext}) + \tfrac{1}{R^*_v} \cdot (T_{int} - T_{ext}) + \varepsilon \end{aligned} \tag{3.3}$$

From Equations 3.2 and 3.3, the following values are considered as target values:

R^*	overall thermal resistance (K/W)
R^*_{tr}	transmission thermal resistance (K/W)
R^*_v	ventilation and infiltration thermal resistance (K/W)
R^*_{dir}	transmission thermal resistance directly towards the exterior (K/W)
R^*_u	transmission thermal resistance through unheated neighbouring spaces (K/W)

Not all target values are strictly speaking constant over time. Ventilation and infiltration vary with temperature differences and with wind. R^*_v and R^* vary therefore too. It is however possible to infer an averaged value to be representative of the heat transfers over the considered period.

Calculation of the target values is then done on basis of the 3 months long simulation, with daily averaged data with a least square method. The fictive steady-state conditions guarantee excellent fit of the data to the linear energy balance models.

R^* is determined on basis of Equation 3.2. R^*_v and R^*_{tr} however depend on the exact same confounding variables. For each value to be determined separately, there is a need for an additional simulation where the ventilation is set to 0. The resulting difference accounts for the ventilation thermal resistance R^*_v. A similar derivation is used to separately determine R^*_{dir} from R^*_u.

Actual calculated values are given in the application chapters 4 and 5 as the composition of the reference model varies with each application.

3.3 Model assessment and comparison : a quantitative indicator

The last step of model assessment consists in assessing how close the estimated thermal performance of the building is to the theoretical target value. As the theoretical thermal performance is accurately known, it is possible to quantify the closeness of the estimation to the target value.

In Senave et al. (2020a) for example, such comparison is made with error ε to the target value defined as in equation 3.4 and with a conditional indicator, whether or not the target value is included in the confidence intervals (with a yes or no answer).

$$\varepsilon = (R_{estimated} - R_{target})/R_{target} \times 100\ (\%) \tag{3.4}$$

Applied in preliminary works Juricic et al. (2019, 2018b), such indicators proved to be limited for model comparison at large scale. Indeed, it is rather inconvenient that the assessment is made by two different values, one of which is not even scalar. Furthermore, it does not properly reflect the desirability of the results, because it compares a single maximum likelihood estimator to the target value, without reflecting on acceptable margins and uncertainty.

Instead, an alternative quantitative indicator is proposed in this work as an attempt to better reflect on the desirability of both the results' accuracy and uncertainty. It is inspired by the Bayesian approach, where parameter estimations are continuous probability densities, see Chapter 1. It can be extended to the frequentist approach with the assumption of Gaussian distribution. Let us remind that the mode is the most probable value and that the credible set is the Bayesian conceptual version of the frequentist confidence intervals.

Accuracy of the estimation to the target value can then be categorized in the 4 following cases, illustrated in Figure 3.6:

- (1) accurate mode and narrow credible set : the target value falls in the 89 % credible set (or 95 % CS), the mode is close to the target value. This category shows high accuracy and low uncertainties of the estimation. Estimations in this category are the most satisfactory results;
- (2) accurate mode and large credible set : the target value falls in the credible set, but the uncertainties are large ;
- (3) inaccurate mode and large credible set : the target value falls close to the tails of the credible set, meaning that the mode is misleadingly far from the target value but that at least the uncertainties are representative of the informativeness of the data and the model ;
- (4) inaccurate mode and narrow credible set : the least satisfactory result. The model and/or the data give bias to the parameter estimation and the target value does not fall in the credible set, which will mislead the expert examining the data.

To make model assessment and comparison easier, these 4 categories need then to be translated into a metric, which needs to be high for category (1), moderate for categories (2, 3) and low for category (4).

Intuitively, the probability mass around the target value fits this objective. For a discrete probability function, the probability mass at the target value exactly would suffice, i.e. for $X : S \mapsto R$ a discrete random variable, the probability at the target value x is:

$$P(X = x) = P(\{s \in S : X(s) = x\}) \tag{3.5}$$

This work deals with continuous probability density functions and probability density of a continuous variable at the target value is exactly *0* by definition. Instead, the probability density of the volume around the target value is an option, i.e. the probability of random variable X of density f to be between limits a and b:

$$P(a < X \leq b) = \int_a^b f(x)dx \tag{3.6}$$

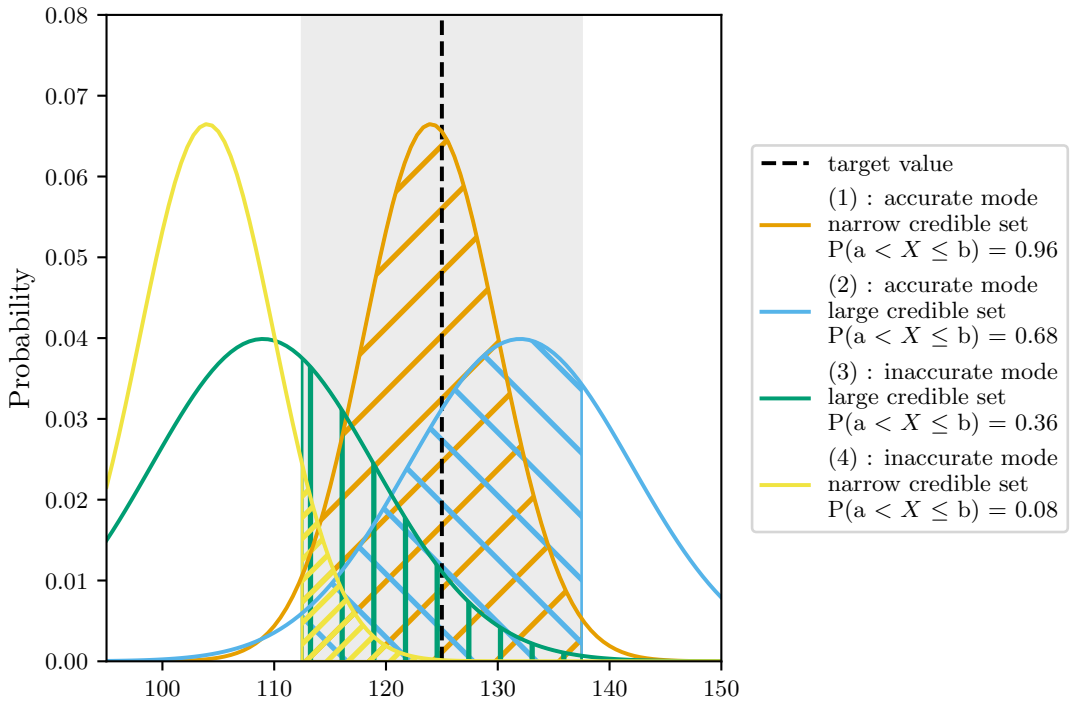


Figure 3.6 – Illustration of 4 categories of posterior distributions (synthetic data). Limits a and b defined as for example target $\pm$ *10 % create a region of interest (in grey) around the target value (dotted black line). The probability of each density to be within these limits (hatched areas) defines a quantitative indicator that takes values between 0 and 1.*

Figure 3.6 illustrates how a probability density is equivalent to calculating the integral of the posterior density function between a and b, i.e. the area under the curve between a and b being on either sides of the target value. The values taken by this metric are bound between 0 and 1, meeting then the criteria stated earlier. For a category (1) type of posterior distribution, the metric is close to 1, as in the illustration from Figure 3.6 $P(a < X \leq b) = 0.96$. For categories (2) and (3), the metric will show values around 0.5. Category (4) results are very unlikely and the metric tends to 0.

Another argument in favour of the probability density on $]a,b]$ is that the metric is standardized between 0 and 1 whatever the order of magnitude of the target value, therefore facilitating model assessment on any of its parameters. The interval $]a,b]$ is perfectly arbitrary. We propose to take as reference a 5 % acceptability error. The assessment metric becomes:

$$P(a < X \leq b) = \int_{a=target-5\%}^{b=target+5\%} f(x)dx \tag{3.7}$$

3.4 Global sensitivity analysis

A major novelty of this model assessment framework lies in the possibility to create and manage multiple configurations of the reference model, as suggested earlier in Figure 3.2. The existing literature assessing physical interpretation and based on synthetic data either relies on a single set as in Deconinck and Roels (2017) or at best on a basic design of experiments

with a one factor at a time approach as in Senave et al. (2017). Such assessment might then not be generalisable and fails to clearly uncover interactions between building configuration and environment and estimation interpretability. In the worst case scenario, a poor design of experiment can lead to erroneous conclusions (Saltelli et al., 2010), due to the limited view on the influence of the variable reference model simulations on interpretability.

In the present methodology, the interpretability assessment is not based on a single case but on an overall design of experiment as to allow global sensitivity analysis. The influence of the variable configurations on the interpretability of the parameter estimates can then be quantified and better grasped.

This section 3.4.1 first details how global sensitivity analysis allows a comprehensive, qualitative and quantitative assessment of influence of variability in the reference model on the results. Subsections 3.4.2 and 3.4.3 describe how global sensitivity analysis can be carried out for evaluation of the influence of respectively weather variables and thermal property of the reference model.

3.4.1 How to perform global sensitivity analysis

The objective of the framework is to assess how estimated model parameters are influenced to variations of simulations of a reference model, which thermal performance is known. In a sensitivity analysis semantic, the thermal properties that vary are called the input, the interpretability of the parameter estimates (or the estimates themselves) are called the output. Let us underline that "input" and "output" for a sensitivity analysis have a different meaning than for a model.

To study how the inputs influence the output, several methods are available. Local sensitivity analysis methods are immediately discarded as they do not fit the purpose of this study as they do not provide a satisfactory exploration of the parameter space (Saltelli et al., 2008).

Graphical methods, provided with the adequate input sampling, give a useful overview of the input-output relationships (Saltelli et al., 2008). Also called main effects visualisation (Iooss and Lemaître, 2015), graphical methods usefully enable to detect trends in the output data through scatter plots or cobweb plots.

As illustrated by Figure 3.7, scatter plots for example may be used to visualize first and higher order effects on the output:

- first order effects graphs show input parameter x_i against y for the entire set of x_i,
- higher order effects graphs show two inputs parameters on x- and y-axes and as a third dimension the output value (colour or marker type depending on whether the output is discrete or continuous).

Yet, graphical methods cannot provide any ranking nor quantification of the influential parameters on the studied output. To do so, Iooss and Lemaître (2015) suggests that as numerical cost is not too high, variance based methods are appropriate. The principle is to decompose the total variance of the output into a sum of the partial variances representing the marginal effect of each input parameter independently (i.e. uniquely due to the input parameter) and the effect of the interactions at second, third to m order (m the number of parameters) hence Equation 3.8.

$$V = \sum_{i=1}^{m} V_i + \sum_{i>j} V_{ij} + ... + V_{1...m} \tag{3.8}$$

Then, the sensitivity indices are the ratio of each variance to the total variance of the output 3.9. The total then necessarily equals to 1. The sum $\sum_{i=1}^{m} S_i$ relates to the first order effects only, i.e. the marginal effect of the input parameters on the output. If $\sum_{i=1}^{m} S_i$ is close to 1, the interaction effects are negligible. On the contrary, lower $\sum_{i=1}^{m} S_i$ means that the part of the variance is explained by

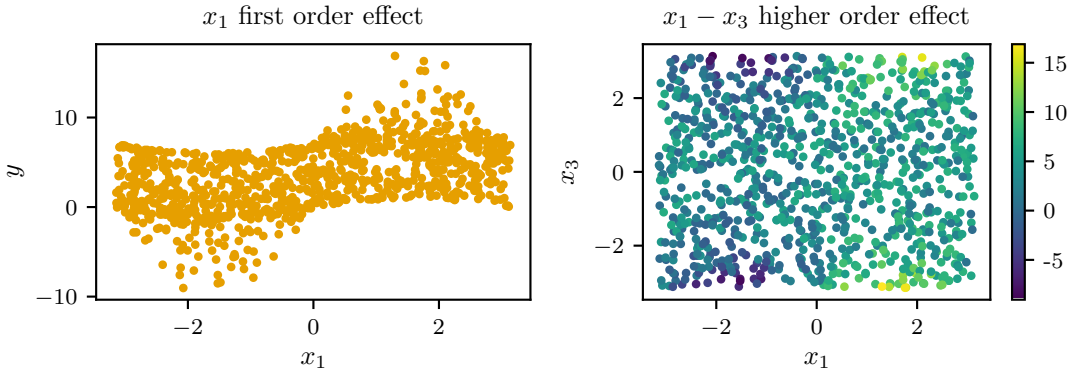


Figure 3.7 – Illustration of graphical methods for the study of a continuous output variable y compared to input variables x_1 and x_3 (application of the Ishigami function). First order indices on the left are visualized by a scatter plot of one input and the output. Higher order interactions are visualized on the right by a scatter plot of two inputs and colours for the output y.

interaction effects. Finally, S_i close to 1 means that the total variance is almost solely explained by the variance of input parameter x_i (Goffart et al., 2017a).

$$1 = \sum_{i=1}^{m} S_i + \sum_{i>j} S_{ij} + ... + S_{1...m} = \sum_{i=1}^{m} \frac{V_i}{V} + \sum_{i>j} \frac{V_{ij}}{V} + ... + \frac{V_{1...m}}{V} \tag{3.9}$$

To calculate the indices, the variance of the output is obtained through thorough sampling of the input space: although significantly costly, Monte-Carlo sampling methods enable to calculate first order and interaction indices or first order and total indices (Iooss and Lemaître, 2015), with the total indices of x_i being the sum of all order indices related to x_i. To reduce the overall numerical costs, quasi Monte-Carlo sampling has proven to achieve robust indices (Iooss and Lemaître, 2015).

Screening techniques such as the Morris method only yield a qualitative classification of parameter importance: negligible effect, large linear effect without interaction and large non-linear or interaction effect. Such methods deliver no quantitative insight into the effect of the inputs on the outputs (Goffart and Woloszyn, 2018; Iooss and Lemaître, 2015; Menberg et al., 2016; Petersen et al., 2019).

As a conclusion, variance based methods sensitivity analysis methods are chosen for the present model assessment framework. They are implemented differently, depending on whether the inputs are time series, such as weather variables, or whether the inputs are scalars, such as thermal properties of the building envelope. Both are now discussed in the following sections.

3.4.2 Assessing the influence of weather variables

The difficulty to estimate the influence of weather variables compared to thermal properties is that weather variables are time series. One cannot simply sample values of time dependent variables as is done for static inputs.

Goffart (2013); Goffart et al. (2017a) tackles this issue and proposes an original approach to generate stochastic weather data while providing suitable samples for variance based sensitivity analysis. The idea is based on the possibility to estimate the sensitivity indices with the following formula (applied to input x_1):

$$S_{x_1} = \frac{Cov(y, y^{(1)})}{\sqrt{V(y) \cdot V(y^{(1)})}} \tag{3.10}$$

where y is the output calculated from samples $(x_1, x_2, ..., x_n)$,
$y^{(1)}$ is the output calculated from samples $(x_1, x_2^*, ..., x_n^*)$ and,
x_i and x_i^* are two independent stochastic inputs identically distributed.

All it then takes is to generate $2 \times k$ independent samples of n weather variables over a certain duration (in Goffart et al. (2017a) one month). Such design will allow to calculate the indices with Equation 3.10. The strength of this method is therefore that each index reflects on one individual weather variable and completely independently of the other variables.

The weather variables are generated based on the hypothesis that a stochastic process (weather variable) $x_i(t)$ is defined by a mean $\bar{x}_i(t)$ and a stochastic variable $\varepsilon_i(t)$. The mean is considered deterministic. For weather variables, the mean is defined from representative statistics of the period of interest (winter, summer,...). Coherent stochastic variables for each stochastic process are obtained from the procedure in Iman and Conover (1982) that uses the auto-correlation and the hourly cumulative distribution function of the weather variable of interest.

An exhaustive explanation of the generation method is available in Goffart (2013) or in Goffart et al. (2017a). The representativeness of the generated variables used in application of the present framework is discussed in Chapter 4.

3.4.3 Influence of variable thermal properties of the envelope

The overall goal of this application of the model assessment methodology is to put into practice the idea that a change in the building thermal reference model should translate into a change in the ad hoc parameter of the low order model. For example, a change of glass wool thickness in the walls of the reference model should mean that at least the overall thermal resistance of the trained RC model should vary proportionally.

In this case, the sensitivity analysis will first rely on samples of different thermal properties of the envelope and secondly rely on an adequate sensitivity analysis (SA) method. SA method is detailed in section 3.4.3.1 and sampling in 3.4.3.2. Convergence and uncertainty of SA are finally presented in 3.4.3.3.

3.4.3.1 Principles of the RBD-FAST method

The general presentation of sensitivity analysis in section 3.4.1 drove to the choice of variance based methods as they allow quantification of the effects on the interpretability. Section 3.4.1 also basically explained how variance based sensitivity indices were calculated. However, such calculations rely on a Monte-Carlo sampling and are therefore computationally demanding for convergence of the calculation to be achieved, all the more when several thermal properties are studied.

To obviate this problem, Iooss and Lemaître (2015) and Goffart et al. (2017a) suggest to use the Fourier Amplitude Sensibility Test (FAST) method, that uses a Fourier transform to calculate the indices on basis of a multi-frequential sampling of the studied inputs (here the thermal properties of the envelope).

The Fourier Amplitude Sensitivity Analysis (FAST) method relies on a particular design of experiments of the input space: each input variable x_i is sampled according to a periodic function at frequency ω_i, so that the output y is periodic too. A Fast Fourier Transform allows a frequency analysis of the output y and enables to pinpoint its frequencies to the frequencies of the input parameters.

The principle is to calculate the Fourier spectrum $F(.)$ at the first M discrete harmonics of the reordered output y^R. If N is the number of samples, the computation of the partial variance of input x_i goes as follows:

$$\hat{V}_i = V[E(y|x_i)] = \sum_{l=1}^{M} F(l) = \sum_{l=1}^{M} \left| \frac{1}{\pi} \sum_{j=1}^{N} y^R(s_j) exp(-i\omega s_j) \right|^2 \tag{3.11}$$

The sensitivity index follows straightforwardly by dividing the partial variance thus calculated by the total variance of output y.

As interference and aliasing due to the choices of frequencies may bias the result of the sensitivity analysis (Tissot and Prieur, 2012), the frequency analysis has later on been done on a different design of experiments, the Random Balance Design (RBD) (Tarantola et al., 2006). Instead of sampling each parameter at a different frequency ω_i, all are sampled to an identical frequency ω. In a nutshell, RBD is close to a Latin Hypercube Sampling (see 3.4.3.2) except that the points are taken in the center of each interval (Tissot and Prieur, 2012).

Additionally, Plischke (2010) proposes an algorithm, EASI, that post-processes any initial random or pseudo-random sampling of the input space: each input variable x_i is in turn sorted into a triangular shape. The output undergoes the same permutation allowing then the FAST analysis for the estimation of the sensitivity indices. This sorting principle is illustrated in Figure 3.8. All input and output variables are sorted according to the permutations that sort input x_1 into a triangular shape . Other inputs x_2 and x_3 still remain pseudo-random, which is due to the pseudo-random nature of the input space sampling. Sorted output y however shows a particular behaviour, as it is partially influenced by x_1. Sorted output y may then proceed to the Fourier frequency analysis to quantify the effect of x_1.

Figure 3.9 then illustrates the principle of the Fourier frequency analysis on an EASI-RBD based sampling. The data comes from the Ishigami function usually used as test function for sensitivity analysis methods. The Ishigami function takes 3 inputs which sensitivity indices are analytically known. Figure 3.9 shows the first 15 harmonics and the amplitude of the output y when sorted according to input x_1. The sum of the first $M = 10$ harmonics determines the sensitivity index. An influential input variable such as in Figure 3.9 has large amplitudes for the first harmonics. On the contrary, non influential input parameters have equally low amplitudes for any frequency.

Plischke (2010) showed that the EASI RBD-FAST algorithm performs as well as the most efficient of the FAST and RBD methods but offers the advantage of using low discrepancy samplings too, such as a Latin Hypercube Sampling. Plischke (2010) underlines though that the indices calculated with EASI, as well as any index calculated with a FAST or an RBD method, is biased by the noise from the samples. This noise produces insignificant yet non null Fourier coefficients. As the estimation of the sensitivity indices must be performed on a finite sum that stops at harmonic M, the sensitivity indices do not converge when M grows. Figure 3.10 shows in dotted orange line how the estimation of the sensitivity index grows.

To avoid this bias, Tissot and Prieur (2012) have derived that the correction proposed by Plischke (2010) indeed allows a robust sensitivity index estimation. The correction showed in Equation 3.12 consists in subtracting a fraction $\lambda = 2M/N$ of the sum of the remaining unbiased coefficients $1 - S_i^c$ to the biased estimator of the sensitivity index S_i. The impact of the correction by Tissot and Prieur (2012) is shown in Figure 3.10. The resulting correction for calculation of the indices is straightforward:

$$\begin{aligned} \hat{S}_i^c &= \hat{S}_i - \frac{2M}{N}(1 - \hat{S}_i^c) \\ \hat{S}_i^c &= \hat{S}_i - \frac{\lambda}{1-\lambda}(1 - \hat{S}_i) \end{aligned} \tag{3.12}$$

Let us underline here that the bias correction yields more accurate sensitivity indices for influential input variables, but non influential parameters may then have negative indices. This

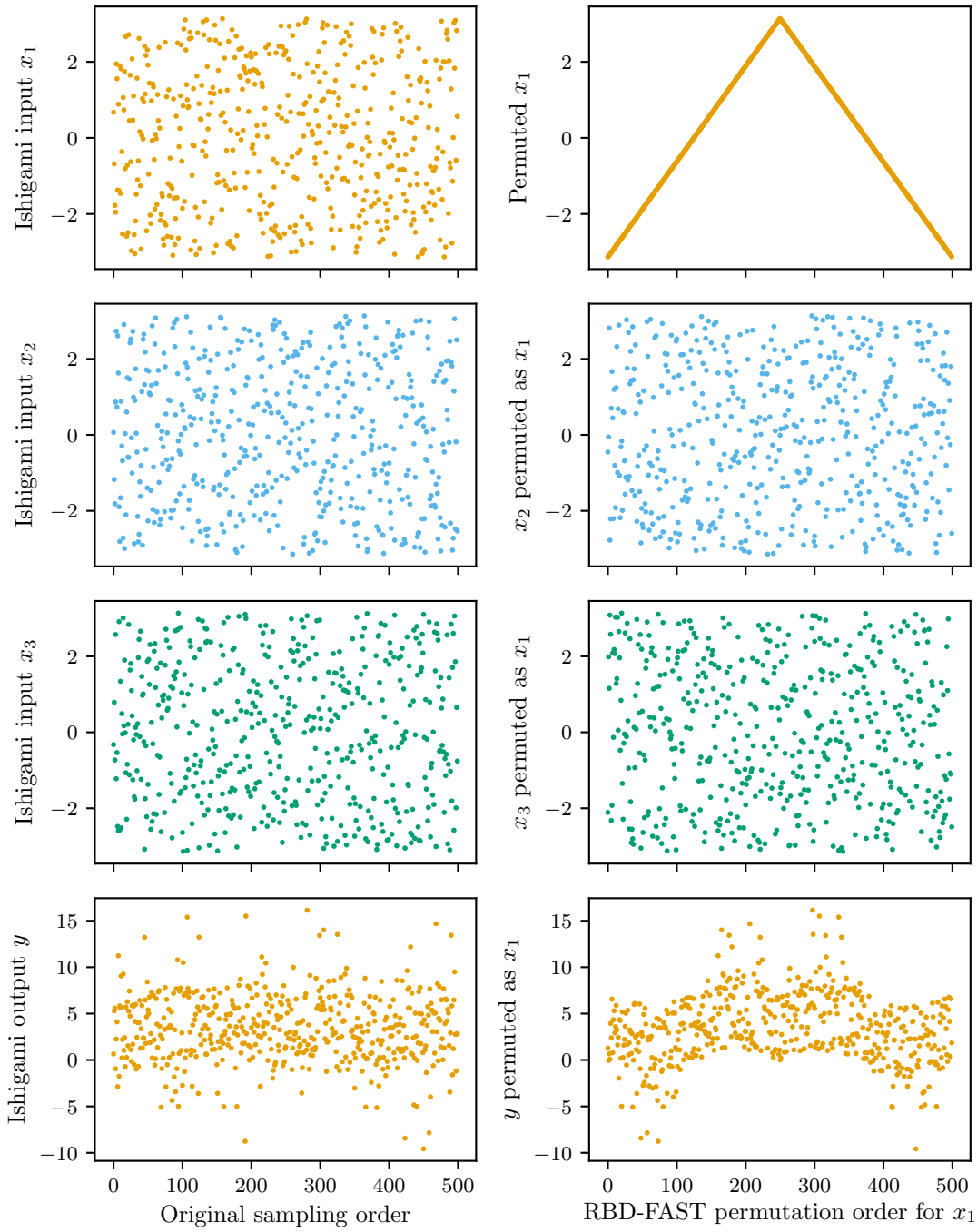


Figure 3.8 – Illustration of the sampling and permutation of the RBD-FAST method with the Ishigami function: the three inputs are sampled with a Latin Hypercube Ssampling and one by one, the inputs and the output is reordered as to form a triangular signal

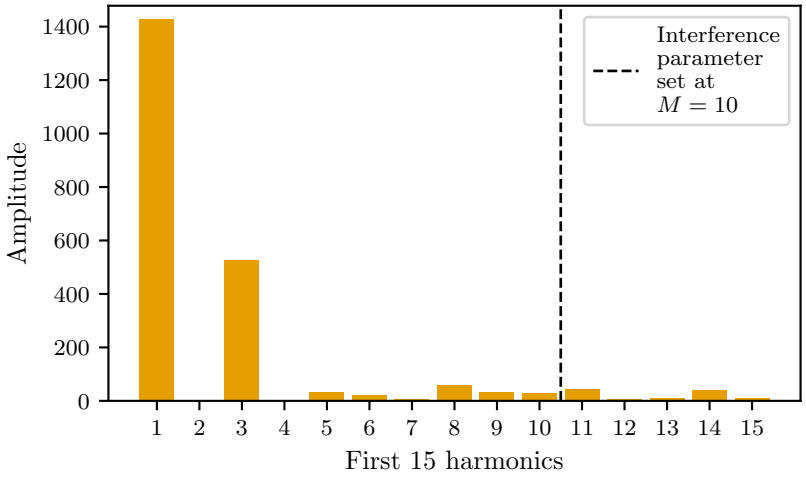


Figure 3.9 – Illustration of the Fourier frequency analysis for the Ishigami function, sorted for input parameter x_1: the first $M = 10$ amplitude coefficients are used for the index calculation (in application of an EASI RBD sampling)

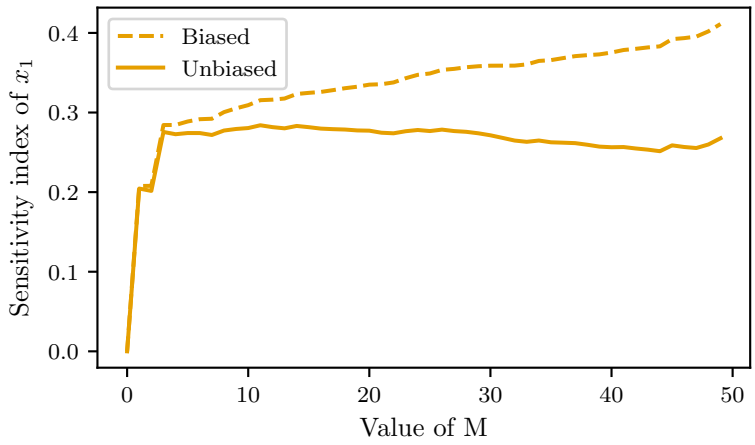


Figure 3.10 – Illustration of correction of the bias introduced by the FAST method: if unbiased, the larger M, the more the sensitivity index grows, as it takes noise related coefficients into account

is due to the fact that their indices were already almost null. Negative indices may therefore be considered as null and in any case insignificant.

This work has been an opportunity to include in the python package for sensitivity analysis SAlib (Herman and Usher, 2017) the EASI RBD-FAST algorithm, which is now conveniently available and open-source in Python.

3.4.3.2 Efficient input variables space coverage by a Latin Hypercube Sampling

The choice of the EASI RBD-FAST analysis method has the advantage to rely on an independent set of samples of the input variables, in the sense that the samples need not to be drawn following a particular design just for the analysis. On the contrary, the EASI algorithm handles any random or pseudo-random samples.

Classical Monte-Carlo random sampling can be used but will only provide reliable indices estimates with a large number of samples (Goffart et al., 2017b). Instead, the efficiency of the space coverage can be enhanced by maximizing the exploration. Algorithms such as Latin

Hypercube Sampling (LHS) or low discrepancy sequences provides such maximized exploration (Janssen, 2013; Saltelli and Annoni, 2010; Saltelli et al., 2008).

The LHS on particular is easy to implement (Goffart et al., 2017a). In a one dimensional space, the principle of LHS is that instead of randomly create N samples over the entire space, samples are picked randomly in each of the N equiprobable intervals constructed from the cumulated distribution. Brought to a multi-dimensional space, the samples for each input are randomized and combined. Samples are therefore randomly located within specified equally probable subvolumes of space (Saltelli et al., 2008). In the end, as illustrated in figure 3.11, for a given number of samples, the LHS sampling covers faster the entire space than random samples. In other words, for a given number of samples, sensitivity analysis will be more robust and achieve convergence faster if a LHS is used.

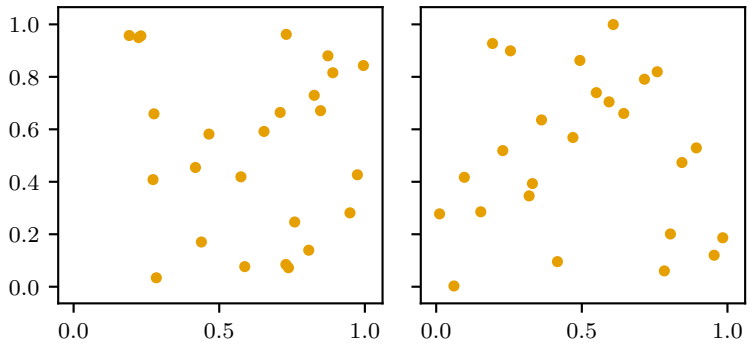


Figure 3.11 – Space coverage for random samples and LHS samples. Given 25 draws, LHS covers almost all regions in a two dimensional space whereas random samples may miss part of it.

3.4.3.3 Convergence and uncertainty

Regardless of the method chosen to perform the sensitivity analysis (SA), the indices are merely estimates themselves. As with any estimates comes the question of their accuracy and how the sample size influences accuracy. Convergence and uncertainty analysis therefore assesses the extent to which the values of the estimated sensitivity indices may be trusted.

The principle of an analysis of convergence is that drawing additional samples should not change the values of the sensitivity indices. The number of samples is large enough and the samples are sufficiently representative of the input space so that drawing more samples does not improve the knowledge of the input output sensitivity relationship.

For a given sample drawn with a Monte-Carlo technique, it is possible to assert the convergence through drawing multiple sub-samples of growing size and calculate the sensitivity indices of each sub-sample. An illustration is shown in Figure 3.14 with the convergence of the sensitivity index of input variable x_1 of the Ishigami function. Low size samples (less than 250) show extreme variability as they are not representative of the input space. As the number of samples grows, the sensitivity indices convergences towards around 0.3. From a number of 500 samples, the indices do not significantly change, meaning that convergence has been reached with a sample size of 500.

At the same time, an estimation of the confidence bounds of the sensitivity indices is necessary. The indices would be uncertain if within the set of samples, some yield dramatically different outputs than the rest of the samples, thus significantly changing the sensitivity index. The index cannot then be representative of the average influence of the inputs on the outputs. The issue is obviously linked to that of the convergence. If a sample size does not allow convergence, the confidence bounds estimation will yield significant variability. Performing it alone is therefore not sufficient on its own.

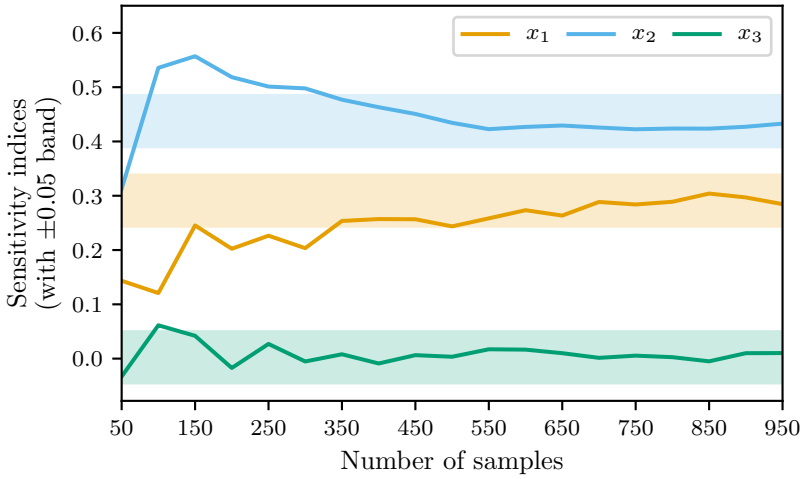


Figure 3.12 – Illustration of the convergence analysis of the sensitivity analysis of the Ishigami function: the analysis is carried out on sub-samples of growing size. If from a certain size and for all input variables the indices are stable within a ± 0.05 *band, convergence is achieved.*

Sensitivity analysis of multiple new draws of the same size is an effective and robust evaluation of the uncertainty. However, it requires many runs of the model or the procedure. If each run is costly, this technique is not affordable. The model assessment framework that includes both simulation of a reference model and model training can be considered as rather costly. Drawing many more samples to grasp the uncertainty of the SA indices is not preferable.

A bootstrap technique may instead be used and does not require new samples and model runs. It consists in drawing multiple times N samples with replacement from the initial draw of N samples. If some samples weight heavily on the output, drawing them multiple times thanks to the draw with replacement will also reflect heavily on the sensitivity index. 1000 new draws with replacement will then give a certain variability, which indicates uncertainty of the index. A certain variability is always to be expected though. Large uncertainty bounds give however a sense of the extent to which the value itself is to be trusted.

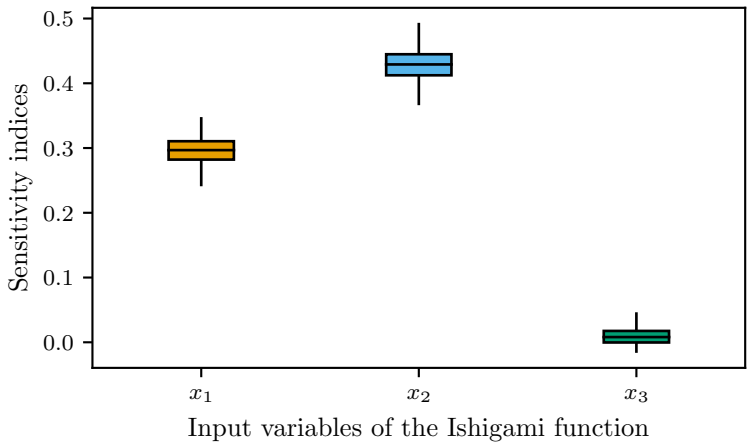


Figure 3.13 – Illustration of a bootstrap technique: random draws with replacement are done among the original samples (size $N = 1000$*) to recreate a set of samples with same total size (1000). This procedure is performed multiple times (say 500) and for each of the 500 artificial sets, the sensitivity indices are calculated. Their variability is shown as boxplots.*

In the illustration given in Figure 3.13, the bootstrap technique is performed on the 1000 samples initially drawn. The convergence has been established in Figure 3.14. The remaining variability is therefore an indication that the sensitivity indices give a general trend. Input variable x_1 (resp. x_2 and x_3) is responsible for $0.31(\sigma = 0.02)$ (resp. $0.45(\sigma = 0.02)$ and $0.01(\sigma = 0.01)$) of the variability of the output. Input variables x_1 and x_2 may be ranked in order: x_2 is more influential on x_1 and both have significant influence on the output. Noteworthy, input variable x_3 has no significant influence: although non null, its sensitivity index shows enough variability for it to be considered insignificant.

Figure 3.14 shows how uncertainty and convergence on a set of existing samples interfere. As seen earlier, 500 samples were sufficient for convergence. For subsamples of size lower than 500, the uncertainty of the indices from the bootstrap technique (the many grey lines) are quite high. From sub samples of size larger than 500, the overall uncertainty provided by the bootstrap barely diminishes, which illustrate a residual uncertainty of the estimation of the SA indices.

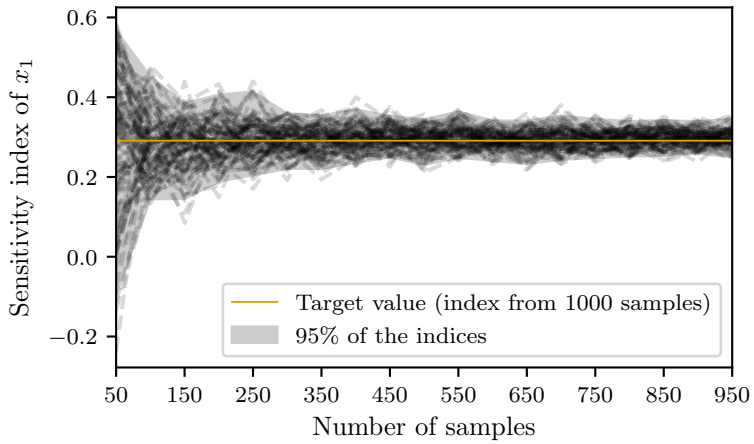


Figure 3.14 – Mixed bootstrap and convergence analysis: from a given set of 1000 samples, 100 sensitivity indices are calculated for subsets of sizes in $[50, 950]$.

Naturally, sub-sampling alters the random or pseudo-random nature of the original set of samples, in particular with low discrepancy sampling algorithms such as the LHS, where coverage of the input space is meant to be efficient. This implies that the uncertainty bounds established by subsampling an existing set will very probably be larger than if new samples had been drawn. Yet, bootstrap uncertainties may still be considered as benchmark values: the actual uncertainty of the sensitivity indices lies most likely within the bounds.

3.4.4 Conclusion on global sensitivity analysis

Using global sensitivity analysis in the model assessment framework allows to efficiently study a wide coverage of the variability of quantities of interest, either weather conditions or thermal properties of the envelope. This section uncovered indeed that variance based methods achieve both global quantification and qualification of the interactions between simulation conditions and physical interpretability of RC models. Variance based analysis can be performed to account for both weather conditions variability and thermal properties variability. The former is based on stochastic synthetic weather variables, generated as to perform sensitivity analysis. The latter uses a Latin Hypercube Sampling of thermal properties and the EASI RBD-FAST method to infer the sensitivity indices.

3.5 Conclusion

This chapter presented the methodology designed to assess the interpretability of estimated parameters of a trained model. It is designed following the idea that a model will be trustworthy if it yields repeatable results under variable weather conditions and that a model is physically interpretable if a change in the envelope means a coherent change in parameter estimates.

The developed framework is therefore based on a reference building energy model that produces simulated data. The simulated data serves as training data for the model which identifiability is under study. Interpretability is then assessed by a quantitative ad hoc scalar indicator, reflecting on both accuracy and uncertainty of the result, with respect to the target physical value. As such, the chain produces an interpretability assessment from a single configuration.

This single assessment is then wrapped up in a global sensitivity analysis, where multiple configurations are designed to follow a particular scheme in order to perform a global sensitivity analysis. In overall, the complete framework may thoroughly study how interpretability is influenced, quantitatively and qualitatively.

The scope of possible applications of this methodology is large. The next 2 chapters will focus on the two following aspects:

- **repeatability of the estimation of an overall thermal resistance** of a building envelope in chapter 4. It is clear that non intrusive measurement conditions provide poorly informative data. In this case, the heating power is not an optimal solicitation for heat transfer in the building. Weather conditions then play a larger role in the energy balance than in controlled experiments. The question is then not only to assess the extent to which weather conditions bias the accuracy of the estimation, but also to determine if there exists a minimal measurement duration that allows a repeatable estimation, regardless of the outdoor conditions.
- **detailed diagnosis of the thermal properties of a building envelope**: a detailed and accurate diagnosis of the building envelope could better drive the effort of thermal retrofit of a building. chapter 5 applies the model assessment framework to determine the extent to which individual estimated parameters are physically interpretable by comparison to theoretical values of the reference model.

Assessment of model estimation robustness to occupancy perturbation is not done in this work but would be a logical work prospect. Indeed, occupancy is expected to widen the model discrepancy gap if the RC model does not account for it. Very basically, influence of occupancy can be categorized in three cases:

- constant energy use by occupant related equipment: it introduces a difficulty to determine the precise energy used for heating. It also might produce unmeasured internal heat gains. These elements are present when the building is in use even when occupants are absent.
- random heat and moisture production by occupants, due for example to cooking or showering. These perturbations produce unmeasured and almost unpredictable internal heat gains that again affect the energy balance.
- occupants' actions on the building envelope: occupants are constantly in interaction with the building envelope: briefly opening a window for natural ventilation, leaving indoor doors arbitrarily closed or open, which affects airflows trough the building. These actions infer dynamics that are not accounted for in a usual stochastic RC model.

These elements should logically lead to very large errors on the thermal performance estimation because of systematic errors on the measured variables and because of major model

discrepancies. A prospect could first be to at least assess the extent to which thermal performance estimation is robust to occupancy. Secondly, a Bayesian calibration in the sense of Kennedy and O'Hagan (2001) or Gaussian process latent force models Sarkka et al. (2019) could account for the stochastic aspect of occupancy and maybe achieve satisfactory thermal performance estimation, or at least produce representative uncertainties of the result. Additional measured variables might in any case be beneficial for enhanced identifiability, as the uncertainties grow in experiments under occupancy.

4

Repeatability of parameter estimation under variable weather conditions

In an uncontrolled measurement framework, weather conditions have a significant influence on the estimation of the building's overall thermal resistance. This influence questions the repeatability of an experiment in a non intrusive experimental campaign. The model assessment methodology defined in the previous chapter is applied by running the reference model multiple times under variable weather conditions, such as to enable a global sensitivity analysis. A minimal observation length of 11 days for the repeatability of estimations is established. A sensitivity analysis on the estimated thermal resistance identifies the most influential weather variables. Outdoor temperature and wind speed are found to play a major role in the remaining variability for this particular case study.

Contents

4.1 Introduction

The accuracy of thermal performance estimation is a function of the information carried in the dataset: information carried in the input variables (heating, outdoor weather conditions) and the system response, indoor temperature. The dynamic nature of the indoor temperature is therefore driven by the heating and the outdoor weather conditions.

In a non intrusive experiment, the heating input is itself conditioned by the indoor temperature set point which remains user-friendly as to be non intrusive with respect to possible occupancy. The heating power is therefore not as informative as in a controlled experiment.

The dynamics provided by the outdoor conditions probably play a larger role in the dynamics of indoor temperatures than in controlled experiments. As the effect of weather conditions grows larger, it is possible that the calibration results depend on the actual conditions during the experiment. Poorly informative heating input raises therefore the issue of weather dependency of the results. If then the results are dependent on the outdoor conditions, results are not repeatable.

Naturally, the longer the dataset, the more stable the estimation should be. Yet, when a few days are considered sufficient in a controlled experiment, the question of the minimal duration of experiment in a poorly informative framework remains unanswered.

This chapter proposes to deal with the question of repeatability of the heat transfer estimation, regarding the weather variability. A first short review in section 4.2 lays out how literature handled the study of boundary conditions influence on the estimation's accuracy. This chapter then proposes to apply the methodology described in Chapter 3 to study the influence of weather variability. The chapter explores in section 4.3 how to efficiently vary weather conditions. Then, section 4.4 establishes a minimal duration for repeatable estimation of the thermal performance. Finally in section 4.5, the chapter uncovers the influential weather variables, which gives indication on the improvements to do on the RC model.

4.2 Weather conditions influence: state of the art

Previous work on thermal performance estimation of the envelope in a non intrusive framework has focused on the study of a single wall performance or on the entire envelope of a building. Rasooli and Itard (2018) show how solar irradiation significantly defers stability and convergence of the estimation of the conductive thermal resistance of a wall R_c, i.e. the inverse of the HTC of the wall, using heat flow meters following the ISO 9869 standard (ISO 9869-1, 2014). The authors suggest using flow meters on both sides of the walls to secure a robust and faster estimation of R_c.

Petojević et al. (2018) propose an innovative method to exploit non intrusive data from heat flux and temperature meters to determine dynamic thermal characteristics of a wall. The use of 12,5 days data, although not justified, met acceptable accuracy on the results.

Gaspar et al. (2018) precisely study the duration of a heat flux meter test for estimating the U-value, again following the ISO 9869-1 (2014), and compares its stability to the criteria given by the standard, in which is given that three conditions must be met simultaneously to end the test (Gaspar et al., 2018):

- the first condition is that the test must last 72 h or longer,
- the second condition is that the U-value obtained at the end of the test must not deviate more than 5 % from the value obtained 24 h earlier,
- the third condition is that the U-value obtained from the first N days and from the last N days must not deviate more than 5 %, with N = 2/3 · total duration.

There seems to be no limitation to applying the last two conditions to other calibration techniques than the methods defined in the ISO, although Gori et al. (2018) mention that it has never been seen in literature and that it might be too conservative.

In an application of RC models on heat flux measurements, Gori and Elwell (2018) and Gori et al. (2018) introduce the idea of stabilisation of the estimation : from short datasets, the estimates suffer from the prominent noise in the data. As the dataset grows, the values stabilise towards a final value. Applying the criteria of the ISO 9869-1:2014 standard, they found that up to 10 days were necessary to reach stabilisation in autumn and winter season whereas longer periods were necessary in warmer seasons. The minimum length tested was 3 days, as demanded by the ISO standard, but authors found that shorter datasets sufficed in some cases with the use of a dynamic model.

At wall scale again, Rodler et al. (2019) compared a dynamic model calibrated by Bayesian inference to the average and dynamic methods described in the ISO 9869-1:2014 standard and found that the temperature difference was more determinant than the length of the dataset, thus uncovering the major role played by uncontrolled boundary conditions.

Reddy et al. (1999) point out the issue of data informativeness in a study where non intrusive measurements are used to assess building overall heat loss and overall ventilation rate of a large commercial building using a steady-state equation. They found that daily averaged data over a year combined with a multi-step regression technique, where multiple regressions are performed one after the other to estimate parameters one by one, achieved the best results. Parameter identification over a single season was less accurate: in winter and summer seasons, the combined variability of the outdoor temperature and the relative humidity was narrower than during the spring season. Large variability of these two weather variables yielded less correlated parameters and more accurate overall parameter identification.

Deconinck and Roels (2017) applied dynamic grey box modelling in a non intrusive framework to assess the thermal performance of a single wall based on heat flux measurements. The authors used two different data subsets of 10 days in winter (steady indoor temperature assumed at 20 °C) and 9 days in summer (free floating indoor and outdoor temperatures). They found that winter conditions with constant indoor temperatures were not appropriate to identify the parameters of interest, considering that temperatures are the main variables of the differential equations used for the exploitation of the data. Summer free floating conditions were then found to be more informative and led to identifiable and interpretable parameters. Let us remind here that identifiability relates to the unicity of the parameter estimation and interpretability to the ability to give the estimation a physical meaning. Both may be confounded if the model characterizes perfectly the system.

More recently, Senave et al. (2019) studied the physical interpretation of ARX models, aiming for the estimation of the HTC via on-board monitoring, i.e. in a non intrusive measurement framework. Four different indoor temperature scenarii were tested through 20 days of synthetic data twice: once for training and once for validation. The building modelled in TRNSYS is a single-zone opaque box and the study focused on the estimation of the HTC in case of heat losses to the ground. There is no mention of the influence of the length of the dataset on the results, but in order to fit at best the available data, a model selection process is applied to check that the residuals can be considered as white noise and that the parameters of the ARX models are significant through a marginal t-test.

4.3 The reference model undergoes variable weather conditions

Using a numerical building energy model presents the advantage of knowing the exact thermal performance of the case study. The simulation environment allows to perform simulations under any weather conditions. This section presents different ways to consider such variations and why in the end stochastically generated weather data has been chosen. The section also describes the dataset selection and details on the calibration of the RC model.

4.3.1 Adaptations of the reference model methodology

The case study described in chapter 3 is used to produce synthetic datasets. Its thermal properties in this specific application are described in Table 4.1. To focus on the study of weather variability and avoid effect of heat losses through unheated adjacent spaces (attics and crawl space), the insulation thicknesses of horizontal walls has been set at 30 cm, which basically ensures very low heat transfers. At the same time, vertical walls have medium insulation and the air change rate is quite large.

Vertical insulation thickness	10 cm
Attic insulation thickness	30 cm
Ground floor slab insulation thickness	30 cm
Air change rate	1.0 h^{-1}

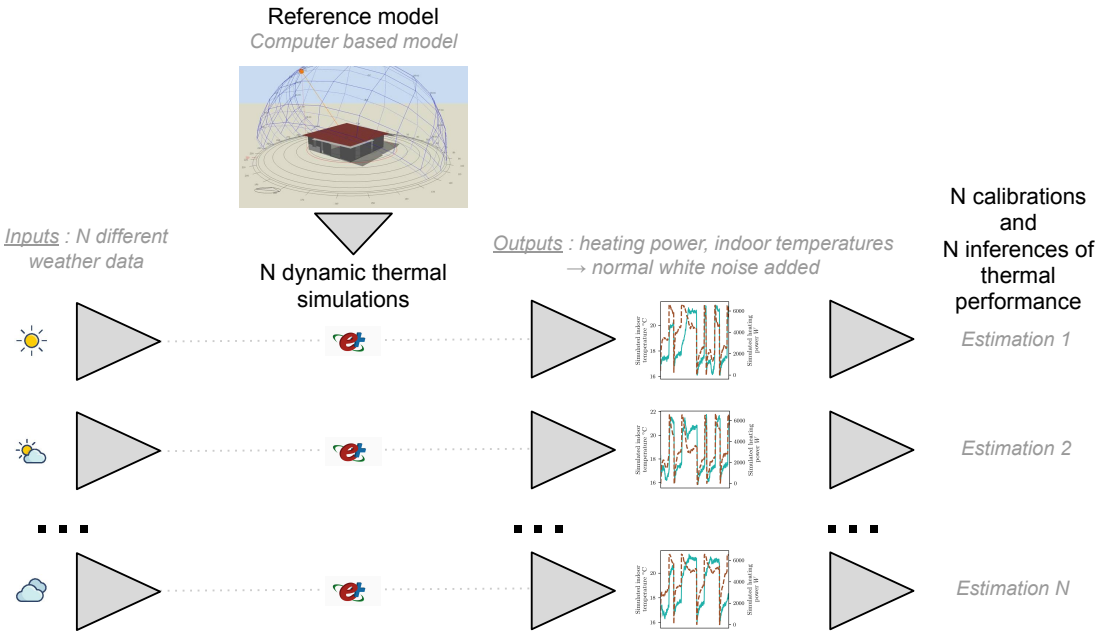
Table 4.1 – Thermal characteristics of the numerical building energy model used in this application

The theoretical *target* thermal resistance of this reference model is calculated by the methodology described in the previous chapter, section 3.2.3.2. This case study configuration yields an overall theoretical thermal resistance $R^*_{eq} = 5.19 \times 10^{-3}$ K/W. This theoretical value is later on called ***target value*** R^*_{eq}, in opposition to the estimated values of the overall thermal resistance R_{eq} (without *), estimated from the "collected" data.

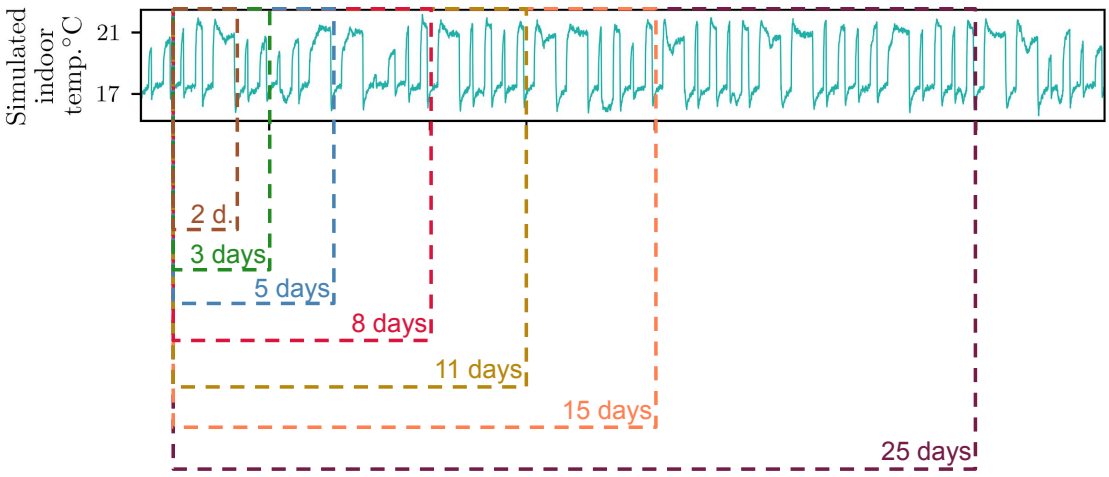
In essence, as illustrated by Figure 4.1a, the case study is simulated with N different weather conditions datasets. As a result, there are N synthetic simulation outputs, from which indoor temperature and heating power are saved. Each synthetic dataset serves as training data for a stochastic RC model. Finally, an overall thermal resistance can be inferred from the estimated parameters of the RC model.

Noteworthy, the data is produced at a time step of 10 minutes, which should be sufficient for practical identifiability (see subsubsection 2.4.2.2).

To explore adequate experiment duration under variable weather, each synthetic dataset gives data subsets of growing length, each serving as training data. As shown in Figure 4.1b, starting at date $\mathscr{D}$, subsets of 2, 3, 5, 8, 11, 15 and 25 days are considered for model training.



(a) Overall framework for the study of weather conditions influence



(b) Selection of data subsets from each synthetic simulated dataset

Figure 4.1 – Framework for repeatability assessment

4.3.2 Calibration and model validation

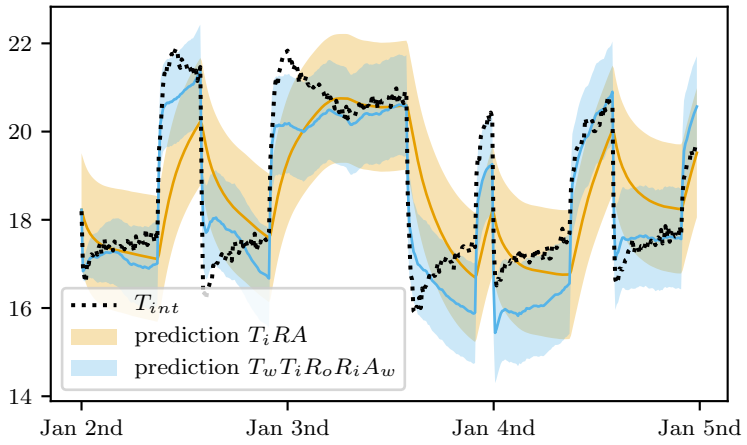
For a given synthetic dataset, the frequentist approach has been chosen for model training. Concededly, a Bayesian approach would have made more sense in a poorly informative framework, as it concatenates all available information on the thermal performance of the building (section 1.5). Yet, seeing that a very large number of calibrations is performed in this application, the associated computations would be too costly. Indeed, for illustration, on a 40 cores computation server, a single Bayesian estimation with 4 chains running in parallel with 15 days data at 10 minutes time steps takes several hours whereas a frequentist approach a few minutes.

At the same time, regardless of the approach, this chapter studies how weather variability influences experimental repeatability. This study therefore deals with the information gained from the *data only*. Under this angle, a frequentist approach makes as much sense as a Bayesian approach.

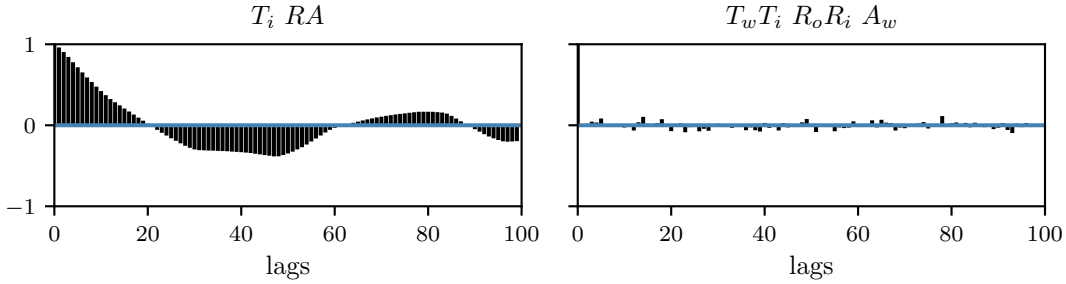
Given the available techniques reviewed in Chapter 1 Section 1.5, stochastic RC model training is performed by a quasi-Newton optimization using the BFGS algorithm in the PySIP python library (Raillon et al., 2019).

Good practice then demands that model selection be made as to infer results from one best fitting the data (see in Chapter 2 sections 2.4.3.1 and 2.4.3.2). The best fitting model might be different between short and long duration datasets. For this reason, even if model selection is performed on a 3 days training, the residuals of prediction of a 15 days training are verified as well, to ensure the selected model still performs well on larger datasets.

First order RC model may be quickly discarded as they visibly fail to catch the physics compared to a second order model, as can be seen in Figure 4.2a. The residuals of a first order model are highly auto-correlated, see Figure 4.2b, which again is proof that such a model does not correctly fit the data. A second order model's residuals are indeed much closer to white noise for 3 days training. When a 15 days training is performed, residuals of prediction show that the first order model is still highly auto-correlated and that the second order model still performs well.



(a) Non filtered prediction on a 3 days training set



(b) Autocorrelation of the residuals of filtered predictions of models T_i RA and $T_w T_i$ $R_o R_i$ A_w for 3 days data

Figure 4.2 – Graphical visualisation of how well models T_i RA and $T_w T_i$ $R_o R_i$ A_w fit the data

As an additionnal step in the model selection, more complex models have been tested: model $T_w T_i$ $R_o R_i$ A_w c_v with an infiltration/ventilation input and third order model $T_w T_i T_m$ $R_o R_i R_m$ A_w. To keep the results simple, the main outcome are listed below:

- residuals all show equivalent white noise patterns;
- likelihood based comparison, also with correction on number of parameters and/or length of data as BIC or AIC criteria, show that model $T_w T_i T_m$ $R_o R_i R_m$ A_w is the best fit;
- however, as shown in Figure 4.3, correlations between parameters are significantly large

for model $T_wT_iT_m\ R_oR_iR_m\ A_w$, which makes it less desirable for this study than the simple second order model $T_wT_i\ R_oR_i\ A_w$.

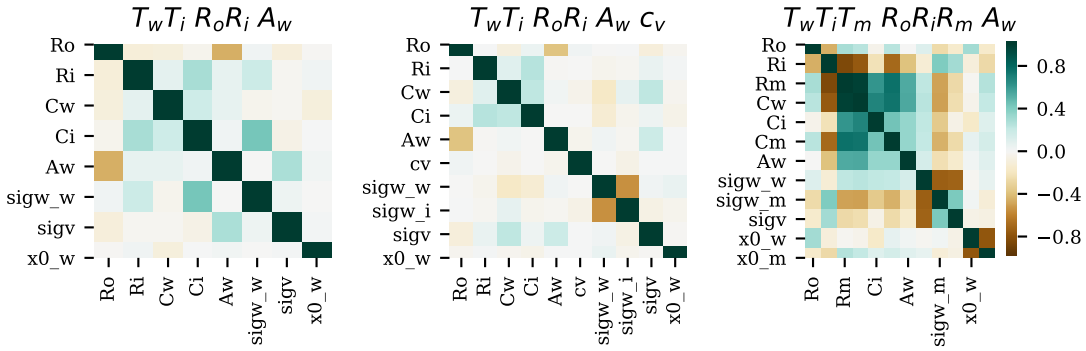


Figure 4.3 – Parameter correlation of model $T_wT_i\ R_oR_i\ A_w$ compared to models $T_wT_i\ R_oR_i\ A_w\ c_v$ and $T_wT_iT_m\ R_oR_iR_m\ A_w$

As a conclusion, in a frequentist approach, model $T_wT_i\ R_oR_i\ A_w$ shows satisfactory residuals and is a sufficient compromise between data agreement and number of parameters to avoid overfitting.

4.3.3 Weather variability in a numerical methodology

In application of the aforementioned methodology, the idea is to use various weather datasets to study how these boundary conditions influence the stability of overall heat transfers estimation. This section first shows how simulations with actual weather data were insufficient in deeply understanding the relationship between weather and R_{eq} estimation and then presents a set of 2000 stochastically generated weather files, built for sensitivity analysis. This set will prove to be an informative alternative to actual weather data.

4.3.3.1 Estimation from 10 years of actual weather

In a first attempt to understand the influence of weather conditions, the reference model was run with 10 years of actual weather data in Geneva, and for each year on subsets of data of variable lengths: 2, 3, 5, 8, 11, 15 and 25 days. Each subset serves as training data for the calibration of stochastic RC model $T_wT_i\ R_oR_i\ A_w$. An estimation of R_{eq} is inferred from the estimated parameters of the model for each subset and shown in Figure 4.4, along with their uncertainties.

As the length of the dataset increases, the variability decreases to converge approximately within a 5 % error band around the target R^*_{eq}. This would mean that at around 15 days are necessary to provide a R_{eq} estimation robust to weather conditions.

This application raises however some concerns:

- study of the variability with only a few estimations is statistically difficult. Affirmation that 15 days are sufficient to have satisfactory robustness in the estimation is far fetched from such a small sample.
- each dataset suffers from the effect of internal mass. Depending on whether the previous days were colder or warmer, there is a heat debt in the thermal capacities in the building. Using actual data cannot eliminate this effect. Growing the number of samples by creating multiple data subsets starting at different dates will not alleviate this issue.
- it remains a challenge not to say impossible to determine what weather conditions have a significant influence on the outcome.

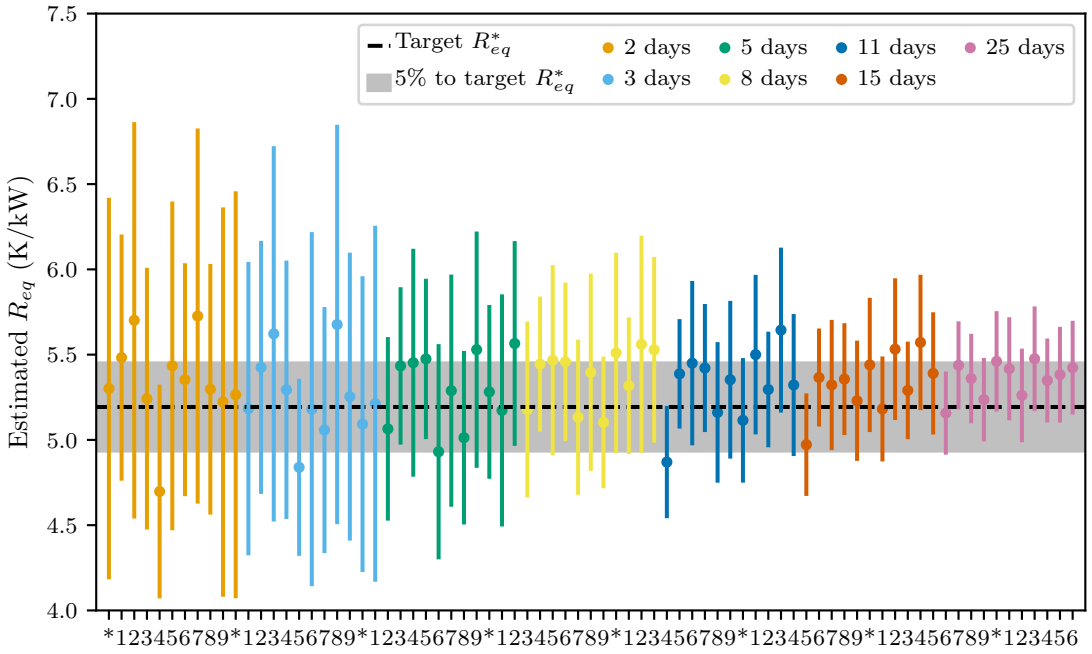


Figure 4.4 – Variability with simulations and calibrations from actual weather data

4.3.3.2 Stochastically generated weather data

This study proposes to use synthetic weather data from which a variance based sensitivity analysis is possible (Goffart et al., 2017a). A total of 6 weather variables are stochastically generated to be representative of usual weather conditions in Geneva in winter, following the methodology described in Anstett-Collin et al. (2015); Goffart et al. (2017a), as a time series constructed by a combination of statistical and deterministic features. The charateristics are extracted on the basis of representative weather data : the TMY weather file (Pernigotto et al., 2014). The TMY file, standing for Typical Meteorological Years, is built by concatenation of typical months. Each month is chosen from 30 years actual data: each monthly dataset is weighted as a sum of 13 Finkelstein-Schafer statistics (Finkelstein and Schafer, 1971) from the temperature, wind and solar radiation data. In the end, the chosen monthly dataset is the one that shows statistics closest to mean, median of the 30 years data distribution, after having discarded years with exceptionally long periods of consecutive warm, cold or low radiation days. The stochastic generation (Goffart et al., 2017a) contains then as much variability as in the TMY file : if the TMY has for one particular variable a lower variability than the rest of the 30 years actual weather data, it will reflect in the synthetic data.

From the TMY file, Goffart et al. (2017a) select 6 weather variables to stochastically generate 2000 weather files, the rest of the variables are left unchanged. The generated variables are exterior dry bulb temperature, relative humidity, direct normal solar irradiation, horizontal diffuse solar irradiation, wind speed and wind direction.

Finally, the weather data is generated as to calculate sensitivity indices through a Sobol variance method able to cope with groups of time-dependent inputs, such as here time dependency of each weather variable. Sensitivity indices by groups estimate the effect of the entire time series of the meteorological variable under study. The sensitivity indices are therefore scalars even when the variables are time series. The indices are calculated from two sets of 1000 samples, each sample of the first 1000 being defined by the characteristic features extracted from the TMY file of each weather variable, the second 1000 samples being a rearrangement of the first.

In this study, the output of interest for the sensitivity analysis is the R_{eq} estimation and in

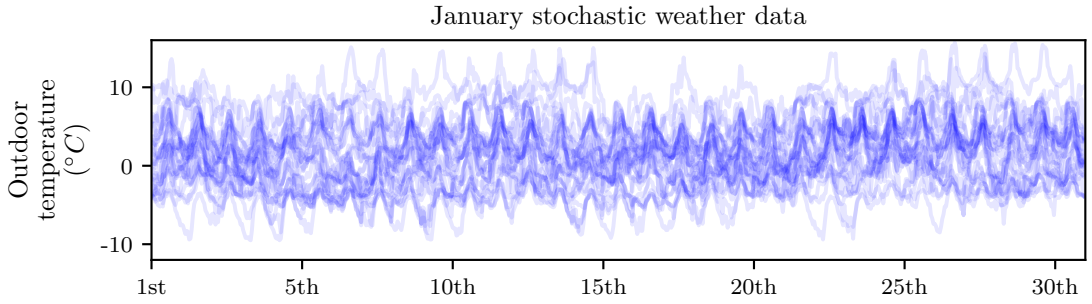


Figure 4.5 – Stochastically generated weather data: example of the outdoor temperature profiles in January. The stochastic data varies between $-10°C$ and $+13°C$, as usual in January in Geneva

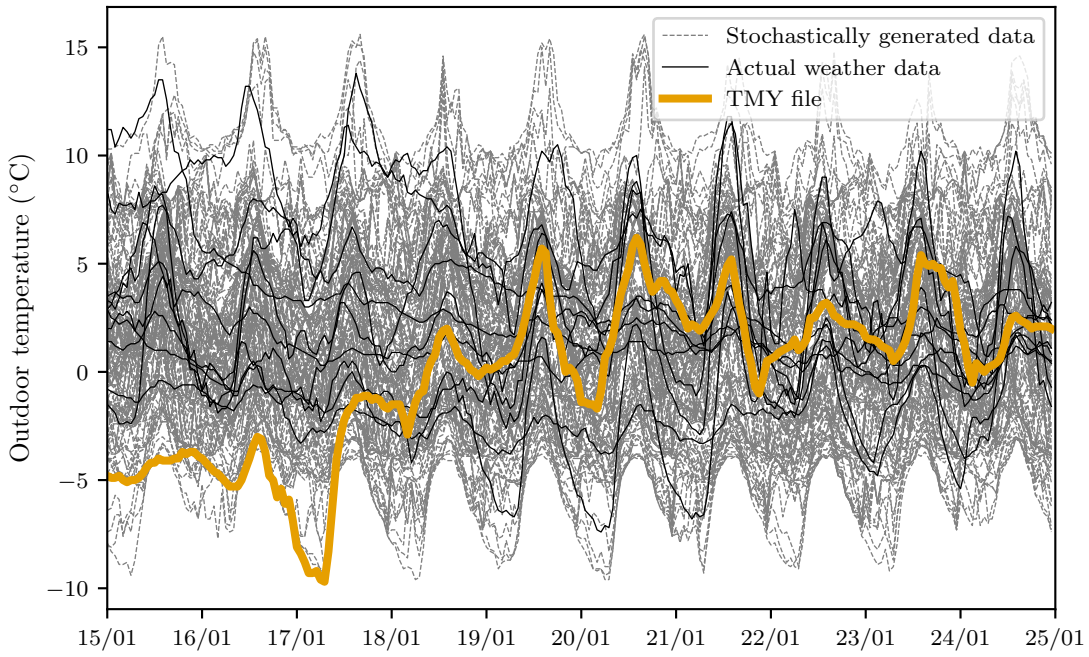


Figure 4.6 – Comparison of the stochastic outdoor temperature profiles with actually measured temperatures: stochastic data seems rather realistic

particular whether the weather conditions lead to increased or decreased estimates. The results from 4.4 suggested that the variability of the R_{eq} estimations with respect to weather variability is expected to become negligible as the data becomes more informative with longer datasets. To understand more in depth how each weather variable influences the output, partial variance and sensitivity indices are calculated.

In order to check the representativeness of the generated weather data, Figure 4.7 compares the synthetic data to the actual data from Geneva. The figure shows the empirical cumulative distributions of the 6 weather variables for the month of January of the actual weather data in black and in orange the data used to generate the synthetic data. The grey areas represent the 50 %, 75 % and 95 % quantiles of the synthetic weather data.

Synthetic outdoor dry bulb temperatures seem to be representative of the actual measured temperatures. Synthetic wind direction is in good agreement with actual measurements as well.

Synthetic relative humidity seems to be lower than some of the actual measurements. The synthetic diffuse radiation however seems slightly overestimated, as does the wind speed. This is due to the fact that the *TMY* weather dataset has indeed higher wind speed than other years: in

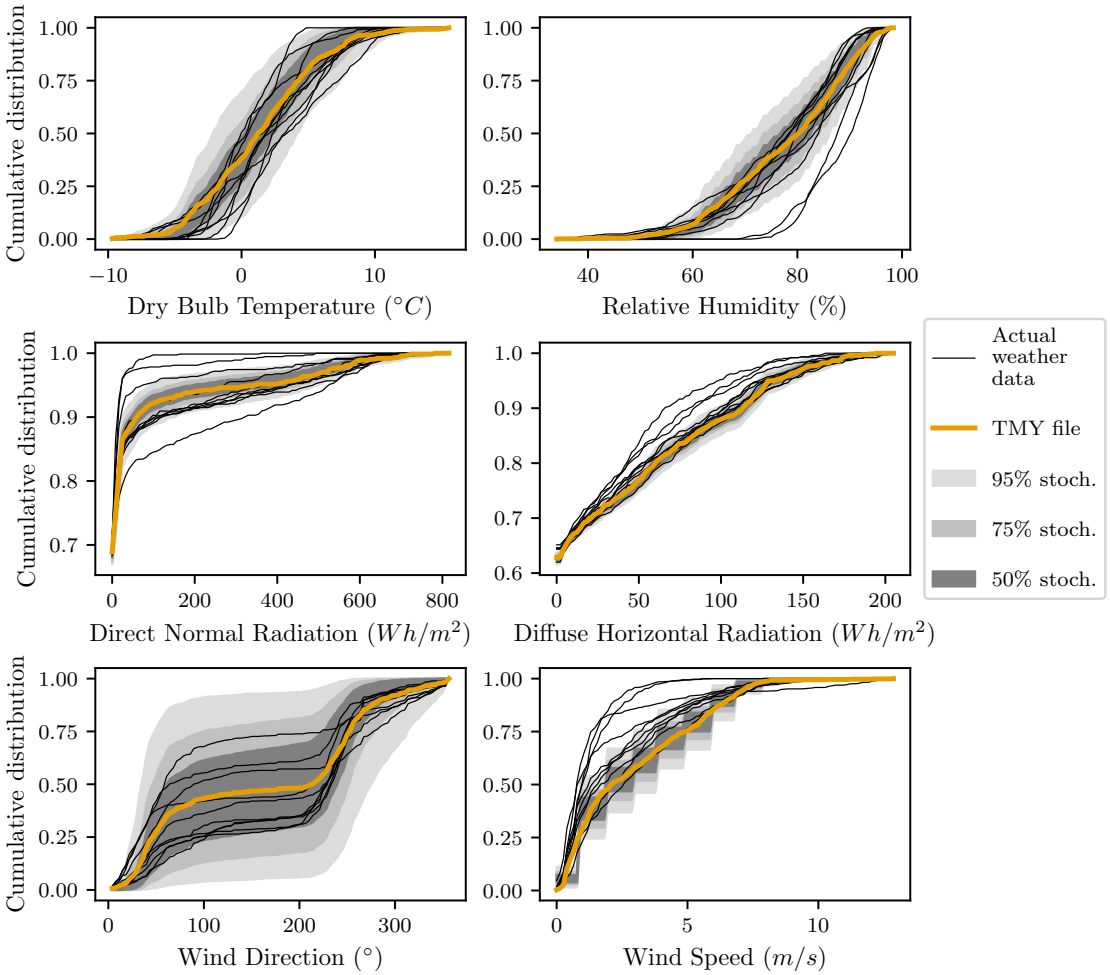


Figure 4.7 – Cumulative distribution quantiles of the 6 stochastic weather variables against cumulative distributions of actual weather data. Representativeness of the stochastic generation can be assessed by the position of the actual data distributions in the quantiles.

Figure 4.7 the orange line representing the *TMY* file is indeed significantly lower than the black lines representing 10 years of actual weather data.

The generated direct normal radiation data does not cover a range as wide as the actual data: some of the real data may have much higher or lower direct radiation. This might have an impact on the following results.

Finally, considering that the generated weather data concerns January only, all subsets to be submitted to model training should begin as early in the month as possible. Datasets will thereby be less biased by accumulated thermal inertia as they would have undergone similar weather in December. At the same time, there is small discontinuity in the weather data between December 31^{st} and January 1^{st}. All data subsets start therefore on January 2^{nd}.

4.4 Decrease in variability of R_{eq} estimation with experiment duration

4.4.1 Variability with a 2-days model training

For each of the 2000 data sets and for each subset, the stochastic RC model $T_wT_i\ R_oR_i\ A_w$ is calibrated. In each case, R_{eq} is inferred as the sum of the resistive parameters estimations. Figure 4.8 shows on the left hand side 50 randomly picked R_{eq} maximum likelihood (ML) estimations with their confidence interval.

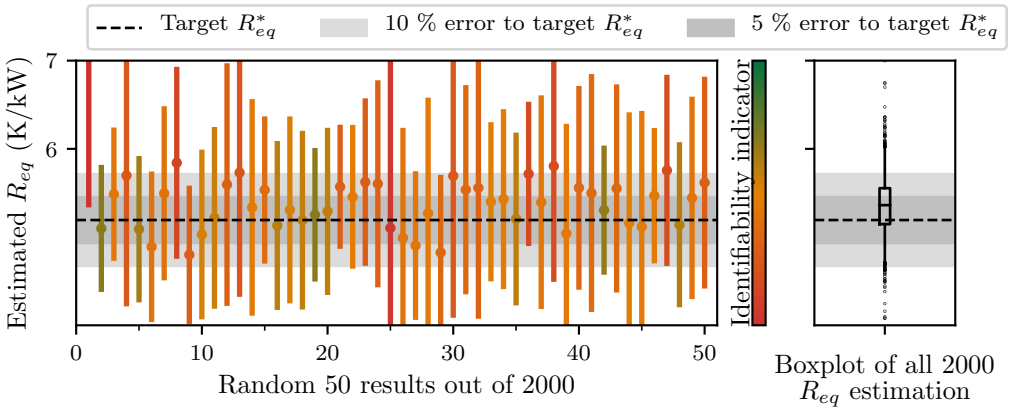


Figure 4.8 – Variability of the R^*_{eq} estimation with 2-days training:

Looking at these individual results, there are three cases to distinguish:

- the R_{eq} estimation is close to the target R^*_{eq} value: the estimation is accurate and the credible interval includes the target R_{eq}. This case is the most desirable case;
- the R_{eq} estimation is far from the target R^*_{eq} value but the credible interval includes the target R^*_{eq} : the estimation is not accurate but the credible interval relates to this inaccuracy which keeps the result trustworthy;
- the R_{eq} estimation is far from the target R^*_{eq} value and the credible interval does not include the target R^*_{eq} : not only is the result inaccurate but also give a false sense of confidence on an inaccurate result.

The latter case is the most sensitive one. It is therefore paramount to understand under what conditions over- or under- estimation occur. In order to discriminate these unwanted estimations from the others, an interpretability indicator is calculated, under the frequentist gaussian hypothesis. This indicator, defined earlier in the previous chapter (see 3.3), represents the area under the bell curve that is ± 5 % of the target R^*_{eq}. The interpretability indicator takes values between 0 and 1. For example, the case (c) with strong error on the estimation has an interpretability indicator close to zero.

On the right hand side, Figure 4.8 displays the boxplot af all R_{eq} ML-estimations showing a wide variability: the median of the 2000 estimations falls at 5.36 K/kW with a standard deviation of 0.35 K/kW (5^{th} quantile 4.82 and 95^{th} quantile 5.98). The outlier estimations show absolute errors beyond 20 % of target R^*_{eq}. This variability shows that the influence of weather conditions on the ML-estimation of R_{eq} is not negligible. A data subset longer than 2 days is certainly needed to decrease this variability.

4.4.2 Minimal measurement duration for model training

Figure 4.9 shows boxplots of all 2000 R_{eq} ML-estimations with the 7 data subsets: model training from 2, 3, 5, 8, 11, 15 and 25 days data. From the figure can be inferred that the longer the

data subset, the lower the variability. There is distinctively a decrease in total variance, towards a median value slightly above the target value R^*_{eq} . Calibrations from 11 days data and more show all estimated R_{eq} values within 10 % of their median value, hence ensures low variability in the R_{eq} estimation with respect to weather influence.

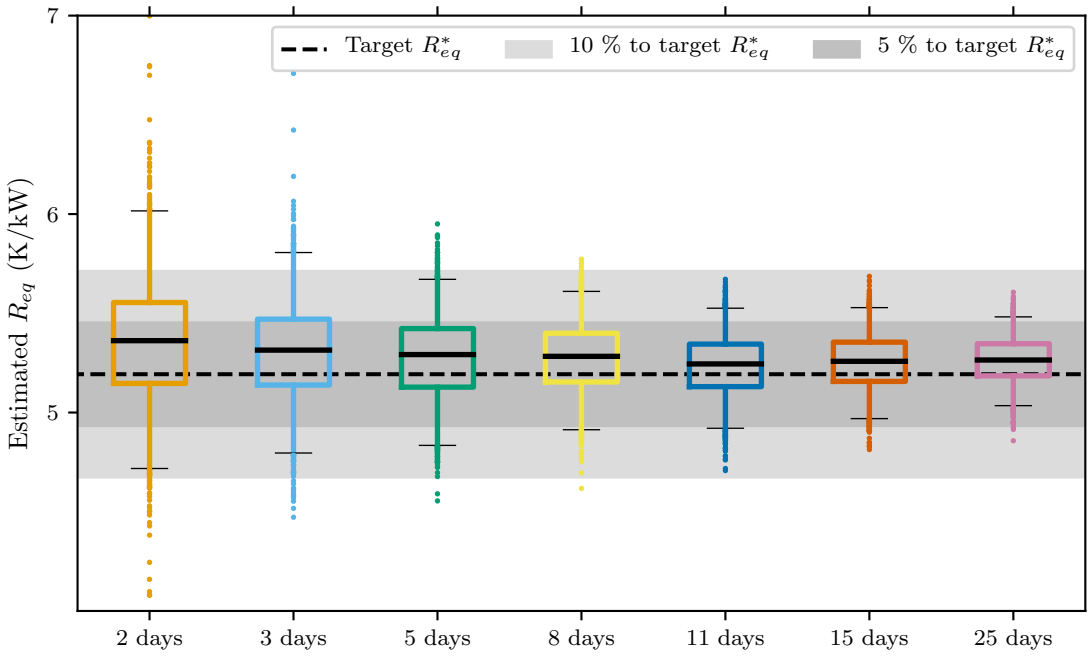


Figure 4.9 – 2000 R_{eq} ML-estimations for growing duration datasets: datasets over 11 days are all within $\pm$10 % error to the target R^*_{eq}.

To validate the impression of decrease in variability from Figure 4.9, Figure 4.10 shows for each data subset the evolution of the total variance of ML-estimations.

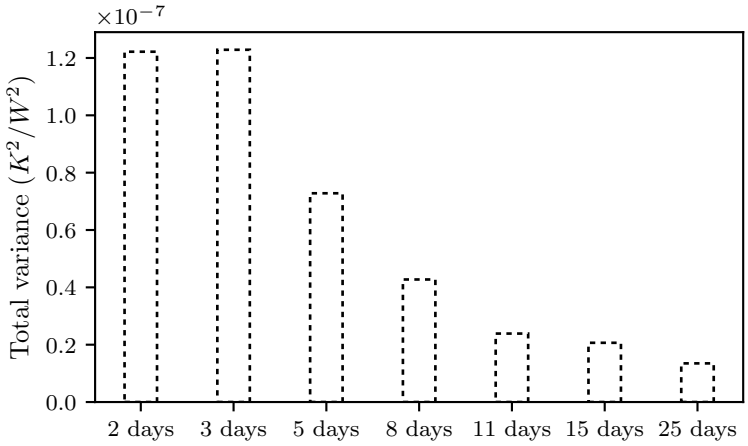


Figure 4.10 – Decrease of the total variances with longer datasets

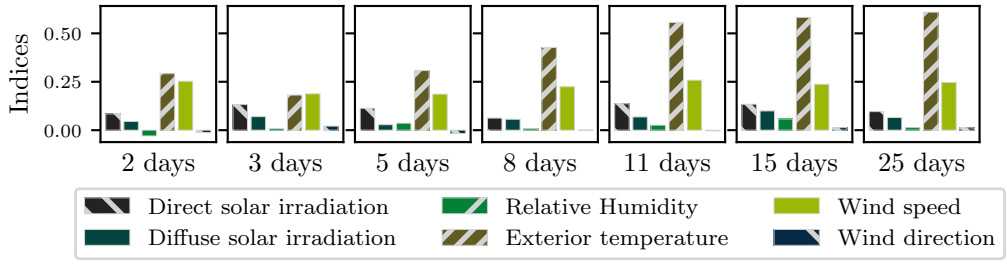
From the 2 days and 3 days training, the R_{eq} ML-estimations have a total variance around 1.2×10^{-7} K^2/W^2. With 8, 11 and 25 days training, the total variance decreases respectively by a factor 3, 4 and 6. Partial variance evolution will be further discussed in the next section.

As a partial conclusion, 11 days training suffices to reduces the error below the $\pm$10 %. Longer training still significantly reduces the overall variance. However, from a practitioner's point of

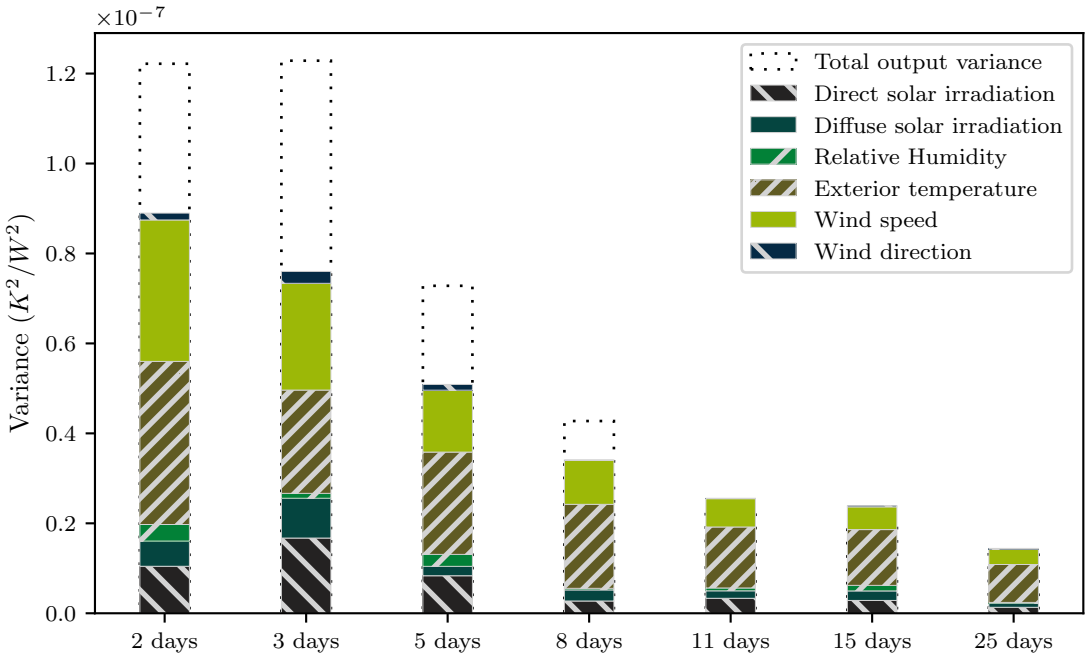
view, longer experiments might be unnecessary, as it would immobilize the experimental setup almost twice as long for an all in all relative decrease in uncertainty.

4.5 Influential weather variables on an R_{eq} estimation

As mentioned in section 4.3.3.2, the synthetic weather files allow a global sensitivity analysis on the output with respect to 6 weather variables. Figure 4.11a shows the sensitivity indices of the estimations of some parameters with respect to the weather variables: R_{eq}. The sensitivity indices are calculated for all 7 data subsets. The indices shown in Figure 4.11a are the first order indices, meaning that they only show the direct influence of each weather variable. If the sum of each first order indices is close to 1, it would imply that there were almost no second order effects, i.e. combined effects of the weather variables.



(a) First order sensitivity indices of the R_{eq} ML-estimations: outdoor temperature and wind speed are the main factors to variability



(b) Decrease of the total and partial variances with longer training datasets

Figure 4.11 – Individual effect of weather inputs on the estimations of R_{eq}

Let us also finally remind that, as mentioned in the previous chapter (see 3.4.3.3), the values of the sensitivity indices are always simply estimated. The indices given in Figure 4.11a should

mainly be interpreted as order of magnitudes. Indices below 0.1 may be considered insignificant, given the uncertainty of their estimation.

In Figure 4.11a can be seen that the variability of the R_{eq} ML-estimations are influenced by the outdoor temperature and the wind speed. With shorter datasets, the sum of the first order indices is significantly inferior to 1. This means that the variability is also explained by interactions of weather variables. Variability from with longer datasets is on the contrary almost only explained by the variability of outdoor temperature and wind speed, seeing that the indices add up to 1. Let us also note that wind direction was not expected to have an influence on the estimations as it is not used in the infiltration and ventilation model of EnergyPlus. Its sensitivity indices are indeed insignificant.

This outcome is also visible in the evolution of the partial variances of each weather variable shown in Figure 4.11b. Let us remind that the total variance is the sum of all order partial variances. The figure shows the first order partial variances, i.e. the partial variances due to the effect of each weather variable individually. With these elements in mind, it is visible that the total variance of the 11, 15 and 25 days datasets is solely explained by first order effects of the weather variables, mainly outdoor temperature and wind speed.

Let us now examine how outdoor temperature and wind speed influence the R_{eq} ML-estimations. Figure 4.12 shows how the R_{eq} ML-estimations from 11 days training vary with the average outdoor temperature on the abscissa and the average wind speed on the ordinate. Warmer training periods tend to produce over-estimations and colder days under-estimations. At the same time, non windy days produce in overall over-estimations, windy days under-estimations.

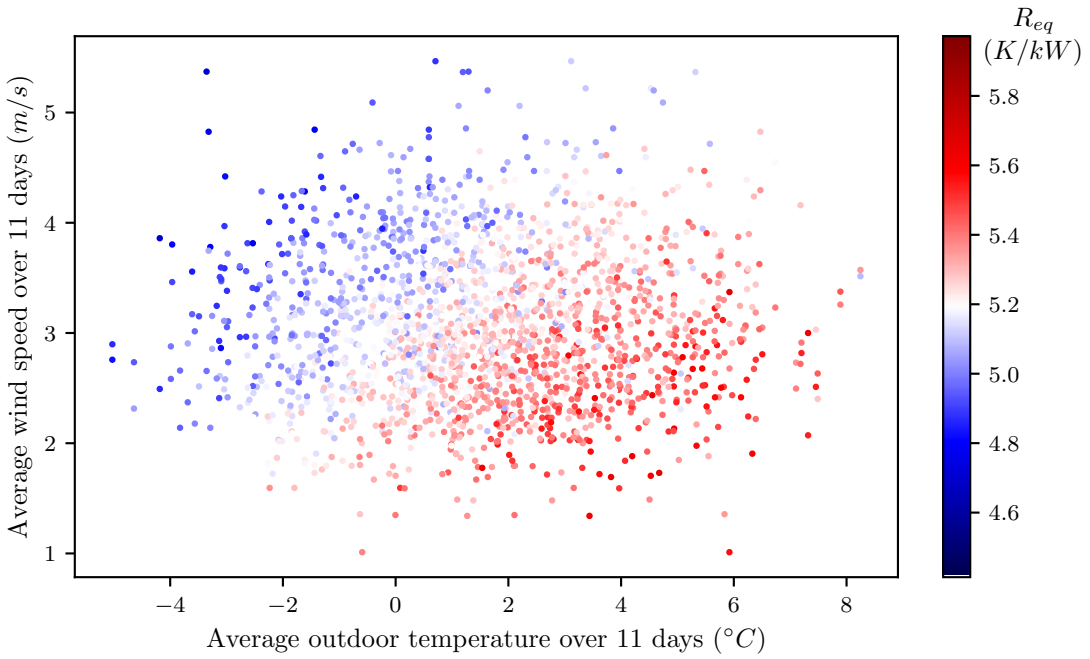


Figure 4.12 – Variability of the R_{eq} ML-estimations from 11 days training with respect to outdoor temperature and wind speed. Colours refer to ±15 % errors to target R^*_{eq}.

At the same time, an interaction can also be seen in Figure 4.12: training from warm and unwindy days results in over-estimation, cool and windy days in under-estimations.

This outcome is in agreement with the hypothesis that the large air change rates in the reference model are a cause of inaccuracy in the estimation of the overall thermal resistance. As the ventilation related heat losses have been modelled in the EnergyPlus simulation environment, there is a direct relationship between temperature difference between indoors and outdoors and wind speed. Ventilation related heat losses are larger with cold outdoor temperatures and with

high wind speed and on the contrary smaller with warmer or unwindy days.

Finally, Figure 4.13 shows more clearly how the influence of outdoor temperature and wind speed evolves from short to longer model training.

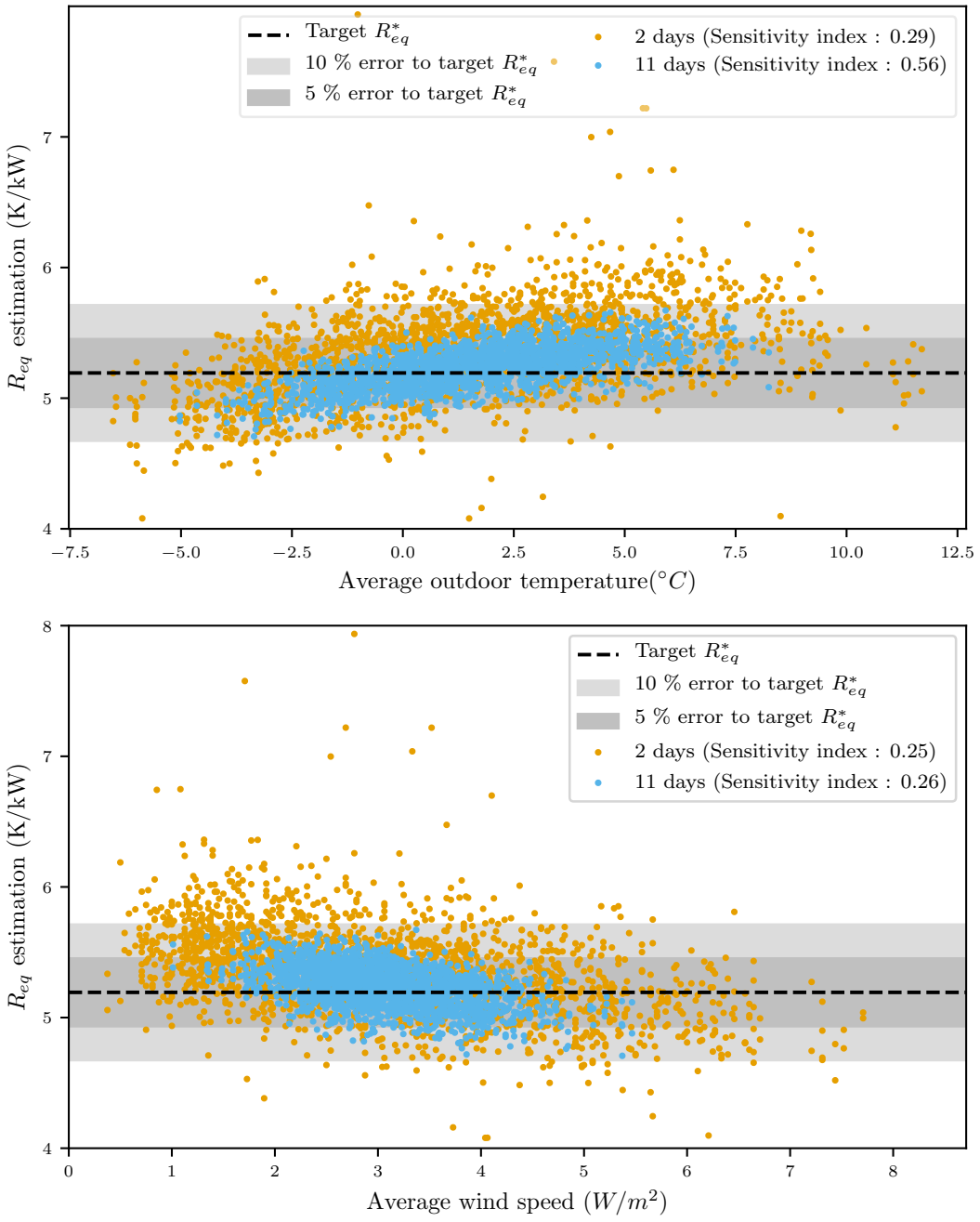


Figure 4.13 – Variability of the R_{eq} ML-estimations with respect to averaged outdoor temperatures and wind speeds, for 2 days and 11 days training.

As seen earlier, this figure too shows the decrease in total variance of the R_{eq} ML-estimations with longer training sets: the vertical spread of all estimations are narrower with the 11 days training. Interestingly, while the total variability does decrease, the angle representative of the correlation remains quite similar whatever the dataset length. Longer datasets produce averages that are less spread horizontally, but the relationship between temperature and R_{eq} estimation is almost unaltered.

A natural assumption would have been to consider that colder days lead to more accurate

estimations than warmer days, as colder days increase the heat losses and thus the heating power needed to keep up with the indoor temperature set point. This assumption does not seem to hold here. If it were, the variance would be significantly narrower under cold days than under warm days. Here, there is no significant difference in vertical spread between cold and warm days, nor is there any between windy and unwindy days.

4.6 Discussion

The results have shown that the calibration needs to be based on at least 11 days to ensure convergence within ± 10 % of the target value R^*_{eq}. They also showed that the longer the dataset, the smaller the total variance of separate estimations. Calibration from shorter datasets will lead to uncertain results, which variability will mainly be due to the variability of the weather conditions. Compared to controlled tests with optimized heating or temperature patterns for which a few days is sufficient, 11 days or more is longer. But considered that optimized heating or temperature patterns create richer and less correlated data, it is quite consistent to find 11 days as a minimum in uncontrolled conditions.

In addition, as non intrusive measurement design is considered here, 11 days or more is not a prohibitive duration: as long as the test remains user-friendly, leaving data loggers for a few weeks is most probably neither burdensome for the building occupants nor for the expert carrying the diagnosis. All the more, compared to data exploitation by other low order models such as linear regressions or auto-regressive models (at least as suggested in Senave et al. (2019)), this stochastic RC model shows a faster approach to exploiting the data, as long as temperature is not kept constant by the occupants. All day long constant temperatures would necessarily lead to choose models which have the heating power as output to exploit the data.

On another note, the results are certainly specific to this case study, although the order of magnitude of model training duration should be approximately similar. Upon exploitation of an actual measurement campaign, an important indicator would actually be to consider convergence of the estimation.

Let us therefore take the opportunity to make a distinction between convergence of a single estimation and repeatability of such an experiment. At the scale of a punctual measurement campaign, the stop factor would be convergence of the estimation results: more measurements do not significantly change the results.

Although tools for assessing the convergence are not the purpose of this chapter, one may extrapolate the well-established ISO 9869-1 (ISO 9869-1, 2014) criteria for wall-scale characterisation. First, the R_{eq} estimation should not deviate more than 5% from the 24 h earlier estimation. Secondly, with N the total duration of measurements, the R_{eq} value inferred from the last $2/3N$ days training are within 5 % of the first $2/3N$ days training.

The first ISO 9869-1 criterion is roughly applied with the available data in Figure 4.14. The estimation from 3 days training is compared with that at 2 days, the estimation from 5 days training to that at 3 days and so forth. All 2000 deviations calculations are represented with 50 %, 75 % and 95 % quantiles. Interestingly, a large majority of cases show a convergence in the sense of the second ISO 9869-1 criteria within 5 days training. 8 days training is sufficient for convergence in more than 95 % of all cases. Yet at the same time, repeatability as defined previously is not quite achieved: the variance caused by weather conditions is still significant.

The second criterion is not literally applicable to the data from this chapter, but one may extrapolate that there could be up to 10 % deviation between the first $\frac{2}{3}N$ days inference and the last $\frac{2}{3}N$ days inference, for example when the first days are particularly cold and the last particularly

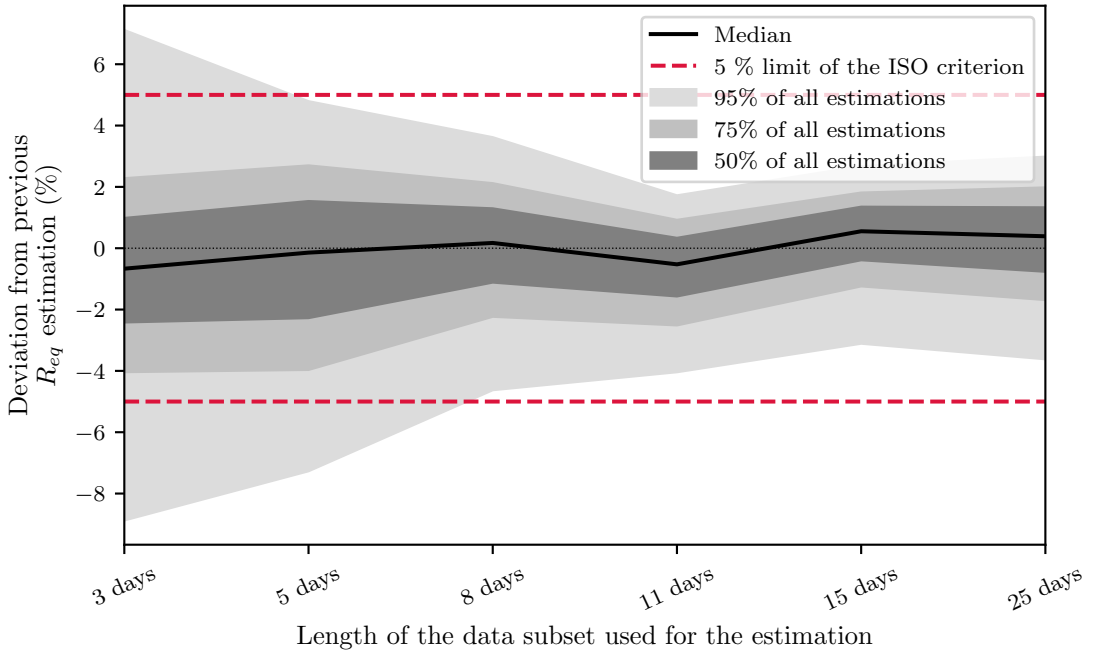


Figure 4.14 – ISO 9869-1 convergence criteria: there should be less than 5 % deviation with the previous estimation.

warm or inversely. However, with similar consecutive weather conditions, convergence would be considered "achieved" rather quickly.

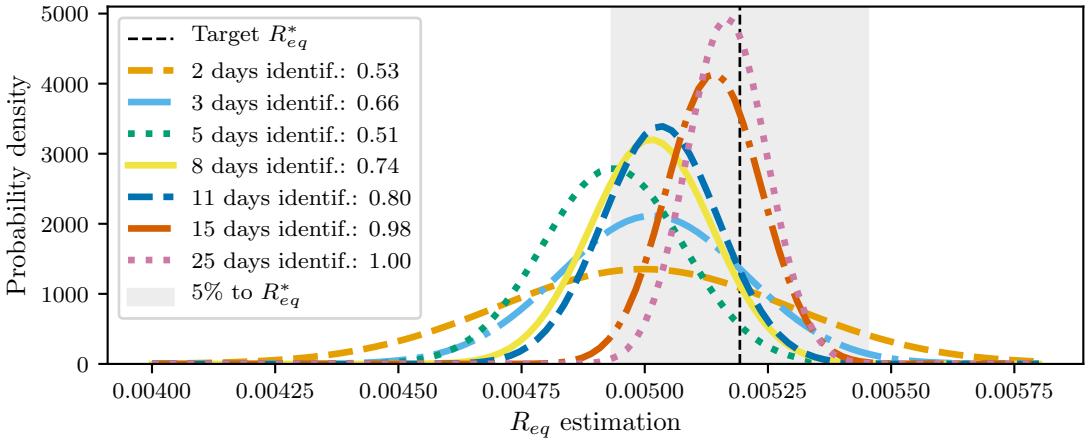
In summary of the convergence topic, according to these criteria, convergence could in some cases be considered as achieved with fewer measurements than the results from this chapter would suggest to reach repeatability.

The question is then maybe not to look at deviation in the punctual R_{eq} result but rather quantify the information learnt from the data and its evolution, with in mind the representativeness of the weather conditions. In this light, it is a call for a Bayesian perspective on the results: the important outcome to consider is the complete posterior distribution and not the single most probable value of interest. If upon significant variation of weather there is no more learning from the data, i.e. there is no further change in the posterior distribution, then the measurements may stop.

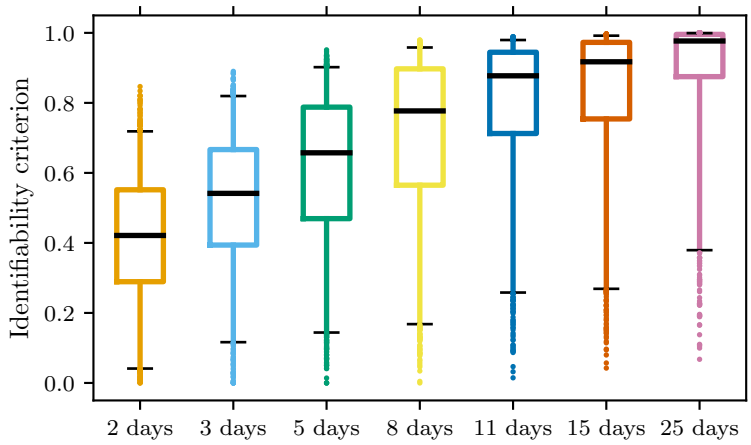
Repeatability can also be looked at with a posterior distribution view. The interpretability indicator defined in the previous chapter (see 3.3) is defined as the cumulated distribution within the boundaries $R^*_{eq} \pm 5\%$. Although it is primarily defined for Bayesian approaches that yield posterior probability distributions, it may be extended to a normally distributed posterior assumption.

Figure 4.15a illustrates how the interpretability indicator can be applied to the results of this chapter, with the frequentist assumption of normally distributed ML-estimations. The interpretability indicator is represented by the area under each curve that is in the grey target band. In the first days of model training, there is in this figure a distinct learning process. In Figure 4.15a, 15 and 25 days training however yield quite similar results, learning could have been considered sufficient at 15 days. Figure 4.15b shows again how 11 days training yields satisfactory results in almost all cases but a few outliers. 25 days training scores in overall higher than 0.85 which means accurate estimates with low uncertainty.

The interpretability indicator can obviously not be a tool for exploitation of actual measurement campaigns as the target value is unknown. But it supports the idea of judging convergence through posterior distribution, by the use of a divergence metric like the Kullback-



(a) Illustration of the identifiability indicator: the most important is that most of the posterior distribution is satisfactorily close to the target value. Although not perfectly accurate at peak, the posterior distribution is in overall within boundaries of the grey target area, provided sufficient training.



(b) Evolution of the identifiability indicator: the larger the dataset, the closer the indicator to 1. Longer training provides more accurate posterior distributions.

Figure 4.15 – Application of the identifiability indicator

Leibler divergence (see Chapter 2). Furthermore, it would make sense to exploit data in a Bayesian approach and for example use in line calibration algorithms like Sequential Monte Carlo (see Rouchier et al. (2019)).

4.7 Conclusion

Establishing reliable methods for estimating the thermal performance of buildings remains a challenging issue under the constraint of a non intrusive measurement framework: the collected data is uncontrollable and thus less informative. Such data may therefore lead to errors in the thermal diagnosis and the outcome may be uncommonly influenced by the test conditions.

In this context, this chapter has adapted the model assessment framework as designed in Chapter 3 to investigate the influence of weather conditions on the repeatability of a non intrusive

experiment. The methodology relies on a sensitivity analysis of the overall thermal resistance R_{eq} estimation with respect to 6 weather variables independently.

A stochastic RC model is used to exploit the data generated by the model assessment framework. After comparing 2, 3, 5, 8, 11, 15 and 25 days of model training, it has been found that 11 days and longer provide repeatable results regardless of the outdoor conditions.

There is however a residual variability of a few percent, that is in the present case study exclusively due to the variability of outdoor temperature and wind speed. This case study has indeed large air change rates which would emphasize the effect on heat transfers of these weather variables, all the more that the weather files used.

As the minimum limit of 11 days is to a certain extent specific to the case study used in the model assessment methodology, the results indeed call for further effort on establishing tools for the assessment of convergence from a given dataset.

5

Decomposition of heat losses in a building

Physical interpretation of parameters of RC models remains an open issue in the literature. The model assessment methodology proposed in Chapter 3 is applied to the purpose of assessing the interpretability of models parameters to actual physical properties. 300 configurations of the building serve to produce simulated datasets. Two appropriate models then undergo the model assessment framework to identify heat transfer through ventilation and transmission heat transfer through neighbouring spaces. Conclusions find that if both perform well to identify an overall thermal resistance, none achieves the initial objective, at least from poorly informative data.

Contents

5.1 Introduction

Diagnosis of the thermal performance of a building envelope from non intrusive data would benefit most from an accurate estimation of different sources of heat loss. Determining how much energy is lost on ventilation and infiltration or through specific parts of the envelope is more informative in the perspective of building retrofit than a simple overall thermal resistance. It is however unknown whether on-site measurements in non intrusive conditions are sufficiently informative to make this type of detailed estimations. The dynamical nature of data collected in these conditions could however be beneficial for training of RC models.

This chapter proposes to apply the model assessment framework defined in Chapter 3 to this purpose. The idea is to determine whether stochastic RC trained models deliver individually interpretable parameter estimates.

So far, the model assessment framework has been applied to assess repeatability of a RC model when weather conditions are variable. Convergence has been found to be achieved within 11 days for the estimation of an overall thermal resistance. Interpretability of each parameter separately has however not been established.

In particular, this chapter will examining if identifiability and interpretability of heat transfers by ventilation and heat transfer to an unheated crawl space can be achieved. The corresponding parameters will be identifiable if they can be estimated even from poorly informative data, as is simulated in the model assessment framework designed in 3. Interpretability is then assessed by comparing the estimations from the theoretical target values.

Section 5.2 describes in detail how the model assessment framework is applied to the aforementioned objective. Section 5.3 then presents the results of the application to the assessment framework on the interpretability of heat losses by ventilation. Section 5.4 in trun presents the results of interpretability of heat losses to a crawl space.

5.2 Model assessment framework for heat transfer decomposition

Interpretability of parameter estimates taken individually means that they actually represent the physical property (or aggregated properties) they are supposed to represent. In other words, a given parameter estimate should take the value of the exact actual (aggregated) property(ies) it is supposed to represent, without lumping the dynamics or the influence of an unrelated other physical property. The interpretability assessment framework proposes to verify that a change in the actual property value translates as a change in the value of the ad hoc parameter.

The general idea is therefore to make a number of variations in the thermal properties of the reference model and to study the induced variability in the parameters' estimations. This section first establishes reasonable objectives for the estimation of decomposed heat transfer, secondly exposes how the numerical model assessment methodology is applied to answer the stated problem, then details what state space models are fit to the simulated data and finally establishes the convergence of the sensitivity analysis performed in the model assessment framework, proving that the results exposed in the following sections can be trusted.

5.2.1 What decomposition can be reasonably expected?

In the expectation of diagnosing energy performance of a building and in particular thermal performance of the envelope, the more detailed the diagnosis, the more useful it will be for establishing pertinent retrofit strategies. However, in a non intrusive design of experiment, the data available reduces the possibilities, in particular as only temperatures of the heated space are

measured. Not all physical characteristics or systems may be diagnosed from only temperature measurements.

This restrictions are similar to issues in calibrating whole detailed building models from experimental data. Strachan (1993) adresses them by suggesting a global sensitivity analysis to determine the adequate spatial and temporal measurements. The idea behind this suggestion is that an input parameter in a whole building model may be estimated from data if it has an influence on the measured output. Josse (2017) too, aiming at fault detection in the building envelope, found through a Morris sensitivity analysis that thermal bridges and local insulation defaults could not be estimated from temperature measurements as their influence is found insignificant on temperature differences.

Back to the model assessment framework, it can be understood that the only estimable thermal properties are the ones that have a detectable influence on the measured model output (indoor air temperature). This leads to determine the physical characteristic that have a significant influence on indoor temperature.

Josse (2017) found that the influential parameters detectable in indoor temperature are the insulation thicknesses, in particular when considered equal in all walls. When a locally faulty insulation is considered, for example in the wall of a particular room, that default would not be detectable on a south oriented wall, but would be on the other orientations. An application of the method on a case study aiming at quantifying a linear thermal bridge was not successful, showing that a local default like a thermal bridge does not have a detectable influence on indoor temperatures. The method and results in Josse (2017) are developed on a case study of a newly built highly energy efficient dwelling, which has specific heat transfer dynamics: solar gains contribute largely to the energy balance. For this reason, a similar analysis on a poorly insulated house would probably show different influential parameters. Also, for the sake of simplicity, Josse (2017) did not vary the ventilation characteristics, although heat transfers through ventilation can be considered significant.

All in all, decomposition of heat transfers of different parts of the envelope seems to be an achievable objective. Heat losses to neighbouring unheated spaces, where the boundary temperature can be measured is a possible focus and presents an actual interest in thermal diagnosis. In the same order of magnitude, heat losses through ventilation seems to be a reasonable objective too.

Notations for the targetted values vary from one author to the other. But with the acknowledged notations of the Section 3 of ISO 13789 (ISO 13789, 2017), this chapter aims at the following heat transfer decomposition:

$$HTC = H_{ve} + H_{tr} = H_{ve} + H_d + H_u \tag{5.1}$$

with:

- HTC the Heat Transfer Coefficient (W/K): heat flow rate from indoors to outdoors divided by the temperature difference,
- H_{tr} the Transmission Heat Transfer coefficient (W/K): heat flow rate due to transmission through the fabric, divided by the temperature difference,
- H_{ve} the Ventilation Heat Transfer Coefficient (W/K): heat flow rate due to air entering the heated space, divided by the temperature difference,
- H_d the direct transmission heat transfer coefficient between the heated or cooled space and the exterior, for the whole building (W/K),
- and H_u (W/K): transmission heat transfer coefficient through unconditioned spaces.

To bridge these objectives with the state space models used for the estimation of said thermal properties, the rest of the chapter will rather focus on the estimation of thermal resistances, as the inverse of the heat transfer coefficients: $R_{eq} = \frac{1}{HTC}$, $R_{ve} = \frac{1}{H_{ve}}$, etc...

5.2.2 Application of the model assessment framework

The objectives of this application of the model assessment methodology are threefold:

- (A) examine the interpretability of the overall thermal resistance,
- (B) identify independently the air change rate,
- (C) decompose heat losses towards unheated neighbouring spaces from losses to the outdoors.

This subsection details in 5.2.2.1 what modifications are done to the reference model in the model assessment framework and in 5.2.2.2 what post-processing is done to the simulated data.

5.2.2.1 Building property sampling

The case study model described in section 3.2.2 undergoes modifications as to cover a wide range of building thermal properties : from poorly to highly insulated, from low to high air change rates, from low to high thermal inertia. However, neither the reference model structure nor its design change.

The modifications for the simulations are done on wall, attic and ground floor insulations, as well as on ground floor slab and brick wall thermal capacity (serving objectives (A) and (C)) and on air change rates (serving objective (B)). As suggested previously in section 3.4.3.2, the modifications of the thermal properties of the reference model, the *inputs*, follow a specific design of experiments allowing both a thorough exploration of the inputs space and a sensitivity analysis. A Latin Hypercube Sampling is therefore chosen, and sampled on uniform distributions on all 6 physical properties. Table 5.1 gives the boundaries between which the parameters vary, as well as complementary thermal properties of interest.

300 samples are drawn, proof of sufficiency of this sample size is shown in 5.2.4.

5.2.2.2 Simulation and selection of data

The 300 building configurations are simulated from January 1^{st} to February 28^{th} on the basis of an actual weather dataset.

The weather has been measured in Le Bourget du Lac (73, France) in 2019. The measurements have a 1 minute timestep and have been adapted to the EnergyPlus weather input file format, such as to contain: dry bulb temperature (°C), field dew point temperature (°C), relative humidity (%), atmospheric pressure (Pa), horizontal infrared radiation intensity (Wh/m^2), global horizontal radiation (Wh/m^2), direct normal radiation (Wh/m^2), diffuse horizontal radiation (Wh/m^2), wind direction (°), wind speed (m/s), total sky cover, opaque sky cover, field visibility (m) and ceiling height (m).

From the two months simulated data, a training period must be chosen. A small selection would, as found in Chapter 4, induce a non negligible effect of the weather conditions on the result. At the same time, a long dataset would just unnecessarily burden the computational cost without adding significant information to the calibration process. A length of 11 consecutive days will be selected in agreement with the outcomes of Chapter 4.

Also, to avoid the influence of the warm-up process of EnergyPlus, which repeats multiple times the first day of simulation to initiate usual temperature conditions, the dataset selection for calibration should not start in the first 2 weeks of the simulation.

	Object Variable in EnergyPlus	**Bounds**	**Distribution**	**Units**	**Other info**
Walls insulation thickness	(Material) Thickness	$[0.05; 0.25]$	Uniform	m	$\lambda = 0.032\ W/(m \cdot K)$
Attic insulation thickness	(Material) Thickness	$[0.05; 0.25]$	Uniform	m	$\lambda = 0.04\ W/(m \cdot K)$
Ground floor insulation thickness	(Material) Thickness	$[0.05; 0.25]$	Uniform	m	$\lambda = 0.022\ W/(m \cdot K)$
Air Change Rate	(ZoneVentilation: DesignFlowRate) Air changes per Hour	$[0.2; 2.0]$	Uniform	h^{-1}	(-)
Thermal capacity brick wall	(Material) Specific Heat	$[0.6; 1.0]$	Uniform	J/kgK	$e = 13$ cm
Thermal capacity floor concrete cast	(Material) Specific Heat	$[1.0; 2.0]$	Uniform	J/kgK	$e = 20$ cm

Table 5.1 – Physical thermal properties variations of the reference model for the decomposition study

Finally, the data used for calibration is resampled with a timestep of 8 minutes. Larger timesteps may imply aliasing, as seen previously in 2.4.2.2. But again shorter timesteps will enhance computational burden without proving to be much more informative than a 8 minutes timestep.

With the aforementioned constraints in mind, the chosen selected dataset runs arbitrarily between January 31^{st} and February 10^{th}. As shown in Figure 5.1, the selected 11 days present rather mild outdoor temperatures, varied solar irradiation with sunny and cloudy days and varied wind speed conditions.

Last but not least, it can be considered that with this data selection, the minimal requirements found adequate for RC models identification in Chapter 2 in section 2.4 regarding sampling, minimal inputs and minimal outputs are fulfilled. This implies that the state space models should be practically identifiable, even if a rigorous verification always needs to be performed.

5.2.3 State space model selection and validation

300 datasets have been produced from 300 different configurations of the reference building. To exploit the datasets and estimate the thermal resistances set as objectives, adequate stochastic RC models need first to be selected. Then, after appropriate model calibration, the most appropriate models are selected through a likelihood ratio procedure. If the selected models also achieve model validation tests, physical interpretability may be assessed.

5.2.3.1 A set of adequate state space models to calibrate

As stated in Chapter 2, the first step is to select an appropriate model : the model best fitting the data and which residuals are significantly not autocorrelated and have white noise properties, in which case the significant heat dynamics are covered by the model. One node models such as

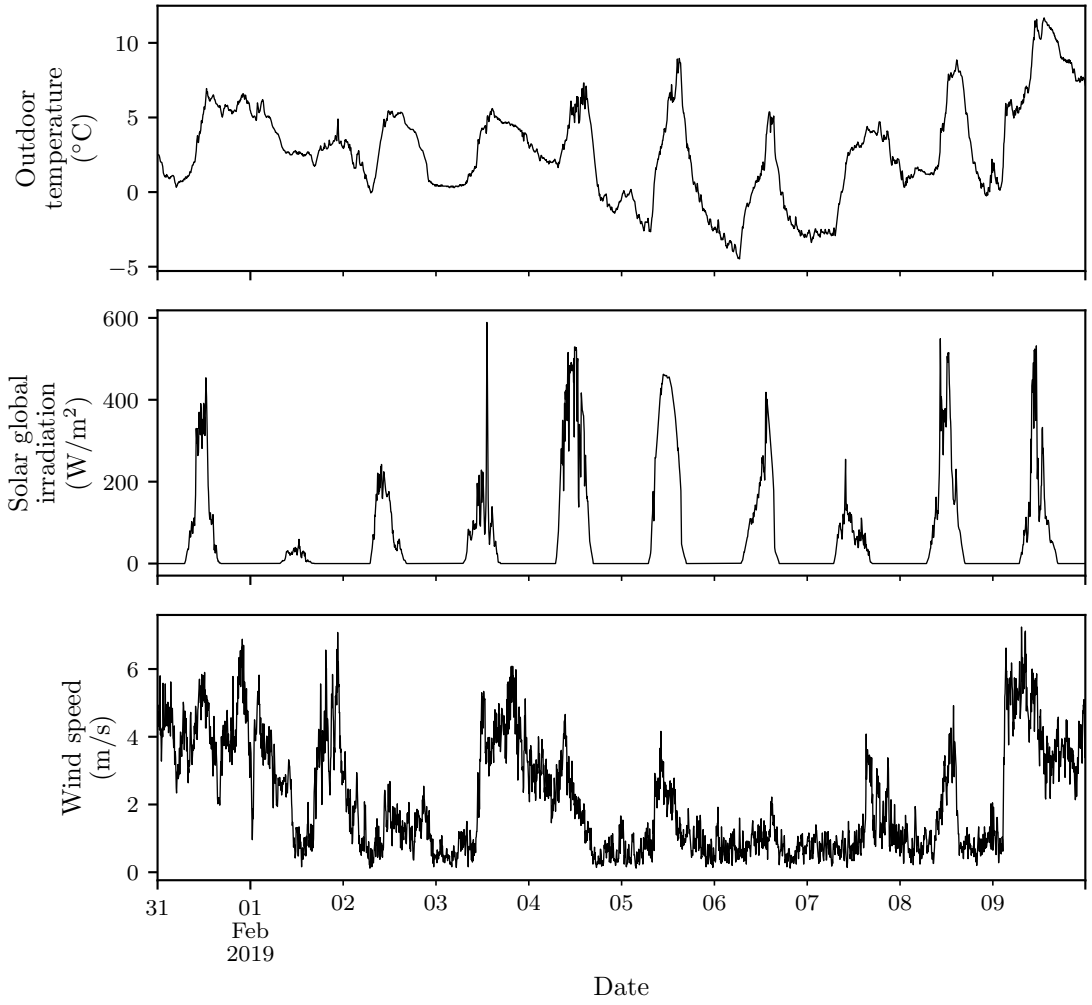


Figure 5.1 – Selected weather dataset used in the 300 simulations: outdoor temperature, solar irradiation and wind speed

2.3 are not proposed in the process as they show very poor fit to the data. In the end, a set of structurally identifiable is selected and tested:

- two nodes: $T_wT_i\ R_oR_i\ A_i$, $T_wT_i\ R_oR_i\ A_w$ and $T_wT_i\ R_oR_i\ A_wA_i$ (see Model 2.7 with or without A_i),
- two nodes with infiltration/ventilation term: $T_wT_i\ R_oR_i\ A\ c_v$ (see Model 2.8),
- two nodes with measured temperature in a neighbouring unheated space: $T_wT_i\ R_oR_i\ A\ R_b$ (see Model 3.4.3.1),
- two nodes & one node towards neighbouring space: $T_wT_i\ R_oR_i\ A\ T_b\ R_{ib}R_{bb}$ (see Model 2.12).

Section 2.2.5 already established structural identifiability. At the same time, the selected dataset theoretically provides data sufficiently informative for practical identifiability, a point that will be verified upon analysing the results.

5.2.3.2 Model calibration

From each of the 300 simulations, a subset of data from January 31^{st} to February 11^{th} is used as synthetic data for parameter estimation through model calibration.

As a first step, model selection is performed, but cannot however be done on a single dataset, as the best fitting model is likely to be related to the reference model configuration. Best fitting

model for one dataset may then not be valid for other datasets. This implies that for all models tested in the selection process and for each of the 300 datasets, a calibration needs to be performed. The number of calibrations adds up to at least 1500.

A Bayesian calibration, although preferred as it draws from the actual posterior distribution, would here bring a considerable computational burden. On the other side, model selection based on likelihoods comparison has been proven reliable (Bacher and Madsen, 2011). Therefore, the model selection process is performed on the basis of a frequentist calibration of a set of adequate state space models, through a BFGS optimisation (see in Chapter 1 section 1.5.1).

5.2.3.3 Model selection through a likelihood ratio test

Each of the 300 configurations undergoes the same model selection process, based on a likelihood ratio test Bacher and Madsen (2011). Each data subset is used to calibrate models in a certain order, respecting the principle of nested models: $\mathscr{M}_{null} \subset \mathscr{M}_{alt}$ if upon setting one or more parameters of $\mathscr{M}_{alt}$ to 0, it becomes identical to $\mathscr{M}_{null}$.

A likelihood ratio test then compares a proposed model $\mathscr{M}_{alt}$, also called alternative model, to a basic model $\mathscr{M}_{null}$, also called the null hypothesis model. These models need to verify $\mathscr{M}_{null} \subset \mathscr{M}_{alt}$. For a dataset, both models are fitted. The logarithmic ratio of their likelihoods is calculated. To test if there is a significant improvement in the likelihood, the significance of the test, the p-value, is estimated through the χ^2 value of the ratio. If the p-value is smaller than for example the usual 0.05, than the alternative model is significantly better than the null hypothesis model.

The model selection test would then normally start with one basis "null" model, the one node model for example. This model however was found to be an extremely poor fit to the data and has been discarded. It is proposed here to rather start with two nodes models and unusually start with choosing between $T_wT_i\ R_oR_i\ A_i$ and $T_wT_i\ R_oR_i\ A_w$. As the reference model originating the simulated data is subject to solar irradiation, a two nodes model without solar parameter makes no physical sense. Starting with choosing between $T_wT_i\ R_oR_i\ A_i$ and $T_wT_i\ R_oR_i\ A_w$ alleviates the computational cost of calibrating a model on another 300 datasets.

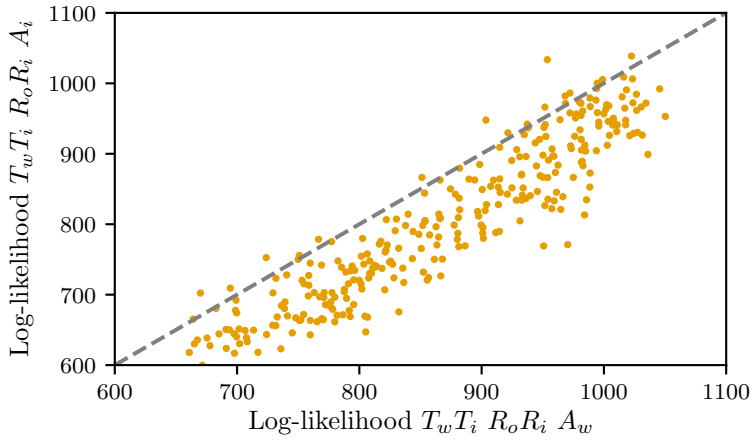


Figure 5.2 – Log likelihoods of best fits for models $T_wT_i\ R_oR_i\ A_i$ and $T_wT_i\ R_oR_i\ A_w$ for each 300 samples: all points under the 45° line favour the x-axis model ($T_wT_i\ R_oR_i\ A_w$).

Models $T_wT_i\ R_oR_i\ A_i$ and $T_wT_i\ R_oR_i\ A_w$ are however not nested and the comparison is made with simple log-likelihoods calculations, as is shown in Figure 5.2: the higher the log-likelihood, the better the fit. The grey dotted diagonal line marks the place where models have an equally good fit to the data. Points **under** the diagonal line have a higher likelihood for the x-axis model whereas points **over** the diagonal line have a higher likelihood for the y-axis model.

From Figure 5.2 can therefore be inferred that the model selection process should start with model $T_wT_i\ R_oR_i\ A_w$ as a vast majority of points are under the diagonal line, favouring therefore the x-axis model $T_wT_i\ R_oR_i\ A_w$.

The next round of test will consider the null hypothesis as being the model $T_wT_i\ R_oR_i\ A_w$, which has just been selected. The alternative models, verifying $\mathscr{M}_{null} \subset \mathscr{M}_{alt}$, are models $T_wT_i\ R_oR_i\ A_w\ c_v$, $T_wT_i\ R_oR_i\ A_w\ R_b$ and $T_wT_i\ R_oR_i\ A_wA_i$. The likelihood ratio test is applied to infer whether there is a significant difference between the null hypothesis and the alternative models, with results of the rest shown in Figures 5.3.

Figure 5.3a shows all 300 p-values of the likelihood ratio tests for all 3 alternative models against null hypothesis model $T_wT_i\ R_oR_i\ A_w$. All values except a few are below the 5% acceptance limit. In overall, it can then be considered that all 3 alternative models are significantly better than null hypothesis model.

To distinguish between different alternative models, Bacher and Madsen (2011) select the one with the better score on likelihood. Figure 5.3b therefore shows for all 300 results the log-likelihoods for all 3 alternative models on the y-axis, while the x-axis shows the log-likelihood of the null hypothesis model.

Similarly to Figure 5.2, if the y-axis models perform better than the x-axis model, the points would be driven on the upper side of the grey dotted diagonal line. Figure 5.3c shows for one that alternative model $T_wT_i\ R_oR_i\ A_wA_i$ in yellow crosses is spread around the diagonal line and rather lower than the other alternative models. Model $T_wT_i\ R_oR_i\ A_wA_i$ may be discarded for now. Figure 5.3c also shows that the two other alternatives seem to perform similarly. Figure 5.3c therefore pictures the log-likelihoods of the two alternative models one against the other. The points are spread around the diagonal line, meaning that the choice between one or the other would be case dependent. A statistical study of their individual residuals could yield additional information on their performance to help for a choice.

Among the three node models, calibration went rather poorly and yielded practical non identifiability, although it could have been expected that higher order models would more likely overfit the data rather than not converge. Details are given in Annex C.

5.2.3.4 Validation of the selected models through statistical analysis of their residuals

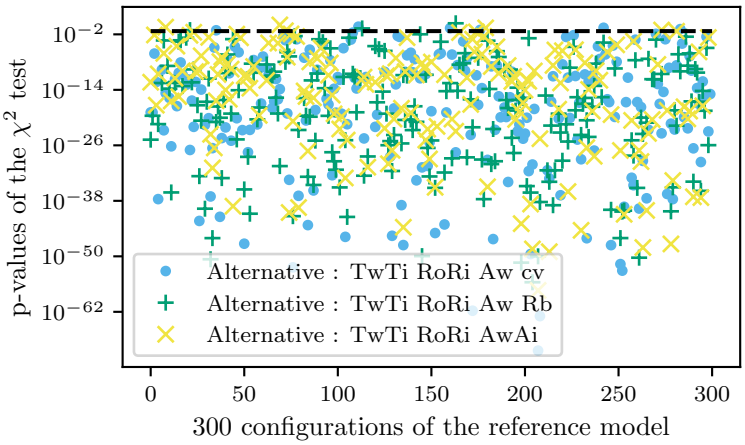
From the model selection based on prediction fitness of the previous section, models $T_wT_i\ R_oR_i\ A_w\ R_b$, $T_wT_i\ R_oR_i\ A_w\ c_v$ and basis model $T_wT_i\ R_oR_i\ A_w$ seemed rather appropriate.

As final validation step, as suggested by the good practice workflow in 2.4.3.2, the white noise property test of the models' prediction residuals is performed. Figure 5.4a shows the autocorrelation of the residuals (with filter) for all 3 models and Figure 5.4b shows the quantile-quantile plot for normality verification of the residuals (on filtered predictions too).

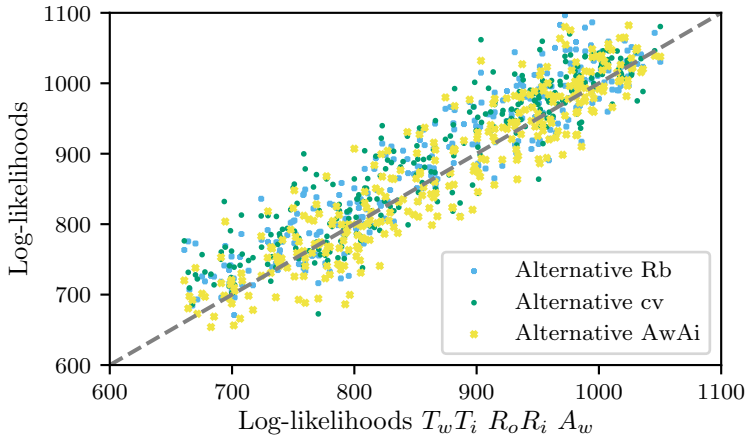
Figure 5.4a shows that the residuals of the two alternative models $T_wT_i\ R_oR_i\ A_w\ R_b$ and $T_wT_i\ R_oR_i\ A_w\ c_v$ are much less autocorrelated than null hypothesis model $T_wT_i\ R_oR_i\ A_w$. This indicates that both the alternatives should be a better fit for the data.

Figure 5.4b shows how the residuals perform compared to a reference normal distribution. Although tails are large and show no normally distributed behaviour, both alternative models' residuals are rather normally distributed. The tails might be larger than for normal distributions because of the temporal resolution of the data. If the building has low characteristic times and the heating power is fast compared to the measurement temporal granularity, the residuals will show a certain lag when the temperature setpoint changes. In other words, the indoor air temperature seems to respond fast to the heating power delivered, which is not well caught by the 8 minutes time step. This would call to future smaller time steps, if the case study or the actual building is thought to have low characteristic times.

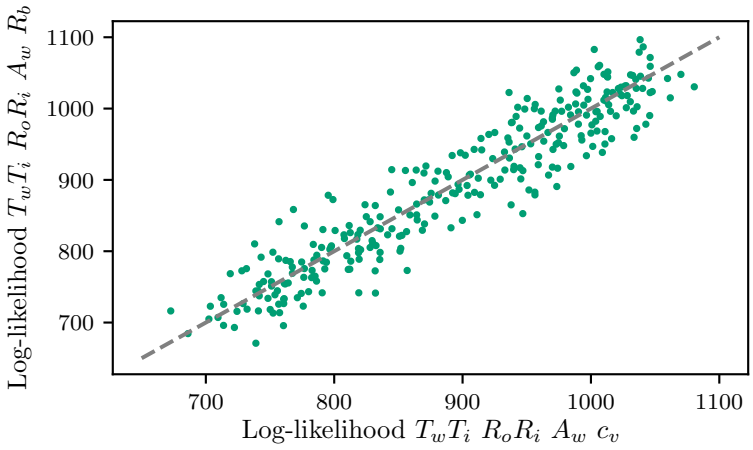
As a conclusion, it seems that models $T_wT_i\ R_oR_i\ A_w\ R_b$ and $T_wT_i\ R_oR_i\ A_w\ c_v$ fit rather satisfactorily the available data, although not perfectly. On the basis of prediction, the good



(a) χ^2 test p-values : values are significant if lower than 0.05 (dotted black line). All 3 tested models are better than T_wT_i R_oR_i A_w.

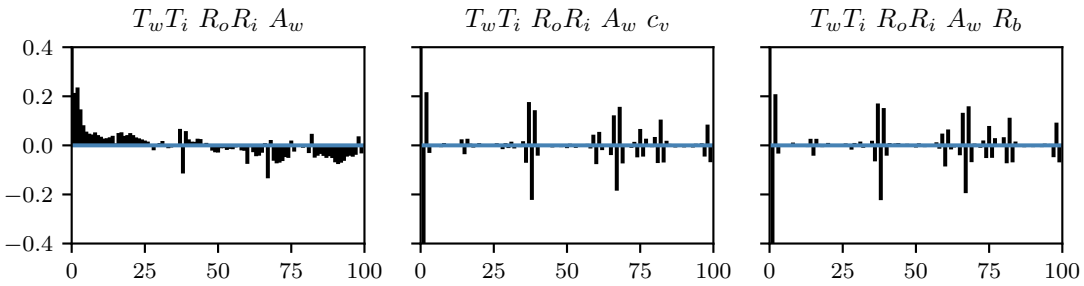


(b) Log-likelihoods of the three alternative models against x-axis null hypothesis model T_wT_i R_oR_i A_w: the yellow cross points are slightly lower than the blue and green points. Alternative model T_wT_i R_oR_i A_wA_i seems not a good as the other two alternatives.

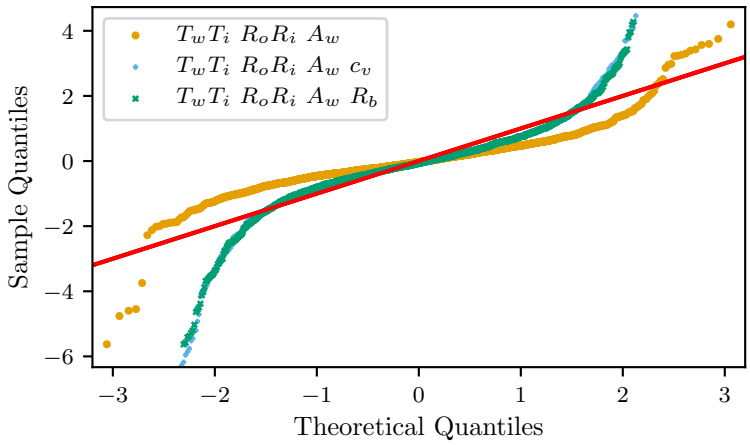


(c) Comparison of the log-likelihoods of the two best alternative models: T_wT_i R_oR_i A_w R_b or T_wT_i R_oR_i A_w c_v

Figure 5.3 – Model selection between alternatives T_wT_i R_oR_i A_w R_b, T_wT_i R_oR_i A_w c_v and T_wT_i R_oR_i A_wA_i against null hypothesis model T_wT_i R_oR_i A_w



(a) Autocorrelation of the residuals of all 3 models. Left plot shows significant autocorrelation while the two others are more satisfactory. Further analysis of their normality would be needed.



(b) Quantile-quantile plot of the filtered prediction residuals for normality verification. The red line gives a reference for what a normal distribution would score. Models $T_wT_i\ R_oR_i\ A_w\ R_b$ and $T_wT_i\ R_oR_i\ A_w\ c_v$ have larger tails than a normal distribution. Model $T_wT_i\ R_oR_i\ A_w$ has smaller tails then both the others, but scores not very satisfactorily in the middle quantiles, which suggest non normal residuals.

Figure 5.4 – Validation tests of the 3 models under consideration. The autocorrelation and the qq-plot for normality of the residuals are ways to verify that the residuals have white noise property, i.e. are sufficient to explain the dynamics in the data.

practice workflow would not prefer any of them. More comprehensive models were found in any case practically non identifiable. They are good candidates to a further interpretability, for both an overall R_{eq} estimation and for individual interpretability assessment.

5.2.4 Convergence of the sensitivity analysis

As underlined in section 3.4.3.3, the convergence and the uncertainty of the sensitivity indices need to be established as to allow drawing conclusions from the values obtained later on.

From the 300 samples, convergence and estimation uncertainty of sensitivity index can be assessed through multiple subsampling on various growing sizes of subsamples. As explained in 3.4.3.3, let us remind shortly here subsampling a Latin Hypercube Sampling will lead to altering its optimal exploration and will exaggerate the uncertainty of the indices. The true variability of the sensitivity indices, although tricky to estimate, will lie in the boundaries given from that methodology.

Figure 5.5 shows the assessment of the convergence and uncertainty of the 3 sensitivity indices of the error in estimation of R_{eq} from model $T_wT_i\ R_oR_i\ A_w$ with respect to walls insulation, attic insulation and air change rate. Variability in the indices seems constant after 200 samples. The convergence is therefore achieved after 200 samples. The uncertainty of the indices is less than 0.1.

A focus on the third figure shows that the sensitivity index related to the air change rate reaches 0.11 with the total 300 samples. With an uncertainty around 0.1, this index can be considered as significant, but any index lower than 0.1 will not.

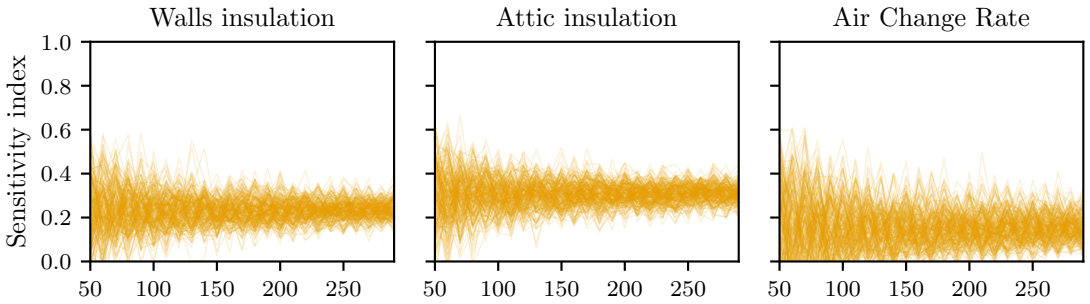


Figure 5.5 – Convergence of the sensitivity indices of the error of R_{eq} with respect to the 6 properties samples

An identical verification is done for models T_wT_i R_oR_i A_w R_b and T_wT_i R_oR_i A_w c_v, shows convergence as well and is shown in appendix D.

5.3 Estimation of the heat losses through ventilation

As heat losses through air change have shorter time characteristics than heat losses through the building envelope, it might be expected that models taking this time characteristic in their formulation might enable to decompose losses through the envelope from losses by air change.

In this section, the model assessment methodology is applied to study the ability of model T_wT_i R_oR_i c_v (Model 2.8) to achieve such a decomposition. Parameter cv should physically represent heat losses through air change whereas parameters R_o and R_i rather account for losses through the envelope, with a longer time characteristic induced by the thermal capacity.

Let us remind here in Table 5.2 the meanings and bounds of the thermal properties used as inputs of the sensitivity analysis.

Abbreviation	Thermal property	Sampling bounds	Units
Floor	Ground floor insulation thickness	[0.05; 0.25]	cm
Walls	Walls insulation thickness	[0.05; 0.25]	cm
Attics	Attic insulation thickness	[0.05; 0.25]	cm
CthFloor	Thermal capacity floor concrete cast	[1.0; 2.0]	J/kgK
CthWall	Thermal capacity brick wall	[0.6; 1.0]	J/kgK
ACH	Air Change Rate	[0.2; 2.0]	h^{-1}

Table 5.2 – Reminder meaning and sampling bounds af all 6 inputs to the sensitivity analysis

5.3.1 Variability of parameters C_w, C_i, R_o, R_i, A_w and cv of model T_wT_i R_oR_i c_v

The inputs variability induce different configurations that may reflect on the estimated values of the parameters of model T_wT_i R_oR_i c_v.

Figure 5.6 shows the variability of parameters R_o, R_i, C_w, C_i, A_w and cv of the state space model with respect to the variability of the 6 inputs.

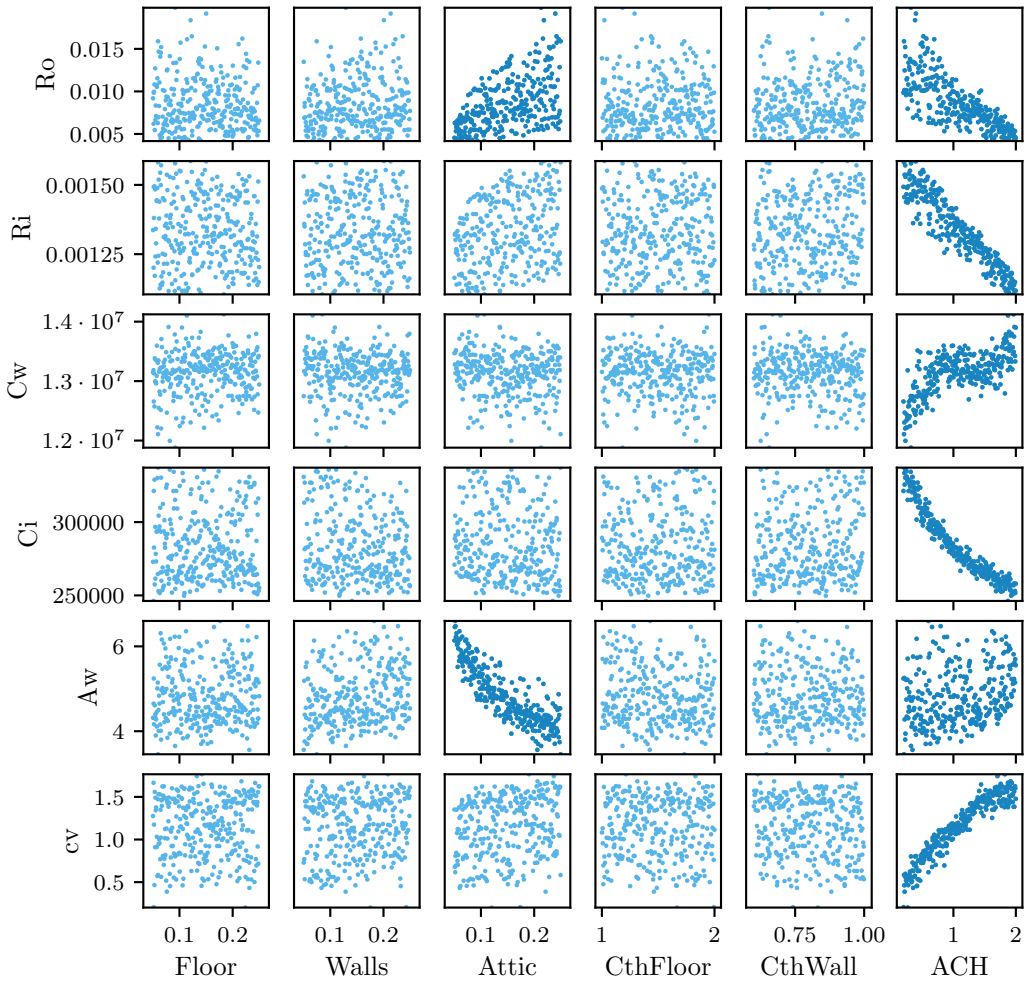


Figure 5.6 – Variability of each parameter of model T_wT_i R_oR_i c_v (y-axis) with respect to the variability of the 6 inputs (x-axis). The darker colours suggest influential inputs on the parameter.

Both thermal resistances R_o and R_i show a correlation with mainly the air change rate and secondly with the attic insulation. The insulation of the walls seems to have on the contrary no influence on the value taken by the resistances, compared to the variability induced by the air change rate and the attic insulation. Parameter cv seems to be only correlated to the air change rate input. The rest of variability looks like random noise.

Both thermal capacities C_w and C_i also have a strong dependence to the air change rate, whereas it could have been expected that the two inputs of thermal capacities (that of the ground floor slab and that of the brick wall) had the major influence.

Parameter A_w shows a large correlation to the attic insulation input. It is likely due to solar irradiation on an almost flat roof causing significant temperature variations. These temperature variations induce heat transfers between the attic and the indoor space that are proportional to the thickness of the insulation between the attics and the indoor space, i.e. to the input variable 'Roof insulation'.

5.3.2 Sensitivity analysis of parameters C_w, C_i, R_o, R_i and cv

Previous section established possible correlations between certain inputs and the studied outputs. A global sensitivity analysis through variance decomposition will in addition quantify the part of variability due to each input variable and assert significance of their influence.

Figure 5.7 shows the sensitivity indices calculated with the RBD-FAST method as detailed in section 3.4.3.1. Let us remind first that the convergence is achieved (see 5.2.4 and D) and that the indices may therefore be interpreted. Sensitivity indices have been calculated for quantities of interest R_o, R_i and cv.

The global variability of R_o is mainly explained by the variability of input air change rate (index 0.58) and to a much lesser extent by input attic insulation (index 0.25). A small part of the variability is explained by effect of the interaction of several inputs, as the total sum of the indices of parameter R_o is not quite close to 1, even considered the uncertainty of the indices.

Similarly, the variability of R_i is explained by the variability of the air change rate (index 0.87) and to a lesser extent the attic insulation (index 0.11). Here,the sum is very close to 1 (somewhat larger which can be explained by the uncertainty of the indices that has been established to 0.1 in 5.2.4), meaning no interaction effects. Variability of cv is only explained by air change rate (index 0.93), the rest being insignificant.

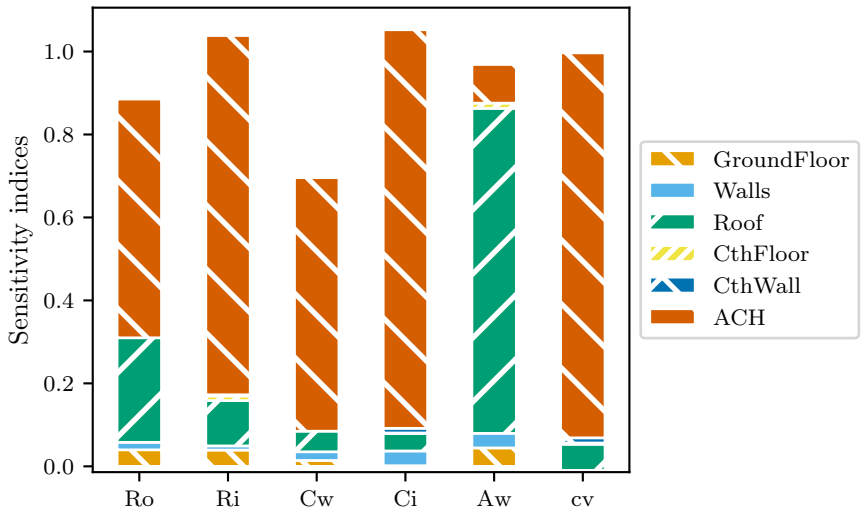


Figure 5.7 – Sensitivity indices of all 6 physical parameters of model T_wT_i R_oR_i c_v

5.3.3 Estimation and physical interpretability of ventilation and infiltration

5.3.3.1 Covariance of the parameters of model T_wT_i R_oR_i c_v

That all R_o, R_i, C_w, C_i and cv show correlation to the air change rate may also suggest that they all just have a strong covariance, from the parameter estimation, which would indeed translate as a fortuitous correlation. To support this hypothesis, Figure 5.8 shows the probability density functions of all 300 covariances between each couple of parameters.

As a reminder from section 1.5.1 in Chapter 1, the covariance between two parameters θ_1 θ_2 is large when around the optimal fit a small variation in θ_1 is correlated to an identical or opposite variation in θ_2 without affecting much the likelihood. Then, it is very likely that both parameters are practically non identifiable. Although the combination of model parameters reaches the best fit, it is impossible to determine the best fitting values for each parameter individually when their covariance is high.

In this figure for example, the covariance of parameters R_o and R_i may be calculated after each calibration. As there are 300 calibrations, all 300 covariances between R_o and R_i may be rendered as a probability density function. Functions higher than 0.5 or lower than -0.5 mean that all 300 covariances are higher than 0.5 or lower than -0.5, meaning that the covariance is significant in all configurations. Figure 5.8 then suggests that R_o R_i, R_o cv and R_i cv have strong covariances.

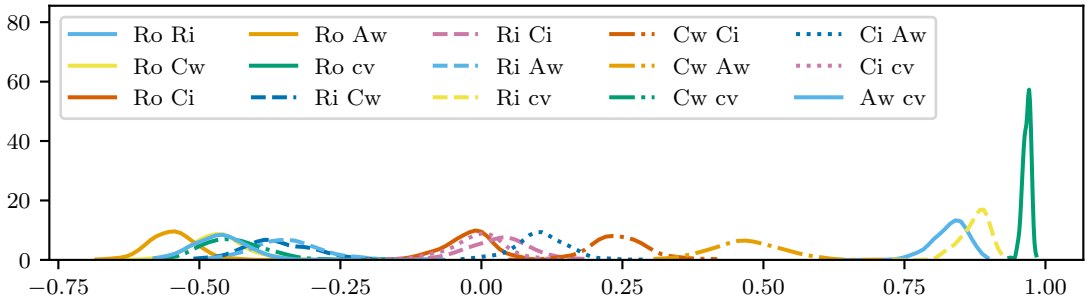


Figure 5.8 – Covariance between parameters of model TwTi RoRi cv: parameter cv has a large correlation with parameters R_o and R_i, as are R_o and R_i together

To a lesser extent, set R_o A_w also shows significant negative covariances. Taken individually, parameters R_o, R_i and cv should probably not be physically interpreted.

5.3.3.2 Physical interpretation of parameters cv and $R_o + R_i$

Model $T_w T_i$ $R_o R_i$ c_v introduces parameter c_v to take into account ventilation and infiltration dynamics as a heat flux on the indoor temperature node. The flux is defined as follows:

$$\Phi_v = c_v V \rho c_p (T_{ext} - T_{int}) \tag{5.2}$$

In 5.2, V the heated volume of the building, ρ the volumic mass of air and c_p the specific heat capacity of air are known, which leaves one unknown in the equation to estimate: c_v. With this definition, c_v is then an average air change rate over the duration of the experiment, with dimension h^{-1}. Let us shortly mention that in the simulation conditions, the air change rate actually varies in time with the temperature difference and with the wind speed.

As cv could be interpreted as an averaged air change rate, let us take a closer look to the variability of parameter cv in Figure 5.9 where the dotted 45° grey line would be the perfect estimation of the air change rate. Parameter cv is positively correlated to the air change rate: the higher the air change rate the higher the parameter cv. However, low air change rates induce a systematic overestimation and high air change rate a systematic large under-estimation.

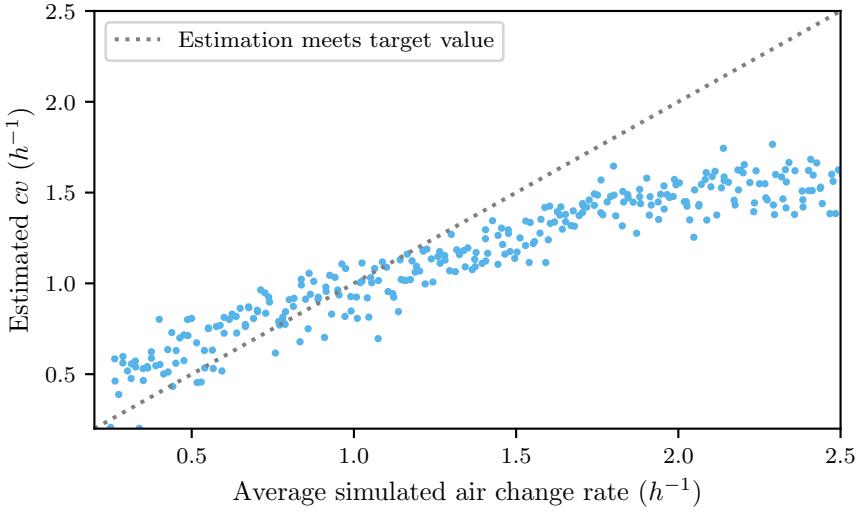


Figure 5.9 – Estimation of parameter cv for varying values of air change rate

Seeing the significant covariances between all resistance parameters and seeing that variability of the estimation of c_v and its error to the simulated value of air change rate, it is expected that ventilation and infiltration losses cannot be properly identified from the current design of experiment. Figure 5.10 shows the estimated sum $R_o + R_i$, $R_{ventilation}$ and R_{eq} against their target values for each of the 300 reference model configurations.

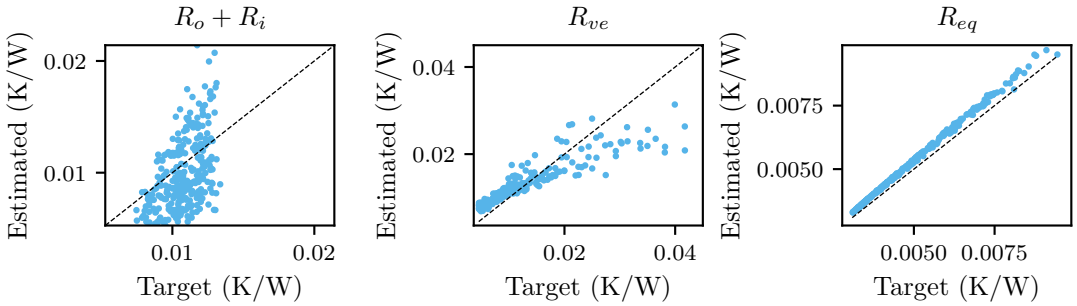


Figure 5.10 – Interpretability of $R_o + R_i$, R_{ventil} and R_{eq} for model T_wT_i R_oR_i c_v

The sum $R_o + R_i$, supposedly representing the heat resistance between the indoor space and the exterior, shows no significant correlation with the target value. As for ventilative losses, the order of magnitude seems to be rather well estimated, but with a high variability: for an identical target value, the error spreads from -50% up to 90%. In overall however, the equivalent thermal resistance is well in agreement with the target value.

5.3.4 Conclusions on decomposing ventilative heat losses

Section 5.3 applied the numerical model assessment methodology to assess the ability of model T_wT_i R_oR_i c_v to separately identify heat losses through ventilation from the rest of the heat losses, by means of the estimation of an averaged air change rate parameter. The results indicate that model T_wT_i R_oR_i c_v estimates the air change rate parameter c_v with errors up to 90% in absolute regardless of the target rate, implying that the order of magnitude is rather well estimated, but that physical interpretation would be risky.

Highly correlated parameters, visible in significant covariances, are in this study a clear hint that the physical interpretation of the parameters separately is not possible. The performed sensitivity analysis was a confirmation that the high covariances translate in unlogical variability in the parameter estimations with respect to the changes in the reference model.

Model T_wT_i R_oR_i c_v was however found in the model selection process as to fit satisfactorily the data. The results of this section confirm it by showing that the overall thermal resistance R_{eq} is in good agreement with the target value. This indicates that high covariances between parameters may still bear physical meaning when said parameters are combined, as in R_{eq}.

5.4 Estimation of heat losses towards unheated neighbouring space

Unheated neighbouring spaces offer different boundary conditions than the outdoors. Measuring the temperature of such spaces would give more information on the heat dynamics of the building with its surroundings and might enhance either the interpretability of the overall thermal resistance or even allow estimation of the direct heat losses to neighbouring spaces.

Earlier results (Juricic et al., 2019) on a simple parallelepiped reference model estimated the

heat loss transfers with in particular the two following linear regressions $P_{heating} = HTC \times (T_{int} - T_{ext})$ and $P_{heating} = H_{dir} \times (T_{int} - T_{ext}) + H_u \times (T_{int} - T_{crawlspace})$, assuming on average daily steady state conditions, as suggested by co-heating tests exploitation.

Results showed that, in overall, taking into account measured temperatures in the crawl space leads to a better HTC estimation than not measuring it. However, interpretability of the heat losses to the crawl space only is not poor with the linear regression models used. Physical interpretability of the heat losses to the outdoors was poor too.

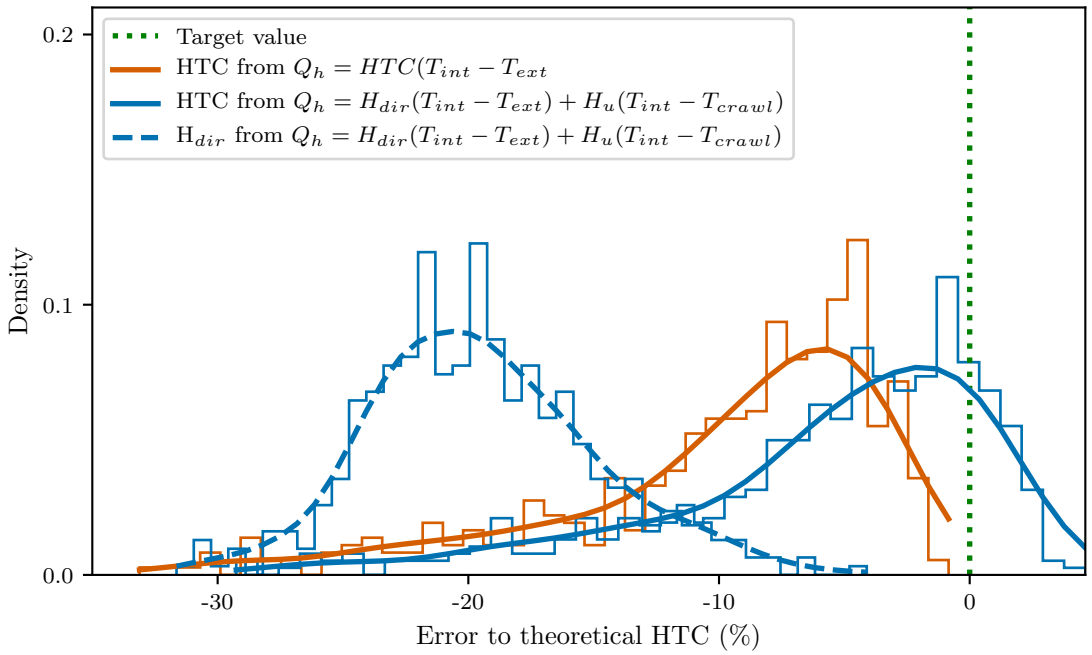


Figure 5.11 – Error variability of HTC estimation from linear regressions (Juricic et al., 2019)

Linear regressions such as the models used in Juricic et al. (2019) however make no use of the dynamics in the data, even in a non intrusive experiment. State space models might therefore achieve better estimation of the overall thermal resistance and of heat losses towards on the one hand the neighbouring space and on the other hand the outdoors. This section applies the model assessment methodology to study, from a more complex reference model, the interpretability of heat losses to an unheated crawl space by an adequate state space model: $T_wT_i\ R_oR_i\ A_w\ R_b$.

Model $T_wT_i\ R_oR_i\ A_w\ R_b$ is tested in the model assessment framework to assess its ability to estimate on one side the direct heat losses to the outdoors with the sum $R_o + R_i$ and on the other side the heat losses to an neighbouring unheated space with parameter R_b.

5.4.1 Variability of the estimated parameters of model TwTi RoRiRb

If model $T_wT_i\ R_oR_iR_b$ were identifiable, its parameters would vary in agreement with physical changes in the reference model. Figure 5.12 shows how the estimated parameters vary according the changes in the reference model.

First, all parameters R_o, R_i, R_b, C_w, C_i and A_w show a significant variability with changes in value of the air change rate. Correlation between R_o and R_i with the air change rate was expected. Indeed, there is in this state space model no parameter designed to represent ventilation and infiltration, suggesting that ventilation and infiltration phenomena will be lumped in the indoor-outdoor branch $R_o + R_i$. Correlations between the air change rate and the other four parameters R_b, C_w, C_i and A_w suggests on the other hand that ventilation phenomena are also lumped on them, or at least that there are some strong covariances in the parameters' estimation.

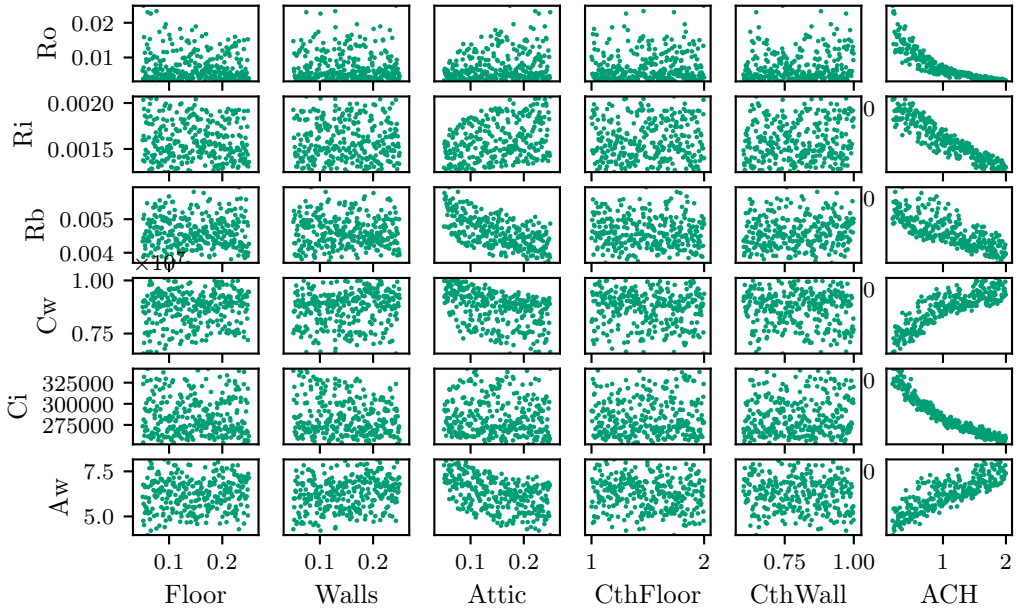


Figure 5.12 – Variability of parameters R_o, R_i, R_b, C_w, C_i and A_w for model T_wT_i $R_oR_iR_b$

Secondly, Figure 5.12 shows no other influence on parameter estimation but that of the insulation in the attic, which has significant influence on parameter R_b and A_w. As in section 5.3, that parameter insulation in the attic is correlated to value of parameter A_w is not surprising, as solar irradiation has an influence on the temperature in the attic, and therefore has an influence on the heat dynamics with the indoor space. However, R_b was not expected to have a correlation with the attic insulation, but rather with the ground floor insulation, which needs not to be the case. A global sensitivity analysis will confirm the conclusions inferred from Figure 5.12.

5.4.2 Sensitivity analysis of the thermal resistance estimation

A sensitivity analysis is performed on the parameters' estimations to quantify the effect of the variability in the reference model on parameter estimation. A global sensitivity analysis is performed with RBD-FAST, as detailed in Section 3.4.3.1.

First, as all sums of indices are close to 1, meaning that the variability observed in the parameter estimation is due to first order effects of the inputs. Sums larger than 1 are due to uncertainties of the sensitivity indices calculation. In this study, the uncertainties were indeed quantified in Section 5.2.4 to be close to 0.1.

Nevertheless, the correlations suspected previously in 5.4.1 remain exact: the variability in air change rate is the main contributor to the variability of all parameters of model T_wT_i $R_oR_iR_b$. Secondly, attic insulation is significantly influent on the estimation of parameter R_b and A_w as well as on C_w.

Let us remind here that the variability of the air change rate in the reference model is significantly large as it spread between 0.2 h^{-1} and 2 h^{-1}. It may therefore be reason why its influence is so large on all parameters. However, without calculating new indices from a subsample of the Latin Hypercube Sampling, which would yield a possibly erroneous analysis, a look at the variability of samples with air change rate below $0.8h^{-1}$ in Figure 5.14 shows a similar pattern: insulation of the attic plays the largest influence on parameter estimation. The rest of the variability looks like white noise.

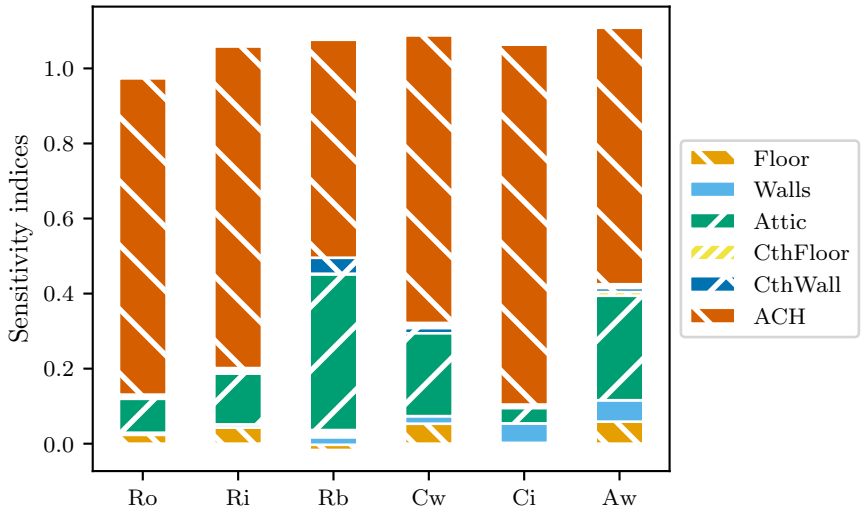


Figure 5.13 – Cumulated sensitivity indices of R_o, R_i, R_b, C_w, C_i and A_w for model T_wT_i $R_oR_iR_b$

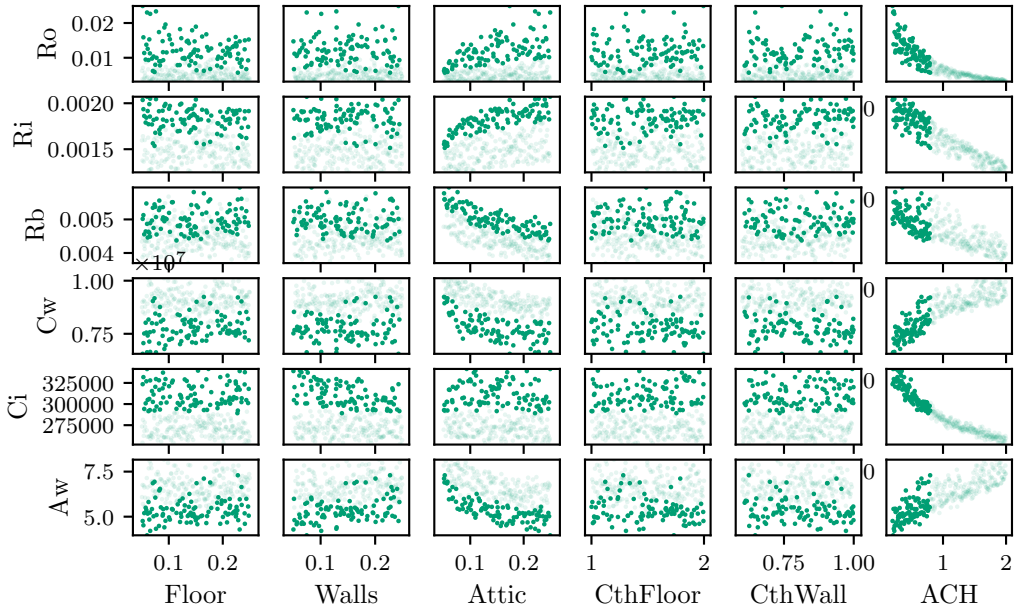


Figure 5.14 – Variability of parameters R_o, R_i, R_b, C_w, C_i and A_w for model T_wT_i $R_oR_iR_b$ with a highlight on small air change rate configurations

5.4.3 Identifiability and interpretability of the proposed model

5.4.3.1 Covariance of the parameters

The sensitivity analysis showed unexpected variability of the parameters with changes in the reference model. Figure 5.15 shows the covariance values of the parameters after the estimation process. As a short reminder, high covariance between two parameters θ_1 and θ_2 means that it is very unlikely that θ_1 and θ_2 are practically identifiable.

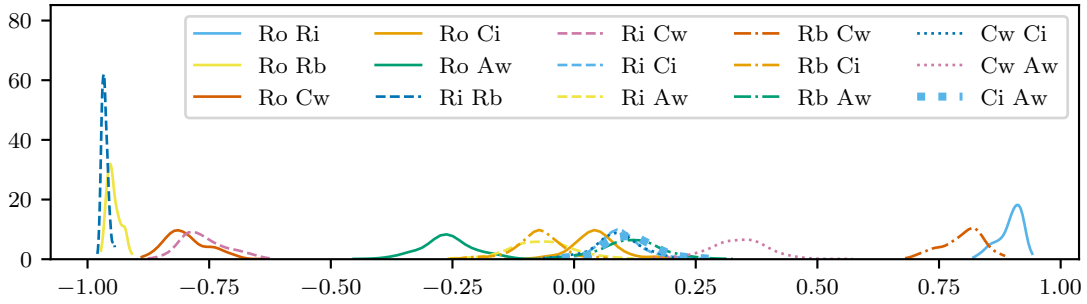


Figure 5.15 – Covariance between parameters of model T_wT_i $R_oR_iR_b$

Figure 5.15 shows that 4 parameters have significantly high covariances: R_i, R_b, R_o and C_w. Each of the 300 calibrations result in covariances between R_oR_b, R_iR_b and R_oR_i larger in absolute than 0.8. Then, C_wR_o, C_wR_i and C_wR_b too have large covariances. It is therefore unlikely that these 4 parameters at least have any physical interpretability.

5.4.3.2 Physical interpretation

Model T_wT_i $R_oR_iR_b$, by taking into account the temperature in the crawl space, may separately identify thermal resistance of the envelope towards the outdoors through the sum $R_o + R_i$ from the thermal resistance towards the crawl space through parameter R_b.

Figure 5.16 shows for $R_o + R_i$, R_b and for the overall thermal resistance R_{eq} their estimation against their theoretical target values. As a comparison, each black 45°dotted line represents an estimation in perfect agreement with the target value.

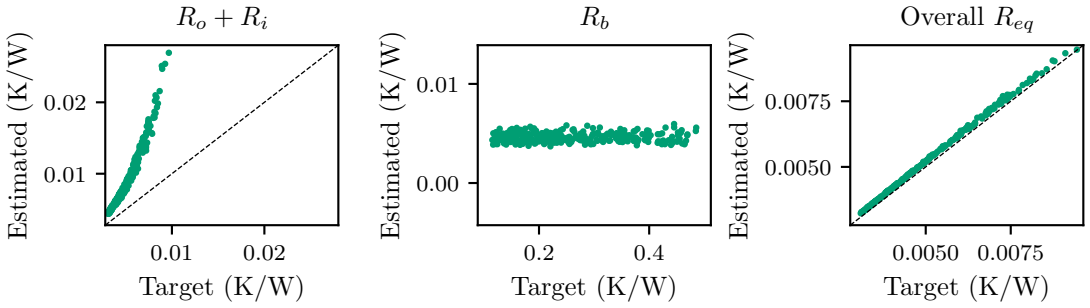


Figure 5.16 – Physical interpretation of thermal resistances of model T_wT_i $R_oR_iR_b$

Estimation of the overal thermal resistance R_{eq}, as shown in Figure 5.16 although slightly overestimated, shows an excellent agreement to the target value. On the contrary, estimation of $R_o + R_i$ and R_b do not show any agreement with the target values, with errors up to 200 % from the target value.

5.4.4 Conclusions on identification of heat losses to neighbouring spaces

The model assessment methodology has been applied to model $T_wT_i\ R_oR_iR_b$ to assess its ability to separately identify heat losses towards a neighbouring unheated space from losses to the outdoors. It has been found that the overall thermal resistance is satisfactorily estimated from the available data. However, the ground floor parameter of the model supposedly representing the thermal resistance towards the crawl space cannot be physically interpreted. This outcome is in agreement with the findings from the sensitivity analysis and the covariance analysis, both suggesting that the individual parameters do not vary adequately to variations in the reference model.

These results would suggest that from measured temperatures only, parameters model T_wT_i $R_oR_iR_b$ are not physically interpretable. Whether this outcome is valid with other building configurations, like measured temperatures in the attics or in a neighbouring garage or cellar is less certain. Heat transfer dynamics between such spaces can be quite different, either because the temperatures may show different profiles than that in an unheated crawl space, or because surface heat transfers are different on vertical partition walls or with ceilings. These differences justify the perspective of an application of the model assessment framework on other reference models for assessing the identifiability of model $T_wT_i\ R_oR_iR_b$ with different types of neighbouring spaces.

5.5 Conclusion and Bayesian prospects

This chapter applied the model assessment framework defined in Chapter 3 to assess the extent to which trained models may be physically interpreted, regarding either their thermal resistance parameter estimates individually or their overall thermal resistance.

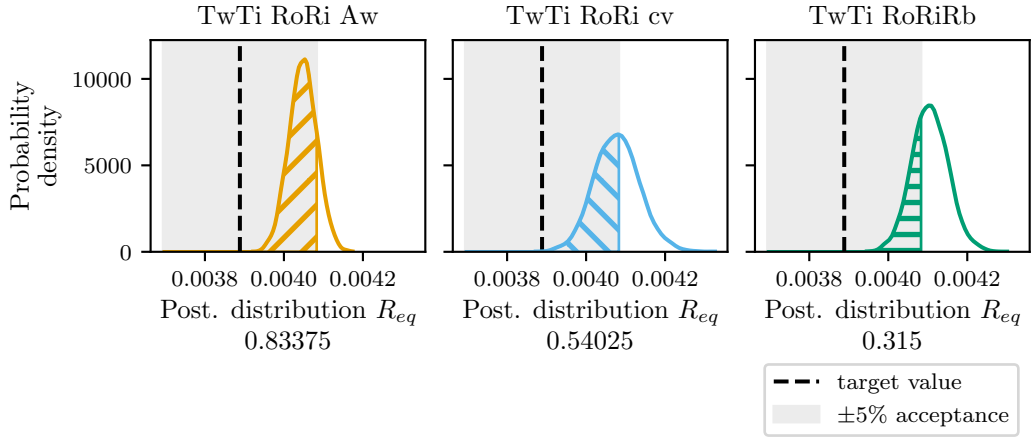
The thermal properties of the reference model envelope have been modified in such design that global sensitivity analysis could be performed. All configurations have been simulated and served as training data for 2 stochastic RC models, selected for their prediction performance, in agreement with the good practice workflow defined in section 2.4.3.2.

Both models have then been assessed for their ability to have physically interpretable estimations of heat transfers through on one hand ventilation with model $T_wT_i\ R_oR_i\ A_w\ c_v$ (2.8) and on the other hand through unheated neighbouring spaces, an unheated crawl space in the case study used, with model $T_wT_i\ R_oR_iR_b\ A_w$ (2.10). The expectation is to see parameter estimates that represent ventilation (respectively heat losses to the crawl space) vary in agreement with variations in the reference model of the air change rate (respectively the ground floor insulation).

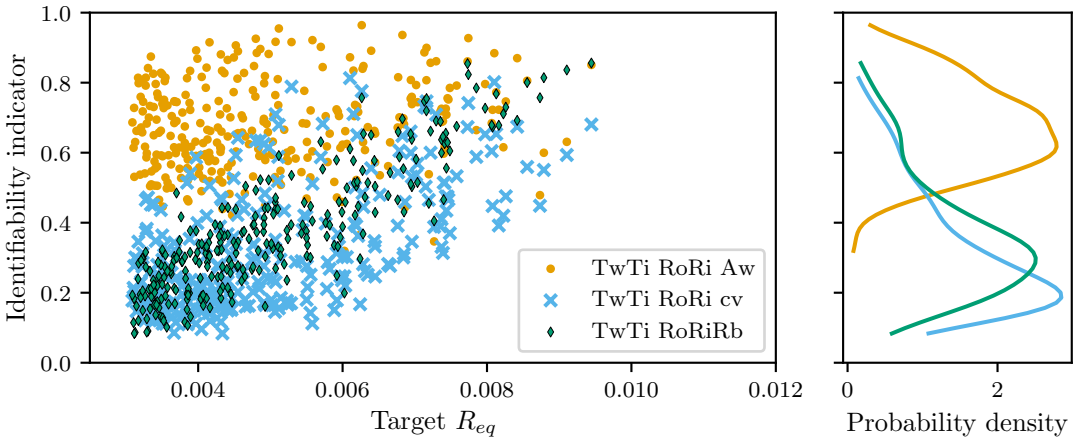
The results of the sensitivity analysis presented in the chapter found that the expected behaviour of the parameter estimates did not match the variations of the reference model. Variability of the air change rate and insulation thickness in the attics were found to be the main influential thermal parameters on any parameter estimates, resistances or thermal capacities. Comparison of the target values and the parameter estimates found in both applications that the parameter estimates should be physically interpreted. Thermal resistances representing the heat transfers by ventilation were concededly found approximately at an acceptable order of magnitude. However, similarly to the other estimates, the sensitivity analysis as well as the covariances of the estimates strongly suggest that physical interpretation cannot be achieved.

A shortcoming of this study might be that the air change rate varies in very large bounds, which might explain why it has such influence on the parameter estimates. Air change rates higher than 1 h^{-1} have indeed a large influence on the energy balance. A focus on smaller variations of the air change rate would allow a finer analysis of the influences of the other input parameters, such as that of the thermal capacities which influence was not visible. Let us for the matter remind that

although invisible, it does not mean that thermal capacities have no influence, it simply means that the large influence of the air change rates flattens the other influences.



(a) Comparison of three models in a single configuration: a single dataset serves as training data for 3 models in a Bayesian approach. The posterior distributions are therefore sampled and known, which allows to assess the interpretability of each model with the interpretability indicator (basically the area under the curve in the grey target area)



(b) Model comparison for all 300 configurations: interpretability indicator of R_{eq} against target R^*_{eq} and distribution of the indicators for the three models. In orange, interpretability indicators of model T_wT_i R_oR_i A_w are in overall quite high and in general higher than for the other two models.

Figure 5.17 – Model comparison through the interpretability indicator

Furthermore, large covariances are found to be good posterior indicator that the parameters are physically independently not interpretable. It does however not imply that the models are unable to achieve a correct estimation of the overall thermal resistance. Figure 5.17 precisely shows how both models can be compared to the simple two nodes model T_wT_i R_oR_i A_w with respect to their interpretability thanks to the interpretability indicator defined in 3.3.

Figure 5.17a first recalls how the indicator is calculated from the posterior distribution (with Bayesian approach calibrations). Depending on the model, the posterior distribution of the overall thermal resistance is determined and compared to the target value and its 5 % acceptable uncertainty in grey. The area under the curve that matches the grey area is the interpretability indicator.

The indicator is calculated for each model (3×300 trained models) and allows a global comparison of their ability to estimate the overall thermal resistance R^*_{eq}, shown in Figure 5.17b. From the figure can be inferred that all 3 models achieve satisfactory and equivalent results in

configurations with high thermal resistances. In low thermal resistance configurations, the simple two nodes model performs rather well compared to the others, although it was not an excellent fit to the data regarding the statistical validation tests from section 5.2.2.2.

Noteworthy is that the interpretability indicator is relative to the 5 % band around the target value. It is the reason why model $T_w T_i\ R_o R_i R_b\ A_w$ (2.10) seems to perform rather badly whereas it seemed from Figure 5.16 that the estimation was rather close. The estimations were then probably quite close, but always somewhat outside of the 5 % target band around the target value. If the estimation of the overall thermal resistance were fundamentally inaccurate, the interpretability indicator would be quasi-null.

All 3 models provide then moderate to good estimation of the overall thermal resistance but their parameters are individually not interpretable. It does not come as a surprise because the model selection tools are based on the likelihood on the training data. Such model selection process is a measure of a fit to the overall behaviour, inducing a good fit of an overall thermal characteristic as is R_{eq}. Model selection based on likelihood yields then appropriate R_{eq} but this application of the model assessment framework suggests that it is no proof for interpretable identification of the heat transfers.

Put in perspective, it does not mean that with more informative data, both models would not achieve interpretability of their parameters. Appropriate heating signals with uncorrelated frequencies, especially with respect to the weather conditions characteristic frequencies, or additional measurements might lower the covariances and maybe improve the interpretability.

Further work could also cover other designs and structure type : external insulation, several storeys with the afferent stack effect influence, different opaque/window surfaces,...

General conclusion and research prospects

Wrap up of the findings of this work

Estimation of the thermal performance of a building from on-site measurements faces two major scientific issues: the uncontrolled measurement conditions provide poorly informative data and the occupants might lead to significant bias in the collected data. This thesis has precisely tackled the former: feasibility of estimating the thermal performance of a building envelope from poorly informative data. This work proposes to address it with the use of stochastic RC models. The issue of identifiability of low order RC models has first been examined with respect to their structural formulation and to the quality and quantity of the data available. Poorly informative data and model discrepancy to the actual building thermal behaviour have been found to still be a threat to physical interpretation of the estimation. An original numerical framework is subsequently proposed to assess the physical interpretability of parameters of stochastic RC models as well as the repeatability of their estimation. With the prospect of a building energy retrofit, the findings of this work aim in the end at driving the choice of appropriate RC models for identifying target thermal characteristics of a building with satisfactory accuracy.

Background literature review on similar inverse problems and how to solve them

Chapter 1 established how the objective of thermal performance estimation in uncontrolled conditions data constitutes an ill-posed problem because of the poorly informative nature of the collected data. This implies that, depending on the data, thermal performance might not be estimable and if it is, the accuracy of such estimation might be highly sensitive to the measurement errors. Chapter 1 further proposed to use stochastic RC models. Their physics-based formulation may allow physical interpretation, and literature suggests that it may be done from reasonably short data collection. Algorithmic tools for numerical estimation have also been presented, in both frequentist and Bayesian approaches.

Distinguishing identifiability from interpretability

Numerical estimation of a target thermal characteristic means finding the solution of the inverse problem thus defined. The existence and unicity of the solution, i.e. identifiability of the model parameters, is in any case not guaranteed and needs to be verified before any physical interpretation. As laid out in Chapter 2, identifiability is a twofold issue. It relies first on a structural property of the chosen model: given ideally informative data, parameters may individually and uniquely be identified. Then, the quality and quantity of the information in the data conditions how the estimates may be in practice determined. The differences between frequentist and Bayesian approaches upon tackling an identifiability issue have also been discussed. Literature suggests in particular that estimation in a Bayesian approach is feasible also on non identifiable problems, thanks to its regularisation ability. While practical identifiability may be established, model discrepancy and possibly biased measurements still constitute threats to physical interpretation of the estimates. A workflow of good practice for thermal characterisation therefore also relies on model selection and validation as well as adequate data collection.

An original methodology for model assessment

Physical interpretation is not necessarily implied by structural and practical identifiability. An original methodology has therefore been proposed in Chapter 3 to assess in numerically simulated conditions the physical interpretability of RC models. A computer-based building energy model serves as reference model and is simulated in known conditions. The simulation output of the reference model serves as training data for model calibration. As such, the procedure lets us relate variations of estimated parameters to variations of the reference values. The novelty of this methodology is to simulate the reference model multiple times in conditions that are designed as to enable global sensitivity analysis. An original quantitative indicator for physical interpretability is defined to account for both accuracy and uncertainty of the estimations and allows model assessment and comparison.

Repeatability of an estimation by RC models under variable weather

The methodology is applied in Chapter 4 to evaluate the influence of weather variability on the estimation of thermal performance of an evaluation. In particular, the assessment framework is used to determine minimal duration for repeatable results, regardless of the meteorological conditions. Synthetic weather datasets have therefore been used, stochastically generated as to enable global sensitivity analysis of 6 weather variables. Repeatability of the estimation of the overall thermal resistance of the envelope can be achieved in uncontrolled conditions, in this case study with 11 days training. Interpretability is also found to be satisfactory. The remaining variability is due to the variability of the outdoor temperature and wind speed, which can be explained by the large air change rate in the case study used as reference model.

Interpretability of RC model parameters for heat loss decomposition

Finally, the methodology is used in Chapter 5 to assess the extent to which parameters of RC models may individually be interpreted. Specifically, the ability of RC model to decompose losses through ventilation or losses towards an unheated crawl space has been assessed by two different adequate models. To better grasp how interpretability is influenced by the envelope configuration, the reference model undergoes changes with respect to its thermal insulation, air change rate and thermal capacities. It has been found that neither were able to perform such identification. However, three RC models were compared for their interpretability of the overall thermal resistance and all achieved rather satisfactorily estimations.

In a nutshell, this work aimed at establishing to what extent stochastic RC models were able to estimate the thermal performance of a building envelope from weather, indoor temperatures and heating power designed as to not disturb occupants. The overall thermal resistance can be robustly estimated from minimally 11 days training data and achieves very satisfactory interpretability. Finer decomposition of the heat losses has however not been found possible from poorly informative data. Finally, regardless of the conditions of data collection, the good practice workflow for meaningful calibration remains valid and in any case necessary for any thermal characterisation.

Research prospects

In a close future, the results could be supplemented by extending the methodology to other building typologies as to examine how thermal inertia and architectural specificities influence the estimation of the thermal performance. Other typologies would indeed widen the issues to tackle, just to name a few:

- Thermal stratification and temperature differences in the different rooms become all the more problematic in large buildings, notably with several storeys. It implies that some part of the information on the thermal behaviour of the building will be missing. These temperature discrepancies seem indeed to have an impact on the estimation's accuracy (Senave et al., 2020b; Thébault and Bouchié, 2018) and invite to further research on the subject. The issue of measurement representativeness actually also apply to the measurement of wind speed and solar irradiation, as well as measurement of the heating power when it is derived from overall energy use data. A quantification of the effect on accuracy of measurement discrepancy would be valuable and would help understanding the extent of the issue. This would help concentrate the effort on tackling on the most influential measurement discrepancies. It could be done by design of appropriate sensor placement in order to reduce uncertainty, although extensive measurements would certainly be more intrusive. At least, acknowledgement of the measurement discrepancy in the modelling process is necessary to reflect the associated uncertainty in the estimation results.
- Measurement of heating power delivered in a single room for centralized heating systems also pushes further the issue of measurement discrepancy;
- Thermal characterisation of the envelope in apartment blocks, large office buildings, or hospitals questions the measurements representativeness. In particular, it cannot be expected at all that all zones follow identical temperature set point schedules and deploying sensors in all rooms is most probably cost prohibitive. This implies that a model would be trained for each measured zone. But then, sampling of zones to measure raises the issue of a compromise between efficiency and representativeness. Then, while sampling may be an option, the question of the accuracy of RC model training in a single zone remains, when there are more heated neighbouring zones than surfaces towards the exterior.
- The question of buildings bound by large surfaces towards heated neighbouring spaces, such as in offices and hospitals but also in terraced houses, is also left to tackle with RC models trained from non intrusive data.
- As suggested in section 2.4.2.1, there remains questions on the most informative indoor temperature set point schedule. If there were significantly more informative and yet occupant-friendly schedules, it becomes relevant to control such schedules when the experiments are performed, in order to lower uncertainties.

In the prospect of assessing the suitability of retrofit strategies, this work could also be applied to the identifiability of the dynamic characteristics of the envelope. Indeed, energy performance is not only influenced by a static thermal resistance but also by the dynamic response of the envelope. This work suggests that the dynamic nature of stochastic RC models would then be adequate for such estimation, although identifiability and interpretability will have to be assessed on basis of data collected in uncontrolled non intrusive conditions.

Next, in the wider prospect of achieving thermal characterisation of a building envelope from data collection under occupancy, the results of this thesis are believed to enlighten the extent of the scientific problem thus defined. Very basically, measurements in an occupied space imply

issues with first data collection which will be aggregated and possibly biased and secondly with unpredictable and barely measurable heat and moisture production as well as unpredictable occupancy related activity such as open doors and windows.

In a non intrusive and uncontrolled experimental framework, data exploitation will first necessarily rely on disaggregation of the energy consumptions to distinguish energy use for heating from energy use for electrical appliances. If it cannot be done, it constitutes an additional measurement discrepancy error with most probably a large effect on accuracy of a thermal characterisation.

Regarding interaction of the occupants on the envelope and on the heat and moisture balance, it implies major model discrepancy as the building envelope has multiple states unaccounted for in a usual RC model. Accuracy and uncertainty are therefore not guaranteed to be to the least satisfactory. The question of repeatability under the variability of the actions of occupants is also questioned, which calls for further research on a Bayesian based indicator for convergence of the estimation, as suggested in 4.6.

These occupancy related issues invite to at least account for uncertainties either in the form of a global model discrepancy term as in Bayesian calibration in the sense of Chong and Menberg (2018); Heo et al. (2012), although physical interpretability has been found to be an issue (Chong and Menberg, 2018), or use Gaussian Latent Force Models (Sarkka et al., 2019) that account for unmeasured influential input variables directly in the state space equations.

Both propositions are also arguments in favour of a Bayesian approach. Indeed, introduction of unknown, unmeasurable and systematic uncertainties will certainly have an impact on practical identifiability and all the more on the algorithmic ease to make estimations. As a consequence, physical interpretability would even less be within reach. A Bayesian approach through careful prior choices would at least act as regularisation in the inverse problem.

Bibliography

Afram, A. and Janabi-Sharifi, F. (2015). Black-box modeling of residential HVAC system and comparison of gray-box and black-box modeling methods. *Energy and Buildings*, 94:121–149.

Agbi, C., Song, Z., and Krogh, B. (2012). Parameter identifiability for multi-zone building models. In *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pages 6951–6956. IEEE.

Alzetto, F., Meulemans, J., Pandraud, G., and Roux, D. (2018a). A perturbation method to estimate building thermal performance. *Comptes Rendus Chimie*, 21(10):938–942.

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

Andersen, P. D., Jiménez, M. J., Madsen, H., and Rode, C. (2014). Characterization of heat dynamics of an arctic low-energy house with floor heating. *Building Simulation*, 7(6):595–614.

Andrade-Cabrera, C., Burke, D., Turner, W. J., and Finn, D. P. (2017). Ensemble Calibration of lumped parameter retrofit building models using Particle Swarm Optimization. *Energy and Buildings*, 155:513–532.

Anguelova, M. (2007). *Observability and identifiability of nonlinear systems with applications in biology*. PhD thesis, Chalmers University of Technology and Göteborg University.

Anstett-Collin, F., Goffart, J., Mara, T., and Denis-Vidal, L. (2015). Sensitivity analysis of complex models: Coping with dynamic and static inputs. *Reliability Engineering and System Safety*, 134:268–275.

Arendt, P. D., Apley, D. W., Chen, W., Lamb, D., and Gorsich, D. (2012). Quantification of Model Uncertainty : Calibration, Model Discrepancy and Identifiability. *Journal of Mechanical Design*, 134(10):40.

ASHRAE (2013). *ASHRAE Handbook - Fundamentals*. ASHRAE.

Audoly, S., Bellu, G., D'Angio, L., Saccomani, M., and Cobelli, C. (2001). Global identifiability of nonlinear models of biological systems. *IEEE Transactions on Biomedical Engineering*, 48(1):55–65.

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

Bagheri, A., Feldheim, V., Thomas, D., and Ioakimidis, C. S. (2017). The adjacent walls effects in simplified thermal model of buildings. *Energy Procedia*, 122:619–624.

Bauwens, G. (2015). *In situ testing of a building ' s overall heat loss coefficient*. PhD thesis, KU Leuven.

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

Bayarri, M. J. and Berger, J. O. (2004). The interplay of Bayesian and frequentist analysis. *Statistical Science*, 19(1):58–80.

Bellman, R. and Åström, K. (1970). On structural identifiability. *Mathematical Biosciences*, 7(3-4):329–339.

Bellu, G., Saccomani, M. P., Audoly, S., and D'Angiò, L. (2007). DAISY: A new software tool to test global identifiability of biological and physiological systems. *Computer Methods and Programs in Biomedicine*, 88(1):52–61.

Bertsekas, D. and Tsitsiklis, J. (2008). *Introduction to Probability*. Athena Scientific.

Betancourt, M. (2016). Diagnosing Suboptimal Cotangent Disintegrations in Hamiltonian Monte Carlo. *arXiv*.

Betancourt, M. (2017). A Conceptual Introduction to Hamiltonian Monte Carlo. *arXiv*.

Betancourt, M. (2018). Calibrating Model-Based Inferences and Decisions. *arXiv*, pages 1–35.

Biddulph, P., Gori, V., Elwell, C. A., Scott, C., Rye, C., Lowe, R., and Oreszczyn, T. (2014). Inferring the thermal resistance and effective thermal mass of a wall using frequent temperature and heat flux measurements. *Energy and Buildings*, 78:10–16.

Bienvenido-Huertas, D., Moyano, J., Marín, D., and Fresco-Contreras, R. (2019). Review of in situ methods for assessing the thermal transmittance of walls. *Renewable and Sustainable Energy Reviews*, 102:356–371.

Big Ladder Software (2016a). EnergyPlus 8.6 Engineering reference Infiltration/Ventilation.

Big Ladder Software (2016b). Input Output Reference - Weather Data Hourly Interpolation.

Big Ladder Software (2016c). InputOutput Reference - Group – Simulation Parameters.

BigLadderSoftware (2020). https://bigladdersoftware.com/epx/docs/8-6/engineering-reference/conduction-through-the-walls.html#conduction-through-the-walls.

Bohlin, T. and Graebe, S. F. (1995). Issues in nonlinear stochastic grey box identification. *International Journal of Adaptive Control and Signal Processing*, 9(6):465–490.

Boisson, P. and Bouchié, R. (2014). ISABELE method : In situ assessment of the building envelope performances. In *System simulation in buildings*, page 19, Liège.

Bouache, T., Ginestet, S., Limam, K., Lindner, G., and Bosschaerts, W. (2013). Identification of Thermal Characteristics of a Building. *Energy Procedia*, 42:280–288.

Bouchié, R., Abele, C., Derouineau, S., and Millet, J.-R. (2014). Conception et validation d'un capteur de mesurage de la température extérieure équivalente d'une paroi opaque d'un bâtiment. In *IBPSA France*, Arras.

Brastein, O. M., Lie, B., Sharma, R., and Skeie, N.-O. (2019). Parameter estimation for externally simulated thermal network models. *Energy and Buildings*.

Brastein, O. M., Perera, D., Pfeifer, C., and Skeie, N.-O. (2018). Parameter estimation for grey-box models of building thermal behaviour. *Energy and Buildings*, 169:58–68.

Brynjarsdottir, J. and O'Hagan, A. (2014). Learning about physical parameters: the importance of model discrepancy. *Inverse Problems*, 30(11):114007.

Burnham, K. P. and Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods and Research*, 33(2):261–304.

Chaffar, K. (2012). *Thermographie active appliquee a la caracterisation in situ de parois de batiment*. PhD thesis, Université d'Artois.

Chambers, J. D. (2017). *Developing a rapid, scalable method of thermal characterisation for UK dwellings using smart meter data*. PhD thesis, UCL.

Chong, A. and Menberg, K. (2018). Guidelines for the Bayesian calibration of building energy models. *Energy and Buildings*, 174:527–547.

Crawley, D. B., Lawrie, L. K., Winkelmann, F. C., Buhl, W., Huang, Y., Pedersen, C. O., Strand, R. K., Liesen, R. J., Fisher, D. E., Witte, M. J., and Glazer, J. (2001). EnergyPlus: creating a new-generation building energy simulation program. *Energy and Buildings*, 33(4):319–331.

Dawid, A. P. (1979). Conditional Independence in Statistical Theory. *Journal of the Royal Statistical Society: Series B (Methodological)*, 41(1):1–15.

Day, A. R., Knight, I., Dunn, G., and Gaddas, R. (2004). Improved methods for evaluating base temperature for use in building energy performance lines. *Building Services Engineering Research and Technology*, 24(4):221–228.

De Coninck, R., Magnusson, F., Åkesson, J., and Helsen, L. (2016). Toolbox for development and validation of grey-box building models for forecasting and control. *Journal of Building Performance Simulation*, 9(3):288–303.

Deconinck, A.-H. and Roels, S. (2017). Is stochastic grey-box modelling suited for physical properties estimation of building components from on-site measurements? *Journal of Building Physics*, 40(5):444–471.

Everett, R. (1985). The Rapid Thermal Calibration of Houses. Technical report, Energy Research Group Open University ERG 055.

Farmer, D., Gorse, C., Swan, W., Fitton, R., Brooke-Peat, M., Miles-Shenton, D., and Johnston, D. (2017). Measuring thermal performance in steady-state conditions at each stage of a full fabric retrofit to a solid wall dwelling. *Energy and Buildings*, 156:404–414.

Fedorov, V. (2010). Optimal experimental design. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(5):581–589.

Fels, M. F. (1986). PRISM: An Introduction. *Energy and Buildings*, 9:5–18.

Ferlay, P. (2012). *Mesure du Coefficient Ubat d'un Bâtiment existant*. Editions parisiennes.

Finkelstein, J. M. and Schafer, R. E. (1971). Improved goodness-of-fit tests. *Biometrika*, 58(3):641–645.

Foucquier, A., Brun, A., Faggianelli, G. A., and Suard, F. (2013a). Effect of wall merging on a simplified building energy model: accuracy versus number of equations. *13th International Conference of the International Building Performance Simulation Association*, pages 3161–3168.

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

Fraisse, G., Viardot, C., Lafabrie, O., and Achard, G. (2002). Development of a simplified and accurate building model based on electrical analogy. *Energy and Buildings*, 34(10):1017–1031.

Gaspar, K., Casals, M., and Gangolells, M. (2018). Review of criteria for determining HFM minimum test duration. *Energy and Buildings*, 176:360–370.

Glover, K. and Willems, J. C. (1974). Parametrizations of Linear Dynamical Systems: Canonical Forms and Identifiability. *IEEE Transactions on Automatic Control*, 19(6):640–646.

Godfrey, K. R. (1979). Correlation Methods. *IFAC Proceedings Volumes*, 12(8):527–534.

Godfrey, K. R., Tan, A. H., Barker, H. A., and Chong, B. (2005). A survey of readily accessible perturbation signals for system identification in the frequency domain. *Control Engineering Practice*, 13(11):1391–1402.

Goffart, J. (2013). *Impact de la variabilité des données météorologiques sur une maison basse consommation. Application des analyses de sensibilité pour les entrées temporelles*. PhD thesis, USMB.

Goffart, J., Mara, T., and Wurtz, E. (2017a). Generation of stochastic weather data for uncertainty and sensitivity analysis of a low-energy building. *Journal of Building Physics*, 41(1):41–57.

Goffart, J., Rabouille, M., and Mendes, N. (2017b). Uncertainty and sensitivity analysis applied to hygrothermal simulation of a brick building in a hot and humid climate. *Journal of Building Performance Simulation*, 10(1):37–57.

Goffart, J. and Woloszyn, M. (2018). RBD-FAST : une méthode d'analyse de sensibilité rapide et rigoureuse pour la garantie de performance énergétique. *Conférence Francophone de l'International Building Performance Simulation Association IBPSA,*, pages 1–8.

Gori, V., Biddulph, P., and Elwell, C. A. (2018). A Bayesian dynamic method to estimate the thermophysical properties of building elements in all seasons, orientations and with reduced error. *Energies*, 11(4).

Gori, V. and Elwell, C. A. (2018). Estimation of thermophysical properties from in-situ measurements in all seasons: Quantifying and reducing errors using dynamic grey-box methods. *Energy and Buildings*, 167:290–300.

Gori, V., Marincioni, V., Biddulph, P., and Elwell, C. A. (2017). Inferring the thermal resistance and effective thermal mass distribution of a wall from in situ measurements to characterise heat transfer at both the interior and exterior surfaces. *Energy & Buildings*, 135:398–409.

Grewal, M. and Glover, K. (1976). Identifiability of linear and nonlinear dynamical systems. *IEEE Transactions on Automatic Control*, 21(6):833–837.

Gustafson, P. (2015). *Bayesian Inference for Partially Identified Models*. Chapman and Hall/CRC.

Haario, H., Saksman, E., and Tamminen, J. (2001). An adaptive Metropolis algorithm. *Bernoulli*.

Hadamard, M. (1907). Les problèmes aux limites dans la théorie des équations aux dérivées partielles. *J. Phys. Theor. Appl*, 6(1):202–241.

Hammarsten, S. (1984). Estimation of energy balance for houses. Technical Report Bulletin M84/18, Swedish Institute for Building Research.

Hammarsten, S. (1987). A critical appraisal of energy-signature models. *Applied Energy*, 26(2):97–110.

Hartland, N. (2018). KL-divergence-estimators.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109.

Heiberger, R. M. and Holland, B. (2015). *Statistical Analysis and Data Display*. Springer Texts in Statistics. Springer New York, New York, NY.

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

Herman, J. and Usher, W. (2017). SALib: An open-source Python library for Sensitivity Analysis. *The Journal of Open Source Software*, 2(9):97.

Hobert, J. P. and Casella, G. (1996). The Effect of Improper Priors on Gibbs Sampling in Hierarchical Linear Mixed Models. *Journal of the American Statistical Association*, 91(436):1461–1473.

Huebner, G. M., McMichael, M., Shipworth, D., Shipworth, M., Durand-Daubin, M., and Summerfield, A. (2013). Heating patterns in English homes: Comparing results from a national survey against common model assumptions. *Building and Environment*, 70:298–305.

Iman, R. L. and Conover, W. J. (1982). A distribution-free approach to inducing rank correlation among input variables. *Communications in Statistics - Simulation and Computation*, 11(3):311–334.

Iooss, B. and Lemaître, P. (2015). A review on global sensitivity analysis methods. *Operations Research/ Computer Science Interfaces Series*, 59(around 30):101–122.

ISO 13789 (2017). Thermal performance of buildings - Transmission and ventilation heat transfer coefficients - Calculation method.

ISO 9869-1 (2014). ISO 9869 Thermal insulation — Building elements — In-situ measurement of thermal resistance and thermal transmittance — Part 1: Heat flow meter method.

Jack, R., Loveday, D., Allinson, D., and Lomas, K. (2018). First evidence for the reliability of building co-heating tests. *Building Research and Information*, 46(4):383–401.

Jacquez, J. A. and Greif, P. (1985). Numerical parameter identifiability and estimability: Integrating identifiability, estimability, and optimal sampling design. *Mathematical Biosciences*, 77(1-2):201–227.

Janssen, H. (2013). Monte-Carlo based uncertainty analysis: Sampling efficiency and sampling convergence. *Reliability Engineering and System Safety*, 109:123–132.

Janssens, A. (2016). IEA EBC Annex 58 Reliable building energy performance characterisation based on full scale dynamic measurements : ST b Overview of methods to analyse dynamic data. Technical report, IEA EBC.

Ji, Y., Lee, A., and Swan, W. (2019). Building dynamic thermal model calibration using the Energy House facility at Salford. *Energy and Buildings*.

Jiménez, M. J., Madsen, H., and Andersen, K. K. (2008). Identification of the main thermal characteristics of building components using MATLAB. *Building and Environment*, 43(2):170–180.

Johnston, D., Miles-Shenton, D., Farmer, D., and Wingfield, J. (2013). Whole House Heat Loss Test Method (Coheating). Technical report, Centre for Built Environment.

Josse, R. (2017). *Méthode et outils pour l'identification de défauts des bâtiments connectés performants*. PhD thesis, Université Grenoble Alpes.

Juhl, R., Kristensen, N., Bacher, P., Kloppenborg, J., and Madsen, H. (2013). CTSM-R user guide. Technical report, Technical University of Denmark.

Juricic, S., Bacher, P., Goffart, J., Rouchier, S., and Fraisse, G. (2019). Identifiability of the heat transfer coefficient in buildings with unheated spaces. In *Proceedings of the 16th IBPSA Building Simulation Conference*, Rome.

Juricic, S., Goffart, J., Rouchier, S., Foucquier, A., and Fraisse, G. (2018a). Impact de la variabilité naturelle des conditions météorologiques sur l'estimation des paramètres : application aux modèles RC. In *Conférence Francophone de l'International Building Performance Simulation Association*, pages 1–8, Bordeaux.

Juricic, S., Rouchier, S., Foucquier, A., and Fraisse, G. (2018b). Evaluation of the physical interpretability of calibrated building model parameters. In *7th International Building Physics Conference*, pages 1199–1204, Syracuse (NY).

Kaipio, J. P. and Fox, C. (2011). The Bayesian Framework for Inverse Problems in Heat Transfer. *Heat Transfer Engineering*, 32(9):718–753.

Kaipio, J. P. and Somersalo, E. (2006). *Statistical and Computational Inverse Problems*, volume 160. Springer.

Kaipio, J. P. and Somersalo, E. (2007). Statistical inverse problems: Discretization, model reduction and inverse crimes. *Journal of Computational and Applied Mathematics*, 198(2):493–504.

Karlsson, J., Anguelova, M., and Jirstrand, M. (2012). An Efficient Method for Structural Identifiability Analysis of Large Dynamic Systems*. *IFAC Proceedings Volumes*, 45(16):941–946.

Kennedy, M. C. and O'Hagan, A. (2001). Bayesian Calibration of Computer Models. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 63(3):425–464.

Kramer, R., van Schijndel, J., and Schellen, H. (2012). Simplified thermal and hygric building models: A literature review. *Frontiers of Architectural Research*, 1(4):318–325.

Kreutz, C. (2018). An easy and efficient approach for testing identifiability. *Bioinformatics*, 34(11):1913–1921.

Kreutz, C. (2019). Guidelines for benchmarking of optimization approaches for fitting mathematical models. *arXiv preprint*, pages 1–13.

Kreutz, C., Raue, A., Kaschek, D., and Timmer, J. (2013). Profile likelihood in systems biology. *FEBS Journal*, 280(11):2564–2571.

Kristensen, N. R., Madsen, H., and Jørgensen, S. B. (2004). A method for systematic improvement of stochastic grey-box models. *Computers & Chemical Engineering*, 28(8):1431–1449.

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General*, 142(2):573–603.

Kruschke, J. K. (2018). Rejecting or Accepting Parameter Values in Bayesian Estimation. *Advances in Methods and Practices in Psychological Science*, 1(2):270–280.

Kruschke, J. K., Aguinis, H., and Joo, H. (2012). The Time Has Come: Bayesian Methods for Data Analysis in the Organizational Sciences. *Organizational Research Methods*, 15(4):722–752.

Kumar, R., Carroll, C., Hartikainen, A., and Martin, O. (2019). ArviZ a unified library for exploratory analysis of Bayesian models in Python. *Journal of Open Source Software*, 4(33):1143.

Lee, K., Baek, H. J., and Cho, C. H. (2014). The estimation of base temperature for heating and cooling degree-days for South Korea. *Journal of Applied Meteorology and Climatology*, 53(2):300–309.

Leroy, M. (2010). Note technique n°37 Classification de performance maintenue. Technical report, METEO FRANCE.

Li, H., Hong, T., and Sofos, M. (2019). An Inverse Approach to Solving Zone Air Infiltration Rate and People Count using Indoor Environmental Sensor Data. *Energy and Buildings*.

Li, X. and Wen, J. (2014). Review of building energy modeling for control and operation. *Renewable and Sustainable Energy Reviews*, 37:517–537.

Ligier, S. (2018). *Développement d'une méthodologie pour la garantie de performance énergétique associant la simulation à un protocole de mesure et vérification*. PhD thesis, PSL Research University.

Ljung, L. and Glad, T. (1994). On global identifiability for arbitrary model parametrizations. *Automatica*, 30(2):265–276.

Lucon, O., Ürge-Vorsatz, A., Ahmed Zain, A., Akbari, H., Bertoldi, P., Cabeza, L. F., Eyre, N., Gadgil, A., Harvey, L. D. D., Jiang, Y., Liphoto, E., Mirasgedis, S., Murakami, S., Parikh, J., Pyke, C., and Vilariño, M. V. (2014). Buildings. In *Climate Change 2014: Mitiga- tion of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 671–738. Cambridge University Press, Cambridge.

Macarulla, M., Casals, M., Forcada, N., Gangolells, M., and Giretti, A. (2018). Estimation of a room ventilation air change rate using a stochastic grey-box modelling approach. *Measurement*, 124(April):539–548.

MacKay, D. (2005). *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press.

Madsen, H. (2008). *Time series analysis*. Chapman & Hall / CRC.

Madsen, H., Bacher, P., Bauwens, G., Deconinck, A.-H., Reynders, G., Roels, S., Himpe, E., and Lethé, G. (2015). IEA EBC Annex 58 : Report of Subtask3, part 2. Thermal performance characterisation using time series data - statistical guidelines. Technical report, IEA.

Madsen, H., Bacher, P., Bauwens, G., Deconinck, A.-H., Reynders, G., Roels, S., Himpe, E., and Lethé, G. (2016). Thermal performance characterization using time series data - statistical guidelines. Technical report, IEA EBC Annex 58.

Madsen, H., Bacher, P., and Delff, P. (2010). Mathematical and Statistical Models and Methods for Describing the Thermal Characteristics of Buildings. In *DYNASTEE international workshop : Dynamic Methods for Building Energy Assessment*, Brussels.

Madsen, H. and Holst, J. (1995). Estimation of continuous-time models for the heat dynamics of a building. *Energy and Buildings*, 22:67–79.

Maillet, D., Jarny, Y., and Petit, D. (2011). Problèmes inverses en diffusion thermique Outils spécifiques de conduction inverse et de régularisation. *Techniques de l'ingénieur*, pages 1–2.

Maiwald, T., Timmer, J., Kreutz, C., Klingmüller, U., and Raue, A. (2011). Addressing parameter identifiability by model-based experimentation. *IET Systems Biology*, 5(2):120–130.

Makowski, D., Ben-Shachar, M., and Lüdecke, D. (2019). bayestestR: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework. *Journal of Open Source Software*, 4(40):1541.

Mangematin, E., Pandraud, G., and Roux, D. (2012). Quick measurements of energy efficiency of buildings. *Comptes Rendus Physique*, 13(4):383–390.

Menberg, K., Heo, Y., and Choudhary, R. (2016). Sensitivity analysis methods for building energy models: Comparing computational costs and extractable information. *Energy and Buildings*, 133:433–445.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. Technical report, Division of Technical Information Extension, U.S. Atomic Energy Commission.

Meulemans, J. (2018). An assessment of the QUB / e method for fast in situ measurements of the thermal performance of building fabrics in cold climates To cite this version : HAL Id : hal-01737563 An assessment of the QUB / e method for fast in situ measurements of the thermal. In *Cold Climate HVAC 2018 - The 9th International Cold Climate Conference*, Kiruna, Sweden.

Meulemans, J., Alzetto, F., Farmer, D., and Gorse, C. (2017). QUB/e: A novel transient experimental method for in situ measurements of the thermal performance of building fabrics. *Building Information Modelling, Building Performance, Design and Smart Construction*, pages 115–127.

Mosegaard, K. and Tarantola, A. (2002). 16 Probabilistic approach to inverse problems. In *International Handbook of Earthquake & Engineering Seismology (Part A)*, volume 81, pages 237–265. Academic Press.

Muñoz-Tamayo, R., Puillet, L., Daniel, J. B., Sauvant, D., Martin, O., Taghipoor, M., and Blavy, P. (2018). Review: To be or not to be an identifiable model. Is this a relevant question in animal science modelling? *animal*, 12(4):701–712.

Naveros, I. and Ghiaus, C. (2015). Order selection of thermal models by frequency analysis of measurements for building energy efficiency estimation. *Applied Energy*, 139:230–244.

Nocedal, J. and Wright, S. J. (2006). *Numerical Optimization*. Number 2nd edition in Springer Series in Operations Research and Financial Engineering. Springer New York.

NREL (2016). Energy Plus - Testing and Validation reports on https://energyplus.net/testing.

Perasso, A. (2009). *Identifiabilité de paramètres pour des systèmes décrits par des équations aux dérivées partielles . Application à la dynamique des populations*. PhD thesis, Université Paris-Sud 11.

Pernigotto, G., Prada, A., Gasparella, A., and Hensen, J. L. (2014). Analysis and improvement of the representativeness of EN ISO 15927-4 reference years for building energy simulation. *Journal of Building Performance Simulation*, 7(6):391–410.

Petersen, S., Kristensen, M. H., and Knudsen, M. D. (2019). Prerequisites for reliable sensitivity analysis of a high fidelity building energy model. *Energy and Buildings*, 183:1–16.

Petojević, Z., Gospavić, R., and Todorović, G. (2018). Estimation of thermal impulse response of a multi-layer building wall through in-situ experimental measurements in a dynamic regime with applications. *Applied Energy*, 228(June):468–486.

Plischke, E. (2010). An effective algorithm for computing global sensitivity indices (EASI). *Reliability Engineering and System Safety*, 95(4):354–360.

Pohjanpalo, H. (1978). System identifiability based on the power series expansion of the solution. *Mathematical Biosciences*, 41(1-2):21–33.

Poirier, D. J. (1998). Revising Beliefs in Nonidentified Models. *Econometric Theory*, 14(4):483–509.

Raillon, L. and Ghiaus, C. (2018). An efficient Bayesian experimental calibration of dynamic thermal models. *Energy*, 152:818–833.

Raillon, L., Rouchier, S., and Juricic, S. (2019). pySIP : an open-source tool for Bayesian inference and prediction of heat transfer in buildings. In *Congrès français de thermique*, Nantes.

Rasmussen, C., Frölke, L., Bacher, P., Madsen, H., and Rode, C. (2020). Semi-parametric modelling of sun position dependent solar gain using B-splines in grey-box models. *Solar Energy*, 195(October 2019):249–258.

Rasooli, A. and Itard, L. (2018). In-situ characterization of walls' thermal resistance: An extension to the ISO 9869 standard method. *Energy and Buildings*, 179:374–383.

Raue, A., Karlsson, J., Saccomani, M. P., Jirstrand, M., and Timmer, J. (2014). Comparison of approaches for parameter identifiability analysis of biological systems. *Bioinformatics*, 30(10):1440–1448.

Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmuller, U., and Timmer, J. (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(15):1923–1929.

Raue, A., Kreutz, C., Theis, F. J., and Timmer, J. (2013). Joining forces of Bayesian and frequentist methodology: a study for inference in the presence of non-identifiability. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(1984):20110544.

Reddy, T. A., Deng, S., and Claridge, D. E. (1999). Development of an Inverse Method to Estimate Overall Building and Ventilation Parameters of Large Commercial Buildings. *Journal of Solar Energy Engineering*, 121(1):40–46.

Reiersol, O. (1950). Identifiability of a Linear Relation between Variables Which Are Subject to Error. *Econometrica*, 18(4):375.

Rendu, M., Dréau, J. L., Salagnac, P., Doya, M., and Daâge, M. C. (2019). Comparison of models to identify thermal characteristics of multi-layer building walls using inverse methods. In *Proceedings of the 16th IBPSA Building Simulation Conference*, Rome.

Ritt, J. F. (1950). *Differential Algebra*, volume 33. American Mathematical Soc.

Rodler, A., Guernouti, S., and Musy, M. (2019). Bayesian inference method for in situ thermal conductivity and heat capacity identification: Comparison to ISO standard. *Construction and Building Materials*, 196:574–593.

Rogelj, J., Shindell, D., Jiang, K., Fifita, S., Forster, P., Ginzburg, V., Handa, C., Kheshgi, H., Kobayashi, S., Kriegler, E., Mundaca, L., Séférian, R., and Vilariño, M. (2018). Mitigation Pathways Compatible with 1.5°C in the Context of Sustainable Development. In *Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change,*, page 2. Masson-Delmotte.

Rouchier, S. (2017). Solving inverse problems in building physics : an overview of guidelines for a careful and optimal use of data. *Energy & Buildings*, pages 1–30.

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

Rouchier, S., Rabouille, M., and Oberlé, P. (2018). Calibration of simplified building energy models for parameter estimation and forecasting: Stochastic versus deterministic modelling. *Building and Environment*, 134(February):181–190.

Saccomani, M. P., Audoly, S., D'Angiò, L., Pia Saccomani, M., Audoly, S., and D'Angiò, L. (2003). Parameter identifiability of nonlinear systems: The role of initial conditions. *Automatica*, 39(4):619–632.

Sahu, S. and Gelfand, A. E. (1999). Identifiability, Improper Priors, and Gibbs Sampling for Generalized Linear Models. *Journal of the American Statistical Association*, 94(445):247–254.

Saltelli, A. and Annoni, P. (2010). How to avoid a perfunctory sensitivity analysis. *Environmental Modelling and Software*, 25(12):1508–1517.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications*, 181(2):259–270.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. (2008). *Sensitivity Analysis: From Theory to Practice*. John Wiley & Sons, Ltd, Chichester, UK.

Salvatier, J., Wiecki, T. V., and Fonnesbeck, C. (2016). Probabilistic programming in Python using PyMC3. *PeerJ Computer Science*, 2:e55.

Sandberg, N. H., Sartori, I., Heidrich, O., Dawson, R., Dascalaki, E., Dimitriou, S., Vimm-r, T., Filippidou, F., Stegnar, G., Šijanec Zavrl, M., and Brattebø, H. (2016). Dynamic building stock modelling: Application to 11 European countries to support the energy efficiency and retrofit ambitions of the EU. *Energy and Buildings*, 132:26–38.

Sarkka, S., Alvarez, M. A., and Lawrence, N. D. (2019). Gaussian Process Latent Force Models for Learning and Stochastic Control of Physical Systems. *IEEE Transactions on Automatic Control*, 64(7):2953–2960.

Scales, J. A., Smith, M. L., and Fischer, T. L. (1992). Global optimization methods for multimodal inverse problems. *Journal of Computational Physics*, 103(2):258–268.

Schetelat, P. and Bouchié, R. (2014). ISABELE : a Method for Performance Assessment at Acceptance Stage using Bayesian Calibration. *9th International Conference on System Simulation in Buildings*, 1(1):1–16.

Sedoglavic, A. (2001). *Alexandre Sedoglavic Septembre 2001*. PhD thesis, Ecole Polytechnique.

Sedoglavic, A. (2002). A Probabilistic Algorithm to Test Local Algebraic Observability in Polynomial Time. *Journal of Symbolic Computation*, 33(5):735–755.

Seem, J. (1989). *Modeling of heat transfer in buildings*. PhD thesis, University of Wisconsin-Madison.

Senave, M., Reynders, G., Bacher, P., Roels, S., Verbeke, S., and Saelens, D. (2019). Towards the Characterization of the Heat Loss Coefficient via On-Board Monitoring: Physical Interpretation of ARX Model Coefficients. *Energy and Buildings*.

Senave, M., Reynders, G., Verbeke, S., and Saelens, D. (2017). A simulation exercise to improve building energy performance characterization via on-board monitoring. *Energy Procedia*, 132:969–974.

Senave, M., Roels, S., Reynders, G., Verbeke, S., and Saelens, D. (2020a). 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., and Saelens, D. (2020b). Analysis of the Influence of the Definition of the Interior Dwelling Temperature on the Characterization of the Heat Loss Coefficient via On-Board Monitoring. *Energy and Buildings*, page 109860.

Sengupta, M., Habte, A., Kurtz, S., Dobos, A., Wilbert, S., Lorenz, E., Stoffel, T., Renné, D., Gueymard, C., Myers, D., Wilcox, S., Blanc, P., and Perez, R. (2015). Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications(www.nrel.gov/publications). Technical Report February, National Renewable Energy Laboratory.

Sicard, J., Bacot, P., and Neveu, A. (1985). Analyse modale des échanges thermiques dans le bâtiment. *International Journal of Heat and Mass Transfer*, 28(1):111–123.

Sonderegger, R. (1978). Diagnostic tests determining the thermal response of a house. In *ASHRAE Meeting*, Atlanta.

Stoffel, T. L., Reda, I., Myers, D. R., Renne, D., Wilcox, S., and Treadwell, J. (2000). Current issues in terrestrial solar radiation instrumentation for energy, climate, and space applications. *Metrologia*, 37(5):399–402.

Strachan, P. (1993). Model Validation using the PASSYS Test Cells. *Building and Environment*, 28(2):53–165.

Stuart, A. M. (2013). The Bayesian Approach To Inverse Problems. *arXiv preprint*, pages 1–36.

Subbarao, K. (1988). PSTAR – Primary and Secondary Terms Analysis and Renormalization: A Unified Approach to Building and Energy Simulations and Short-Term Testing; A Summary. Technical report, Solar Energy Research Institude.

Tabares-Velasco, P. C. and Griffith, B. (2012). Diagnostic test cases for verifying surface heat transfer algorithms and boundary conditions in building energy simulation programs. *Journal of Building Performance Simulation*, 5(5):329–346.

Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics.

Tarantola, A. (2006). Popper, Bayes and the inverse problem.

Tarantola, S., Gatelli, D., and Mara, T. (2006). Random balance designs for the estimation of first order global sensitivity indices. *Reliability Engineering and System Safety*, 91(6):717–727.

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

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

Thomassin, M. (2005). *Estimation de retard dans des conditions expérimentales passives. Application à l'identification d'un bief de rivière.* PhD thesis, Université Henri Poincaré, Nancy 1.

Tian, W., Heo, Y., de Wilde, P., Li, Z., Yan, D., Park, C. S., Feng, X., and Augenbroe, G. (2018). A review of uncertainty analysis in building energy assessment. *Renewable and Sustainable Energy Reviews*, 93(May):285–301.

Tissot, J. Y. and Prieur, C. (2012). Bias correction for the estimation of sensitivity indices based on random balance designs. *Reliability Engineering and System Safety*, 107:205–213.

Tuominen, P., Klobut, K., Tolman, A., Adjei, A., and De Best-Waldhober, M. (2012). Energy savings potential in buildings and overcoming market barriers in member states of the European Union. *Energy and Buildings*, 51:48–55.

Ürge-Vorsatz, D., Novikova, A., and Sharmina, M. (2014). Counting good: quantifying the co-benefits of improved efficiency in buildings. In *European Council for an Energy Efficient Economy (ECEEE) Summer Study*, La Colle sur Loup.

Vajda, S. (1983). Structural identifiability of dynamical systems. *International Journal of Systems Science*, 14(11):1229–1247.

Van Doren, J. F., Van den Hof, P. M., Jansen, J. D., and Bosgra, O. H. (2008). Determining Identifiable Parameterizations for Large-scale Physical Models in Reservoir Engineering. *IFAC Proceedings Volumes*, 41(2):11421–11426.

Van Schijndel, A. W. M. (2009). The Exploration of an Inverse Problem Technique to Obtain Material Properties of a Building Construction. In *4th International Building Physics Conference*, pages 91–98, Istanbul.

Vehtari, A., Gelman, A., and Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5):1413–1432.

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., and Bürkner, P.-C. (2019). Rank-normalization, folding, and localization: An improved Rhat for assessing convergence of MCMC. *arXiv*, pages 1–26.

Villaverde, A. F. and Banga, J. R. (2017). Structural properties of dynamic systems biology models: Identifiability, reachability, and initial conditions. *Processes*, 5(2).

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, l., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., and van Mulbregt, P. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3):261–272.

Wakefield, J. (2013). *Bayesian and Frequentist Regression Methods*. Springer Series in Statistics. Springer New York, New York, NY.

Walter, E. and Lecourtier, Y. (1982). Global approaches to identifiability testing for linear and nonlinear state space models. *Mathematics and Computers in Simulation*, 24(6):472–482.

Walter, E. and Pronzato, L. (1994). *Identification de modèles paramétriques à partir de données expérimentales*. Masson.

Walter, E. and Pronzato, L. (1997). *Identification of parametric models from experimental data*. Springer.

Wang, Q., Kulkarni, S. R., and Verdú, S. (2009). Divergence estimation for multidimensional densities via κ-nearest-neighbor distances. *IEEE Transactions on Information Theory*, 55(5):2392–2405.

Wang, S. and Xu, X. (2006a). Parameter estimation of internal thermal mass of building dynamic models using genetic algorithm. *Energy Conversion and Management*, 47(13-14):1927–1941.

Wang, S. and Xu, X. (2006b). Simplified building model for transient thermal performance estimation using GA-based parameter identification. *International Journal of Thermal Sciences*, 45(4):419–432.

Watanabe, S. (2013). A widely applicable bayesian information criterion. *Journal of Machine Learning Research*, 14(1):867–897.

Wei, S., Jones, R., and de Wilde, P. (2014). Driving factors for occupant-controlled space heating in residential buildings. *Energy and Buildings*, 70:36–44.

Wit, E., van den Heuvel, E., and Romeijn, J.-W. (2012). 'All models are wrong...': an introduction to model uncertainty. *Statistica Neerlandica*, 66(3):217–236.

Wu, H., Zhu, H., Miao, H., and Perelson, A. S. (2008). Parameter identifiability and estimation of HIV/AIDS dynamic models. *Bulletin of Mathematical Biology*, 70(3):785–799.

Xia, X. and Moog, C. (2003). Identifiability of nonlinear systems with application to HIV/AIDS models. *IEEE Transactions on Automatic Control*, 48(2):330–336.

Xie, Y. and Carlin, B. P. (2006). Measures of Bayesian learning and identifiability in hierarchical models. *Journal of Statistical Planning and Inference*, 136(10):3458–3477.

Yang, Y., Wu, T. V., Sempey, A., Dumoulin, J., and Batsale, J.-c. (2019). Short time non-destructive evaluation of thermal performances of building walls by studying transient heat transfer. *Energy and Buildings*, 184:141–151.

Zayane, C. (2011). *Identification d'un modèle de comportement thermique de bâtiment à partir de sa courbe de charge*. PhD thesis, ParisTech Ecole nationale supérieure des mines de Paris.

List of Figures

List of Tables

A

Synthèse en français

Estimer la performance thermique d'une enveloppe de bâtiment à partir de mesures in-situ non intrusives présente l'intérêt d'être représentatif des conditions réelles de fonctionnement du bâtiment mais induit une qualité des données mesurées faible. Cela peut biaiser le résultat voire rendre impossible son estimation. Cette thèse s'attache à étudier la faisabilité de la caractérisation thermique d'une enveloppe à partir de données peu informatives.

Si l'identifiabilité théorique et pratique peut être aisément vérifiée en approches fréquentiste ou Bayésienne, elle ne constitue pas une condition suffisante à l'interprétation physique des paramètres estimés.

Une méthodologie originale est donc proposée pour évaluer l'interprétation physique de modèles RC stochastiques. Un modèle de référence délivre des simulations thermiques qui servent à l'apprentissage du modèle étudié. L'originalité est de pouvoir quantifier l'influence des incertitudes de l'environnement du bâtiment sur la caractérisation par une analyse de sensibilité globale. La méthodologie est appliquée d'une part à l'étude de l'influence de la variabilité des conditions météorologiques et d'autre part à l'interprétabilité physique individuelle de paramètres estimés. Les résultats montrent que 11 jours suffisent à obtenir une bonne répétabilité et précision de l'estimation de la résistance thermique. Aller plus loin en identifiant les pertes thermiques par la ventilation ou par les parois donnant sur un local non chauffé n'est en revanche pas possible dans ces conditions.

Contexte

Pour avoir une chance de contenir le réchauffement climatique sous la barre de 1.5 $^{\circ}C$ au dessus du niveau pré-industriel, il est devenu nécessaire et urgent d'agir à un niveau mondial et systémique en s'engageant a minima sur le respect des Accords de Paris. Un levier d'action est de s'engager sur la décarbonation du secteur du bâtiment, qui présente un potentiel de réduction des émissions de gaz à effet de serre conséquent. L'effort peut judicieusement porter sur la rénovation du parc de bâtiments existants (Lucon et al., 2014). Si l'on considère qu'une rénovation lourde d'un bâtiment donné ne se fera qu'une unique fois à l'horizon 2050, il est alors nécessaire de conduire ces travaux de manière très ciblée et ambitieuse. Quantifier avec précision la performance des systèmes et de l'enveloppe existante permettra donc de cibler la stratégie de rénovation, quantifier ses gains et donc rendre les investissements plus acceptables.

Déterminer les pertes thermiques par l'enveloppe seule peut se faire avantageusement par le biais de mesures in situ en environnement contrôlé. Par exemple, le contrôle de la quantité de chauffage dans le bâtiment permet, avec la mesure des conditions météorologiques et de l'ambiance intérieure, de déterminer les caractéristiques thermiques de l'enveloppe de façon relativement précise, comme par exemple par un test de co-heating.

Des mesures en conditions contrôlées sont cependant incompatibles avec l'occupation de bâtiments. Pourtant, s'il s'agit de proposer des scenarii pertinents de rénovation de locaux continuellement occupés comme des hôpitaux, des établissements d'hébergement, ou presque continuellement occupés comme dans de nombreux locaux de bâtiments à usage tertiaire, il devient nécessaire de développer une méthode fiable, précise, globale et si possible rapide de diagnostic thermique de l'enveloppe d'un bâtiment.

Deux obstacles se dressent à un diagnostic fait de manière que l'on pourrait qualifier de non intrusive:

- si le chauffage n'est pas contrôlé pour permettre une sollicitation optimisée pour le diagnostic, les données mesurées risquent d'être très peu informatives, peu nombreuses voire biaisées et compromettent l'analyse qui peut en être faite;
- l'occupation même des locaux est une source de perturbation des mesures. L'influence des comportements est non seulement très difficile à mesurer et quantifier, mais est en plus difficilement prévisible.

Cette thèse se propose d'évaluer en quoi le caractère peu informatif de données récoltées affecte la fiabilité d'une caractérisation thermique de l'enveloppe en conditions de mesures non intrusives. La synthèse qui suit reprend en substance les contenus des cinq chapitres développés dans le cadre de ce travail de thèse.

La section A.1 situe la proposition de cette thèse au regard de la littérature existante et en rappelant les principes de résolution du problème inverse ainsi constitué.

La section A.2 explicite en quoi il est nécessaire de s'assurer de l'existence et unicité de la solution au problème inverse, c'est-à-dire de vérifier l'identifiabilité des paramètres recherchés. Un ensemble de bonnes pratiques d'estimation de paramètres est alors développé bien que l'interprétation physique de paramètres estimés, même de manière unique reste incertaine.

Pour cette raison, une méthodologie est proposée pour évaluer sur un banc de test numérique l'interprétabilité physique de modèles stochastiques RC estimés à partir de données peu informatives, synthétiquement générées en conditions non intrusives. Pour quantifier le caractère physiquement interprétable d'un paramètre estimé à sa valeur théorique, un indicateur ad hoc appelé indicateur d'interprétabilité physique a été développé.

La méthodologie est appliquée dans un premier lieu pour évaluer l'influence de la variabilité naturelle des conditions météorologiques sur l'estimation d'une résistance thermique globale de l'enveloppe. La section A.3 en trace les principaux résultats.

La deuxième application, dont les résultats sont décrits en section A.4, consiste à évaluer la capacité de modèles RC à déterminer spécifiquement les pertes thermiques par ventilation d'une part, et déterminer la résistance thermique d'un plancher bas sur vide sanitaire.

Les travaux montrent ainsi la pertinence et les limites des modèles dynamiques RC pour exploiter des données faiblement informatives dans des conditions de mesures non intrusives dans le but d'estimer la performance thermique d'une enveloppe de bâtiment.

A.1 Caractérisation thermique d'une enveloppe à partir de données: un problème inverse

Cette thèse s'intéresse spécifiquement à la caractérisation thermique de l'enveloppe du bâtiment globale, c'est-à-dire à l'échelle du bâtiment. A cette échelle, la caractérisation thermique consiste alors en l'estimation à partir de données de mesures in situ de notamment: le coefficient de transferts thermiques global communément appelé *HLC*, de la résistance thermique globale, des capacités thermiques ou des temps caractéristiques, voire de l'estimation des pertes thermiques par ventilation ou par infiltration.

Cette partie cartographie brièvement les méthodes décrites dans la littérature pour estimer la performance thermique d'une enveloppe et développe une proposition d'exploitation de données peu informatives par modèles RC stochastiques, proposition adaptée à ce genre de problèmes inverses.

A.1.1 Cartographie des méthodes existantes de caractérisation globale

Quelles que soient les données et les options de résolution utilisées, les "méthodes" dans la littérature varient sur trois critères : l'acquisition des données, le type de modèle utilisé pour décrire les dynamiques thermiques du système et les outils numériques pour résoudre le problème inverse et ainsi effectuer l'estimation en elle-même.

A l'échelle du bâtiment entier, deux approches s'opposent quant à l'exploitation des données: considérer ou non que les échanges thermiques sont dans un état d'équilibre. Cela revient à considérer ou non que le stockage et déstockage thermique sont significatifs dans l'équilibre thermique du bâtiment.

S'il est évident que les échanges thermiques ne sont jamais, en réalité, à l'état d'équilibre, il est possible de considérer cette hypothèse comme raisonnable moyennant quelques conditions. D'abord, les températures de l'ambiance intérieure doivent être les plus stables possibles. C'est la raison pour laquelle les tests de co-heating, certes intrusifs (Jack et al., 2018), sont effectués à température constante. Ensuite, pour diminuer la variabilité due aux variations des conditions météorologiques ou due à une éventuelle petite variation des températures à l'intérieur, la granularité temporelle des données doit être au plus bas, le plus souvent moyennée journalièrement. A partir de données moyennées journalièrement, la caractérisation thermique de l'enveloppe se fait par régression linéaire ordinaire. Ainsi, en conditions non intrusives, exploiter des mesures avec l'hypothèse d'état d'équilibre nécessite des semaines, voire des mois de mesures pour espérer un résultat précis et avec une incertitude raisonnable.

Si, au contraire, le système n'est pas considéré comme à l'équilibre, l'exploitation peut se faire à partir de pas de temps de mesures plus court et le système est modélisé avec des modèles dynamiques. Dans la littérature traitant de mesures non intrusives, ces modèles sont par exemple des modèles autorégressifs sur quelques semaines de mesures comme dans Senave et al. (2020a) à l'échelle du bâtiment entier, ou des modèles RC à l'échelle de la paroi sur moins de 10 jours

de mesures comme dans Deconinck and Roels (2017) ou Gori and Elwell (2018). En conditions intrusives, où soit la température soit le chauffage est contrôlé, Thébault and Bouchié (2018) et Bacher and Madsen (2011) utilisent également, sur seulement quelques jours de mesures, des modèles RC pour effectuer une caractérisation thermique à l'échelle de l'enveloppe entière.

A l'échelle de l'enveloppe, il semble donc que les méthodes basées sur des modèles dynamiques permettent une caractérisation plus rapide que celles basées sur une hypothèse d'équilibre thermique de l'enveloppe. Pour autant, en conditions non intrusives, très peu de méthodes ont été étudiées. Si des mesures plus courtes sont une perspective encourageante pour une application future à grande échelle, la plus faible qualité des mesures collectées peut être un frein à la précision acquise.

A.1.2 Résolution d'un problème inverse : choix d'un modèle et d'outils numériques

Pour comprendre en quoi une faible qualité des mesures est un frein à la précision de la caractérisation thermique d'une enveloppe de bâtiment, cette section détaille les étapes de la résolution d'un problème inverse. Cette section dégage ensuite une proposition d'exploitation de données peu informatives par modèles dits RC stochastiques.

A.1.2.1 Un problème inverse mal posé

De manière très générale, un problème inverse consiste à déterminer un ou plusieurs paramètres d'intérêt du système étudié, ici l'enveloppe du bâtiment, à partir de mesures dans l'idéal parfaitement représentatives du comportement du système. Comme l'expression analytique liant les paramètres d'intérêt directement aux mesures n'est qu'exceptionnellement faisable, effectuer une estimation signifie déterminer numériquement une approximation de la solution du problème inverse ainsi posé.

En observant, i.e. mesurant, les conditions aux limites et éventuellement les états intrinsèques du système, la résolution suppose de se servir d'un modèle approprié, qui contient les paramètres d'intérêt et qui décrit le comportement du système, pour essayer d'expliquer les mesures précédemment acquises. Concrètement, par itérations successives, un algorithme adéquat converge, en principe, vers les paramètres les plus problables et ainsi, idéalement, vers les valeurs réelles des paramètres d'intérêt.

La difficulté vient du fait que les mesures idéales n'existent pas, de surcroît si les mesures envisagées se font en conditions non intrusives. Ces mesures ne portent alors que sur les conditions aux limites et pas sur les états du système. Elles sont aussi non exhaustives, peut-être même biaisées, discrètes et bruitées. Dans tous les cas, elles ne peuvent pas être représentatives de l'ensemble des flux thermiques significatifs à l'échelle d'un bâtiment.

En conditions non intrusives, les conditions intérieures sont certes contrôlées pour être compatibles avec une occupation, mais pas optimisées pour la caractérisation thermique. Ainsi, ces conditions rendent le problème inverse d'autant moins bien posé que les données sont peu informatives. Dès lors, en conditions non intrusives, les données collectées sont limitées si bien que la faisabilité et la robustesse d'une caractérisation thermique sont mises en doute.

Pour autant, la revue de littérature a montré que des résultats satisfaisants sont obtenus à l'échelle de la paroi avec des modèles dynamiques, même en non intrusifs et sur des durées de l'ordre de la semaine au plus. Pour exploiter des données peu informatives, le choix du modèle est en effet déterminant.

Ce choix est conditionné par à la fois la capacité du modèle à décrire le comportement du système avec des paramètres interprétables physiquement, et à la fois à être robuste quand les données sont faiblement informatives. Pour la première raison, les modèles purement statistiques dits *noirs* seront écartés et pour la deuxième raison, parmi les modèles physiques, dits *blancs*, les modèles très finement détaillés seront également écartés vu le risque de surapprentissage.

Ainsi, les modèles physiques *blancs* mais dont la description des comportements physiques est simplifiée semblerait être bien adaptés à la résolution du problème inverse précédemment exposé. Notamment, pour conserver une description dynamique des transferts thermiques, la proposition étudiée dans ce travail est d'utiliser des modèles dits RC. Ils sont basé sur des transferts thermiques linéarisés et agrégés. In fine, les transferts thermiques, même à l'échelle du bâtiment entier, sont simplement décrit par des résistances et capacités thermiques dans des équations différentielles du 1er ordre. Pour pallier l'erreur de modèle commise en simplifiant la physique, très probablement importante dans le cas de mesures peu informatives (Macarulla et al., 2018), il est proposé de considérer les modèles RC stochastiques (Bohlin and Graebe, 1995; Madsen et al., 2010). Notamment, comme montré dans l'équation A.1, la formulation choisie dans ce travail consiste à ajouter un terme stochastique au système d'état pour prendre en compte la déviation possible du modèle par rapport aux observations. Une matrice de covariance permet de moduler l'importance de cette déviation.

$$\begin{bmatrix} \dot{T_w} \\ \dot{T_{int}} \end{bmatrix} = \begin{bmatrix} -\frac{1}{C_w}(\frac{1}{R_w}+\frac{1}{R_b}) & \frac{1}{C_w R_b} \\ \frac{1}{C_b R_b} & -\frac{1}{C_b}(\frac{1}{R_b}+\frac{1}{R_g}) \end{bmatrix} \begin{bmatrix} T_w \\ T_{int} \end{bmatrix} + \begin{bmatrix} \frac{1}{C_w R_w} & 0 & 0 \\ \frac{1}{C_b R_g} & \frac{gA}{C_b} & \frac{1}{C_b} \end{bmatrix} \begin{bmatrix} T_{ext} \\ I_{glo,Hor} \\ Q_h \end{bmatrix} + \sigma d\omega \quad \text{(A.1)}$$

Il existe une multitude de modèles RC, fonction du nombre de paramètres pris en compte. Le choix du modèle se fait donc en fonction du bâtiment modélisé et de notamment son inertie thermique qui déterminera l'ordre du modèle RC (i.e. le nombre de capacités thermiques modélisées).

L'interprétation physique des modèles RC stochastiques au travers de la résistance globale équivalente a été plusieurs fois établie dans la littérature, par exemple dans Madsen et al. (2016); Thébault and Bouchié (2018). Cependant, l'interprétation physique des résistances et des capacités séparément n'est pas acquise et nécessitera une étude ad hoc.

L'estimation des paramètres d'intérêt, c'est-à-dire l'approximation numérique à partir des données disponibles, se fait par calibration du modèle choisi. Pour reprendre les termes de Tarantola (2005), déterminer *la* valeur du paramètre d'intérêt vu les données n'a pas de sens. La calibration de modèle doit plutôt permettre d'évaluer l'ensemble des valeurs plausibles qui expliquent les données récoltées. Cet ensemble de valeurs plausibles traduit donc l'état d'information que l'on a sur le paramètre d'intérêt. L'objectif de la calibration est donc de déterminer sous la forme d'un intervalle d'incertitude (ou dans un sens très général la distribution de probabilité) l'étendue de l'information déduite de connaissances expertes et des mesures récoltées.

On considère deux approches à l'estimation de paramètres : l'approche fréquentiste et l'approche bayésienne. La différence principale entre les deux est de considérer connue ou non la forme de la distribution de probabilité. L'approche fréquentiste prend en effet comme hypothèse forte que la distribution de probabilité de chaque paramètre estimé est gaussienne. Cela a pour énorme avantage de faciliter donc accélérer l'approximation numérique puisqu'il suffit de trouver *la* valeur de plus grande probabilité et d'évaluer son incertitude sous hypothèse de normalité. L'approche bayésienne consiste à exprimer la distribution de probabilité comme une fonction de connaissances a priori et de connaissances acquises par les observations grâce à la relation de Bayes. Généralement, l'approche bayésienne ne considère pas la distribution de probabilité comme connue, bien qu'il soit possible d'utiliser la relation de Bayes avec l'hypothèse gaussienne. Généralement donc, l'estimation des paramètres d'intérêt se fait par échantillonnage de l'espace des paramètres pour trouver les valeurs de plus grande probabilité. Ainsi, cette approche est plus coûteuse numériquement, bien que plus réaliste en terme d'évaluation de l'incertitude.

Les travaux de cette thèse feront appel à l'une et l'autre des approches et utiliseront à cette fin la librairie python pySIP (Raillon et al., 2019).

A.1.2.2 Conclusion

La faisabilité de la caractérisation thermique d'une enveloppe de bâtiment par des mesures en conditions non intrusives n'est pas acquise du fait que les données récoltées sont possiblement faiblement informatives. Du fait de données non exhaustives, discrètes voire non représentatives et donc biaisées, l'estimation des paramètres d'intérêt comme la résistance thermique globale de l'enveloppe est a priori compromise.

Cette thèse se propose d'exploiter les données ainsi récoltées par des modèles RC stochastiques, ce qui a pour avantage d'exploiter l'ensemble des dynamiques présentes dans les données en minimisant le risque de surapprentissage donc en maximisant l'interprétabilité physique des estimations. L'objectif de ce travail sera alors d'évaluer si le gain en information à partir de données peu informatives de modèles RC est suffisant pour établir une robustesse de l'estimation de la résistance thermique globale de l'enveloppe.

A.2 Identifiabilité et interprétabilité

Il a été établi que le caractère peu informatif des observations récoltées est une menace à l'estimation, c'est-à-dire une menace :

- à l'existence (les incertitudes sont finies);
- à l'unicité (le set de valeurs admissibles est unique) et;
- à l'interprétabilité physique des valeurs estimées.

Cette section détaille dans un premier temps ce qui conditionne l'identifiabilité, i.e. existence et unicité de l'estimation puis, dans un deuxième temps, explicite les raisons pour lesquelles existence et unicité n'impliquent pas interprétabilité physique. La section présentera enfin un ensemble de bonnes pratiques pour la calibration sous forme d'un ensemble des conditions nécessaires, mais pas suffisantes, à l'interprétabilité.

A.2.1 Identifiabilité structurelle et identifiabilité pratique

L'existence et l'unicité d'une solution à un problème inverse, c'est-à-dire l'identifiabilité de la solution, tient à deux caractéristiques : son identifiabilité structurelle et son identifiabilité pratique.

Le caractère structurel tient à la formulation mathématique du modèle. Si par son expression, plusieurs jeux de valeurs de paramètres produisent la même prédiction de modèle à conditions aux limites identiques, il sera impossible de déduire une solution unique au problème. Dans les faits, dans ce cas des jeux distincts de paramètres produisent la même vraisemblance, et aucun algorithme ne pourra faire la discrimination entre chacun d'eux. Il est alors hautement probable et souhaitable que l'algorithme pour la calibration du modèle échoue et renvoie à une erreur. Il convient donc de vérifier avant tout que le modèle utilisé soit structurellement identifiable.

Dans le cas des modèles d'état RC, la vérification de l'identifiabilité structurelle peut se faire de diverses manières. On retiendra les méthodes implémentées dans des outils prêt à l'emploi, notamment dans Sedoglavic (2001) et Bellu et al. (2007), respectivement sous Maple et Reduce, ce qui facilite leur usage. En l'occurence, une banque de modèles RC structurellement identifiables a donc été établie pour les besoins des travaux présentés ici.

Les modèles retenus sont donnés dans le tableau A.1. Les ordres des modèles font référence au nombre de capacités thermiques de chacun des modèles.

	1^{er} ordre	2^{e} ordre				3^{e} ordre	
	T_i R A_i	T_wT_i R_oR_i A_wA_i	T_wT_i R_oR_i c_v A_wA_i	T_wT_i R_oR_i R_p A_wA_i	T_wT_i R_oR_i R_b A_wA_i	$T_{we}T_wT_i$ $R_oR_wR_i$ $A_eA_wA_i$	T_wT_i T_b R_wR_i $R_{bi}R_{bb}$ A_wA_i
Nom du modèle	2.3	2.7	2.8	2.9	2.10	2.11	2.12
Entrée(s) mesurée(s)	T_{ext} $\Phi_{heating}$ I_{sol}^{global}		T_{ext} $\Phi_{heating}$ I_{sol}^{global} Φ_{vent}	T_{ext} $\Phi_{heating}$ I_{sol}^{global}	T_{ext} $\Phi_{heating}$ I_{sol}^{global} $T_{crawl,space}$	T_{ext} $\Phi_{heating}$ I_{sol}^{global}	T_{ext} $\Phi_{heating}$ I_{sol}^{global} $T_{crawl,space}$
Sortie mesurée	T_{int}						

Table A.1 – Une banque de modèles RC structurellement identifiables

Le caractère structurellement identifiable est une condition nécessaire à l'identifiabilité finale du problème, mais non suffisante. En effet, cette propriété suppose que les données récoltées sont idéalement informatives pour induire l'unicité d'une solution au problème inverse. En conditions non intrusives, il a été au contraire établi que les données sont peu informatives. Or des données mal adaptées ou en trop faible quantité sont un frein à l'estimation. Il s'agit alors ici du caractère identifiable en pratique du problème, autrement appelé identifiabilité pratique (Raue et al., 2013).

Le caractère identifiable ou non d'un problème inverse est relatif à la combinaison du modèle et des données récoltées pour la résolution par ce modèle. Pour exploiter des mesures in situ avec des modèles RC, l'identifiabilité pratique est conditionnée par l'excitation suffisante du système par les conditions aux limites et la mesure adéquate de celles-ci. L'identifiabilité pratique peut être quantifiée a posteriori, c'est-à-dire après calibration du modèle, en évaluant la sensibilité de la vraisemblance à la variation des valeurs des paramètres, sachant la mesure. Si un paramètre à estimer n'a pas d'influence sur la vraisemblance vu les conditions aux limites auxquelles est soumis le système, il n'est pas estimable, donc pas identifiable.

Concrètement, pour évaluer a posteriori l'identifiabilité pratique d'un paramètre θ_i, on détermine son profil de vraisemblance $PL_j(\theta)$. Il s'agit de calculer, pour θ_j fixé, la meilleure vraisemblance possible lorsque les autres paramètres sont laissés libres, comme suggéré dans l'équation A.2. Autrement dit, le modèle est calibré pour chaque valeur de θ_j fixé et la plus grande vraisemblance ainsi atteinte est retenue. Les couples (θ_j, PL_j) doivent indiquer un minima global unique pour assurer l'identifiabilité pratique. Si le maximum de vraisemblance est atteint pour plusieurs valeurs de θ_j, ce paramètre n'est pas identifiable en pratique.

$$PL_j(p) = max_{\theta|\theta_j=p} log\mathscr{L}(y|\theta) \tag{A.2}$$

En conclusion, le processus d'estimation des propriétés thermiques d'intérêt doit être précédé d'une vérification de l'identifiabilité structurelle du modèle utilisé pour estimer ces paramètres. Puis, une vérification postérieure à l'estimation permet de s'assurer de l'identifiabilité pratique du paramètre estimé. Pour autant, unicité de la solution n'est pas synonyme de parfaite interprétabilité. Les freins à l'interprétabilité de l'estimation d'un paramètre sont donc exposés dans la prochaine section.

A.2.2 Insuffisance du principe d'identifiabilité et bonnes pratiques en calibration

La calibration d'un modèle sans erreur de caractérisation et à partir de données récoltées parfaites, comme décrit dans Walter and Pronzato (1997), induit que l'identifiabilité d'un paramètre du modèle signifie identité entre sa valeur du paramètre et la valeur de la propriété physique qu'il représente.

Dans le cas de la caractérisation thermique d'une enveloppe en conditions non intrusives, ces conditions ne sont pas réunies. Les modèles utilisés, aussi détaillés soient-ils, comportent toujours

une erreur de modèle. Dans le cas des modèles RC à l'étude dans ce travail, les paramètres agrègent par linéarisation plusieurs échanges thermiques au travers de différents éléments constructifs. De surcroît, les mesures récoltées sont imparfaites, non exhaustives, peut-être non représentatives voire biaisées et dans tous les cas discrètes. Ces éléments sont autant de sources de corruption du jeu de données parfait que l'on voudrait au contraire pouvoir récolter (Gustafson, 2015).

Devant l'impossibilité d'utiliser un modèle parfait et des données parfaites, le cheminement est le suivant :

- réduire au mieux les incertitudes liées à la récolte des données, quand leur impact est connu, et;
- réduire le risque de non interprétabilité en adoptant un ensemble de bonnes pratiques pour l'estimation de paramètres.

Le premier levier est donc de maximiser l'informativité des données même en conditions non intrusives en déployant un nombre minimal suffisant mais nécessaire de capteurs et ce à un pas de temps adapté au calage de modèles RC. En l'occurrence, la proposition évaluée dans cette thèse est basée sur la mesure a minima des températures intérieures et extérieure, de la puissance de chauffage délivrée à chaque pas de temps et de l'irradiation solaire globale. Il est entendu que ces conditions aux limites ne sont représentatives que d'une partie des échanges thermiques. Une part moins importante mais parfois significative des échanges thermiques est aussi influencée par la vitesse et l'orientation du vent in situ ou les masques solaires. De même, une mesure de l'irradiation solaire globale est moins informative qu'une mesure des irradiations diffuses et directes par paroi. Ainsi, il est primordial de s'assurer, a posteriori, que les prédictions du modèles RC calé se confondent à la mesure à une erreur stochastique gaussienne près. Si ce n'est pas le cas, cela signifie probablement qu'une condition aux limites ou un phénomène est ignoré dans le modèle ce qui nécessite la correction du modèle ou des mesures. Cette étape de vérification sera détaillée en fin de section dans la présentation des bonnes pratiques d'estimation de paramètres.

Ensuite, la granularité temporelle, du fait du choix d'un pas de temps pour les mesures, a aussi un impact sur le biais du fait du risque de repliement de spectre (*aliasing* en anglais) (Madsen et al., 2015). Le théorème de Shannon indique en effet qu'il est nécessaire que la fréquence d'échantillonnage temporel soit le double de la plus haute fréquence parmi les phénomènes observés. Si cette règle n'est pas respectée, les fréquences les plus élevées sont illisibles dans les données mesurées, d'où le repliement de spectre.

La figure A.1 présente l'effet de repliement de spectre quand le pas de temps de mesure varie de 8 à 120 minutes. Les distributions postérieures, déterminées en approche bayésienne par échantillonnage, peuvent être comparées aux distributions a priori. Ces distributions postérieures montrent des profils très similaires quel que soit le pas de temps pour le paramètre R_o par exemple, ou très différents pour notamment le paramètre C_w. Cela signifie que l'estimation de R_o se fait de manière similaire quelque soit le pas de temps, quoiqu'un pas de temps plus faible semble être plus informatif vu la plus faible dispersion de la distribution. Pour autant, les modes sont quasiment identiques. Les modes de chaque distribution de C_w au contraire sont très différents. Les plus grand pas de temps sont non informatifs car les distributions postérieures sont très proches des distributions a priori, tandis que des pas de temps plus courts semblent donner des distributions convergentes.

L'exemple montré en Figure A.1 montre donc en quoi la granularité temporelle peut influer sur l'apprentissage ou non des paramètres d'un modèle RC à partir de données non idéales. Vu les constantes de temps caractéristiques habituelles d'un bâtiment (Sicard et al., 1985), il semble cohérent d'observer que des pas de temps inférieurs à 20 minutes soient plus adaptés à un apprentissage satisfaisant par tous les paramètres d'un modèle RC.

Le deuxième levier pour réduire le risque de non interprétabilité est l'adoption de bonnes

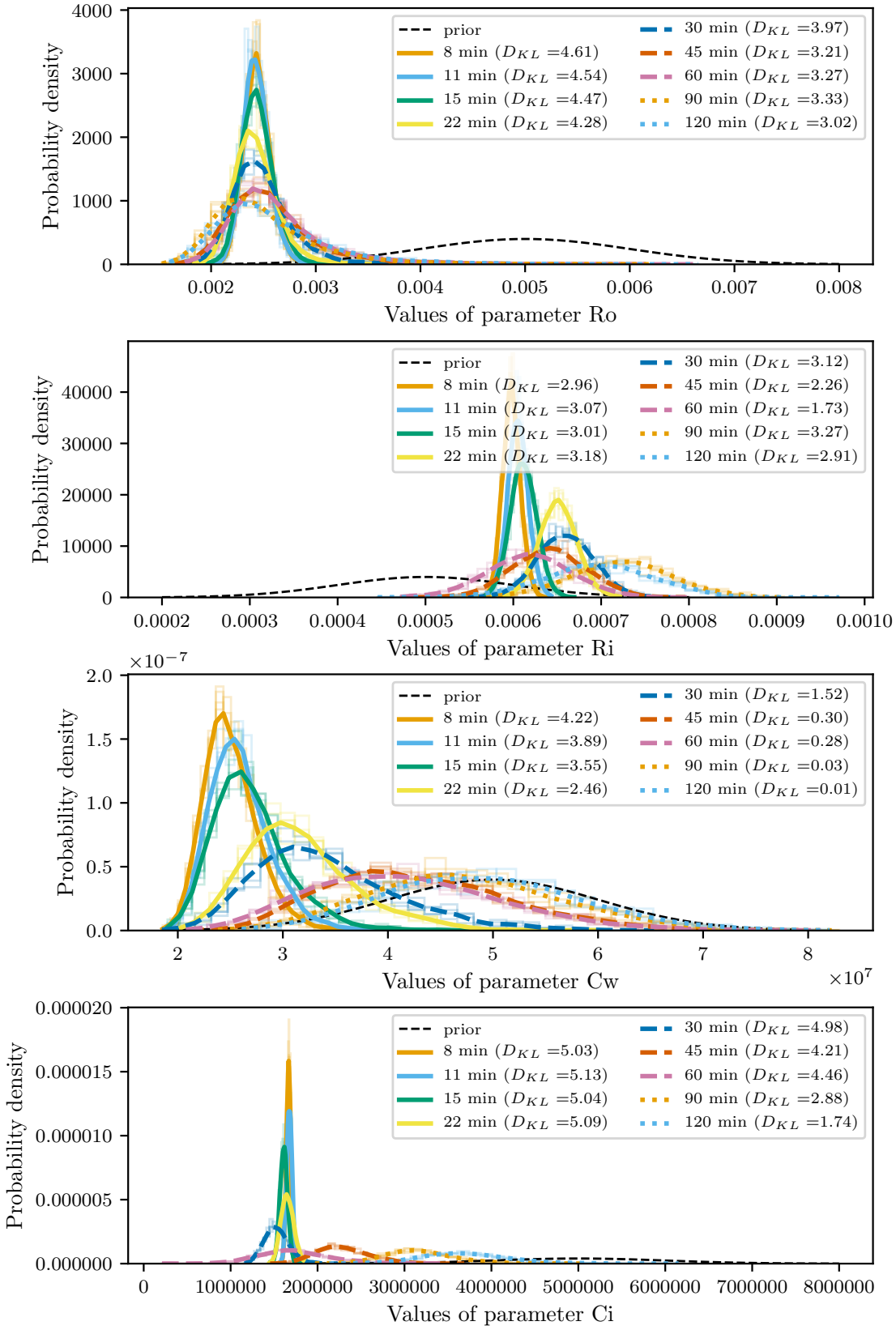


Figure A.1 – Distributions postérieures des paramètres R_o, R_i, C_w et C_i du modèle T_wT_i R_oR_i A_w, calés à partir de 48 heures de mesures ré-échantillonnées à différentes granularités temporelles. Les divergences de Kullback-Leibler D_{KL} mesurent la différence entre distribution a priori et a posteriori et quantifient donc l'apprentissage dû aux seules mesures récoltées.

pratiques au moment de l'exploitation de données récoltées. L'ensemble de ces bonnes pratiques permet à la fois d'effectuer les vérifications basiques d'identifiabilité vues dans la section précédente, mais aussi de détecter les cas d'erreur de modèle, ou de conditions aux limites significatives mais mal ou pas prises en compte. La Figure A.2 présente les étapes à suivre lors de l'exploitation de données, dont on supposera la granularité temporelle satisfaisante et le biais inexistant.

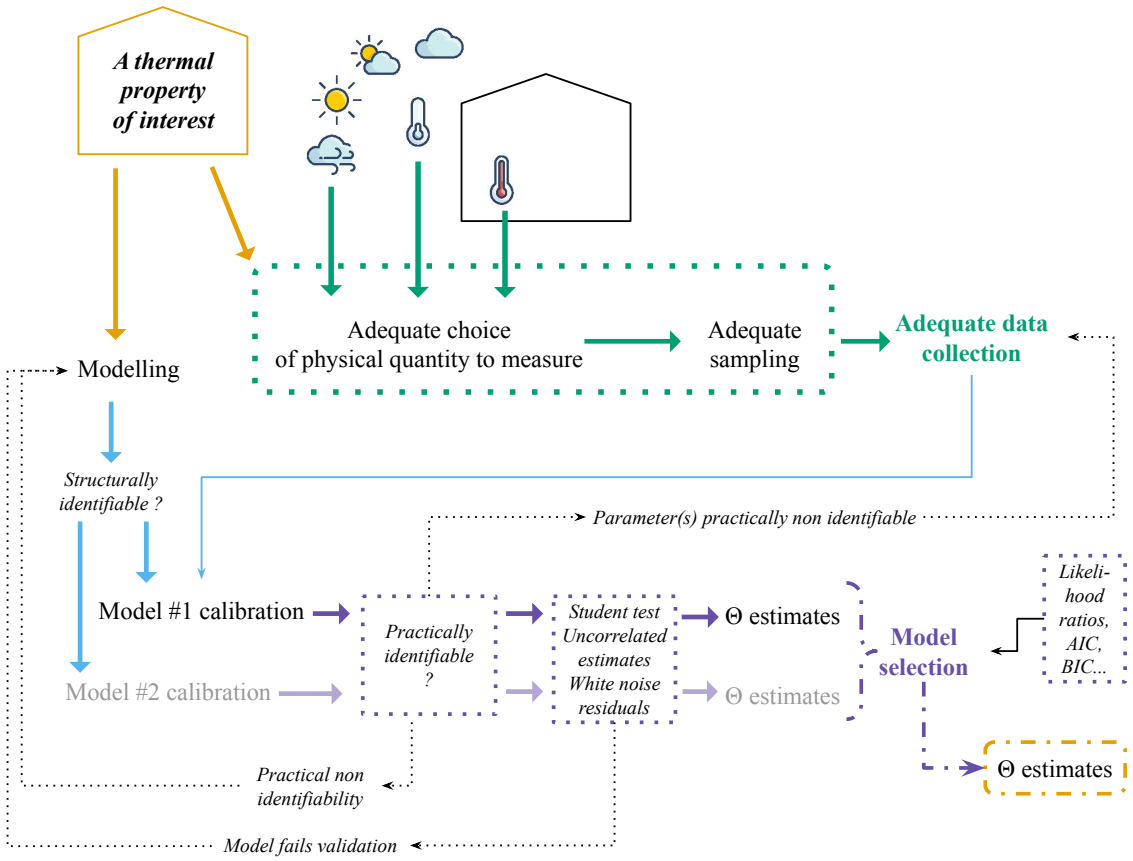


Figure A.2 – Bonnes pratiques pour exploiter des données mesurées: les étapes décrites ne constituent pas une conditions suffisante pour garantir des estimations sans erreur, mais sont en revanche nécessaires.

La procédure présentée vise à éviter le surapprentissage d'un modèle trop complexe, ce qui aurait pour effet de perdre le caractère boîte blanche des modèles utilisés et donc de perdre la signification physique des paramètres estimés. Pour cela, la procédure consiste à effectuer une sélection de modèle. Celle-ci démarre au modèle le plus simple (avec le moins de paramètres) et augmente en complexité à mesure que la vraisemblance s'améliore, jusqu'à ce que le gain en vraisemblance ne soit plus significatif, relativement au nombre de paramètres du modèle sélectionné. La sélection en question se fait donc sur la base d'un critère quantitatif, comme le ratio de vraisemblance (Bacher and Madsen, 2011) ou les critères de type Akaike (AIC, BIC (Madsen et al., 2015)).

La sélection de modèle repose donc sur le calage de différents modèles. A chaque calage, plusieurs étapes doivent être vérifiées: le modèle calé est structurellement identifiable, les paramètres estimés sont identifiables en pratique, ils sont significativement différents de 0, les paramètres estimés ne sont pas corrélés et enfin les résidus (différence normée entre la prédiction du modèle calé et les mesures) ont un caractère aléatoire normal. Ces vérifications a posteriori constituent l'étape de validation statistique du modèle.

In fine, ces étapes mènent à la sélection d'un modèle qui a la meilleure vraisemblance en évitant le surapprentissage. Les étapes de validation du modèle assurent en outre l'unicité du meilleur estimateur des paramètres d'intérêt.

Pour autant, ces étapes ne peuvent constituer une garantie que les estimations, certes uniques, soient identifiables à la propriété physique recherchée, à plus forte raison en conditions de mesures non intrusives. A titre d'exemple, Deconinck and Roels (2017) montrent que la somme des paramètres résistifs de certains modèles RC s'avère interprétable physiquement tandis que ceux-ci n'étaient pas tous considérés comme significatifs donc ne sont pas par ailleurs considérés comme statistiquement validés. L'identifiabilité couplée à la validation statistique ne semble donc pas être une condition parfaitement suffisante ni nécessaire à l'interprétabilité physique des paramètres d'un modèle RC calé.

La suite de ce travail développe une méthodologie pour évaluer spécifiquement l'interprétabilité de modèles RC dans le cadre de calage à partir de mesures en conditions non intrusives. Cette méthodologie est appliquée à la détermination d'une durée de mesures minimales pour estimer une résistance thermique globale de l'enveloppe en s'affranchissant de l'influence des conditions aux limites extérieures au bâtiment. Puis, l'opportunité d'identifier séparément les échanges thermiques par la ventilation ou par une paroi donnant sur un vide sanitaire non chauffé, par des modèles RC ad hoc, séparément du reste des échanges thermiques est également évaluée.

A.3 Répétabilité de l'estimation en conditions météorologiques variables

En conditions de mesures non intrusives, le chauffage est contrôlé de manière à offrir une ambiance intérieure compatible à de l'occupation. Comparé à des méthodes intrusives où la puissance de chauffage délivrée est par exemple pseudo-aléatoire, le chauffage en non intrusif offre une sollicitation moins riche. De même, les températures atteintes dans l'ambiance intérieure présentent des écarts plus faibles qu'en intrusif dans un scénario optimisé. Ainsi, au regard des échanges thermiques à l'échelle de l'enveloppe, les conditions météorologiques, naturellement variables et imprévisibles, prennent comparativement une part plus importante dans la dynamique en non intrusif qu'en intrusif. De ce fait, si la durée de mesure est trop faible, il est très probable que les conditions météorologiques particulières pendant l'acquisition aient un impact sur la justesse et la précision de l'estimation de la performance thermique de l'enveloppe. Démontrer qu'il est possible d'obtenir une estimation fiable à partir de mesures non intrusives équivaut alors à déterminer la durée minimale de mesures nécessaires pour atteindre une robustesse de l'estimation vis à vis des conditions météorologiques.

Cette section développe la méthodologie appliquée et les résultats de l'étude de la répétabilité de l'estimation d'une résistance thermique globale de l'enveloppe en conditions non intrusives, quand des conditions météorologiques sont naturellement variables, sur un site donné.

A.3.1 Méthodologie

L'objectif de la méthodologie est d'évaluer la durée minimale de mesures non intrusives nécessaires à une estimation de la résistance thermique globale de l'enveloppe R_{eq}. Pour cela, un banc de test numérique a été développé à partir d'un modèle de bâtiment numérique appelé modèle de référence, servant à la simulation thermique. Ce modèle, soumis à un fichier de données météorologiques, permet d'effectuer des mesures d'un essai virtuel, dont les données sont ensuite

traitées et servent au calage d'un modèle RC stochastique. Les paramètres ainsi estimés du modèle RC permettent d'estimer R_{eq}. Comme le modèle de référence est numérique, il est possible d'en déterminer une valeur de résistance thermique globale cible, appelée R^*_{eq}. Chaque simulation permet donc une calibration de modèle RC, qui permet une estimation de R_{eq}, laquelle est comparée à la valeur cible R^*_{eq}.

Pour déterminer la durée minimale de mesures pour une estimation robuste, le principe décrit en figure A.3 est appliqué. L'idée est d'effectuer N simulations thermiques dynamiques (STD), toutes à partir du même modèle, mais chacune soumise à un fichier différent de données météorologiques. De chaque STD sont extraits des jeux de données de durées croissantes, tous démarrant le même jour: extraction de 2 jours, puis 3 jours, puis 5 jours, etc. Chaque jeu de données permet alors de caler un modèle RC. Une STD permet donc d'extraire plusieurs jeux de données et donc d'estimer une valeur de R_{eq} avec une durée croissante de mesures virtuelles, ce qui permet d'évaluer la convergence de l'estimation de R_{eq} vers la valeur cible R^*_{eq}.

L'analyse de l'influence des conditions météorologiques sur cette convergence se fait par l'étude du lien entre les conditions météorologiques de chaque jeu de données et l'estimation de R_{eq}. Une durée de mesures trop courte induira une forte influence de ces conditions sur l'estimation, donc une forte variabilité. Au contraire, une durée suffisamment longue atténuera significativement la variabilité de l'estimation de R_{eq} en convergeant, en principe, vers la valeur cible R^*_{eq}.

Les fichiers météorologiques servant aux N simulations thermiques dynamiques du modèle de référence sont des fichiers générés stochastiquement à partir de données réelles hivernales représentatives d'un mois de janvier à Genève, selon la méthode de Goffart et al. (2017a). Les variables météorologiques générées sont la température extérieure, l'humidité relative, l'irradiation solaire directe et diffuse ainsi que la vitesse et l'orientation du vent. Les variations induites par la génération stochastique sont donc représentatives d'une variabilité naturelle de ces variables météorologiques et couvrent une grande diversité de conditions naturellement probables sur un mois de janvier à Genève.

Pour évaluer la fiabilité et la précision de l'estimation de R_{eq} en comparaison à sa valeur cible, le calcul de l'erreur entre le maximum de vraisemblance et la valeur cible seul ne permet pas de rendre compte de l'information portée par l'écart-type de l'estimation, comme le suggère la figure A.4. Les estimateurs de maximum de vraisemblance (haut de la courbe) jaune et verte ont des valeurs très proches, mais leur incertitude est différente. L'estimation jaune n'est notamment pas acceptable, tandis que la verte n'est certes pas précise, mais la valeur cible est comprise dans son intervalle de confiance à 95%. Ainsi, pour discriminer entre ces deux cas de figure, un indicateur a été développé, dorénavant appelé indicateur d'interprétabilité. Il s'agit de l'aire sous la courbe de l'estimation (supposée gaussienne ou échantillonnée en Bayésien) commune à l'aire grise de la figure A.4, c'est-à-dire la valeur cible $\pm \cdot 10\%$, ou 5% selon le degré de discrimination voulu. L'indicateur s'échelonne alors de 0 pour les estimations les plus faussées à 1 pour les estimations très précises. Cet indicateur discrimine ainsi les estimations jaune et verte, avec un score de 0.08 pour la première mais de 0.36 pour la seconde.

La section des résultats montrera donc l'évolution de la variabilité de l'estimation de R_{eq} et de son indicateur d'interprétabilité en fonction de la durée des mesures virtuelles ainsi qu'en fonction des conditions météorologiques lors de l'essai virtuel.

A.3.2 Durée minimale de mesures

Chacune des 2000 simulations faites sous chacun des 2000 fichiers météorologiques a permis d'extraire des jeux virtuels de mesures de durées croissantes, à commencer par une durée de mesures de 48h. Ces jeux virtuels d'une durée de 48h servent donc au calage du modèle RC $T_w T_i\ R_o R_i\ A_w$ qui lui même permet d'estimer R_{eq}.

La figure A.5 montre la variabilité, pour 2 jours de données, de 50 estimations parmi les 2000

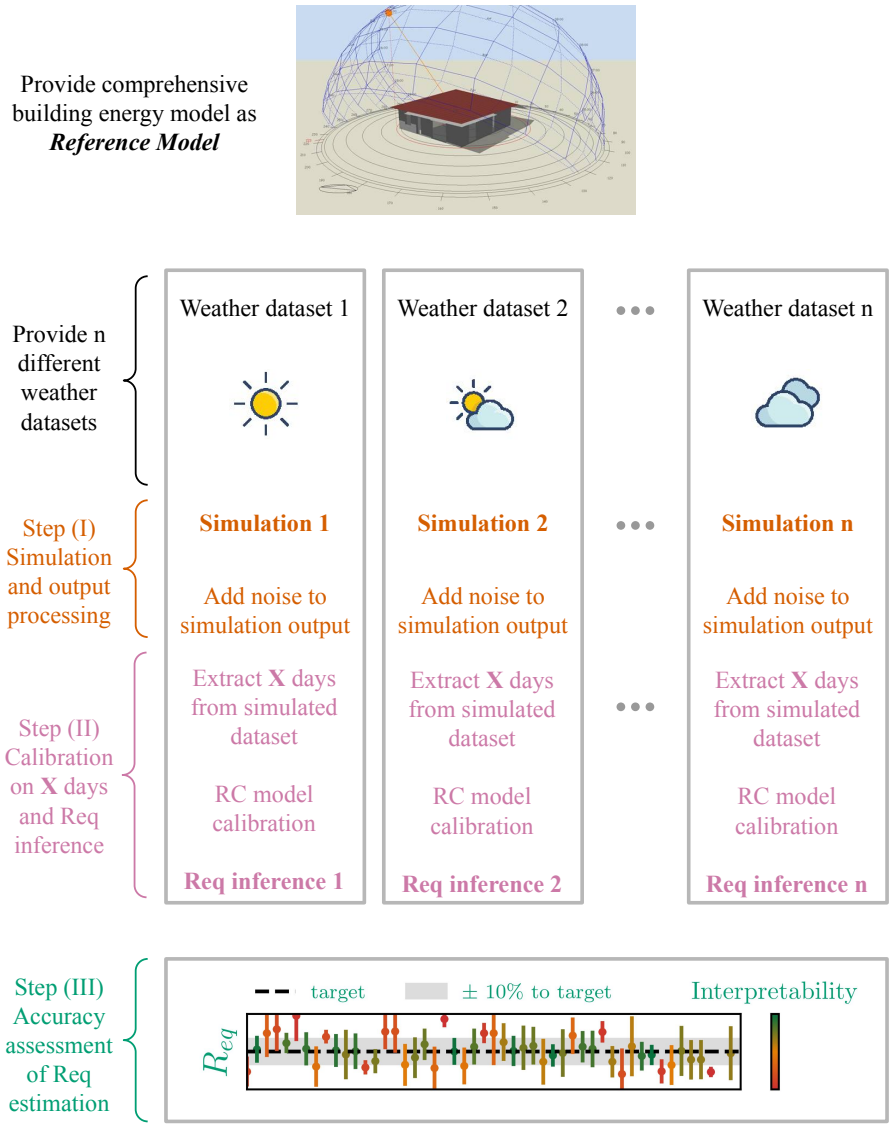


Figure A.3 – Méthodologie pour déterminer la durée minimale de mesures pour obtenir une estimation robuste et fiable de la résistance thermique globale de l'enveloppe, quand les mesures sont non intrusives.

estimations de R_{eq}, comparé en gris à la valeur cible $\pm 5\%$ et $\pm 10\%$. La couleur renseigne sur l'indicateur d'interprétabilité: les couleurs vertes sont souhaitables tandis que les rouges ne le sont pas. A deux jours, les estimations de R_{eq} montrent une variabilité élevée et une fiabilité faible. Cela est confirmé par le graphique de droite qui montre la variabilité du maximum de vraisemblance de chaque estimation. De nombreuses estimations sont très au delà des bornes acceptables autour de la valeur cible, même avec une borne à $\pm 10\%$ pourtant moins discriminante.

Il apparaît donc nécessaire d'augmenter la durée de mesure jusqu'à observer une variabilité moindre dans les estimations de R_{eq}. La figure A.6 montre précisément comment la variance des maximums de vraisemblance diminue avec des durées de mesures plus longues. A partir de 8 jours d'essai virtuel, la variance totale diminue de 60%, à partir de 11 jours de 75%. Au delà, la variance poursuit sa décroissance mais de manière moins significative.

L'analyse de sensibilité permise par l'utilisation des 2000 fichiers météorologiques suivant Goffart et al. (2017a) permet alors de déterminer les variables météorologiques responsables de la variabilité observée dans les estimations. Les parts d'influence de chacune des 6 variables générées

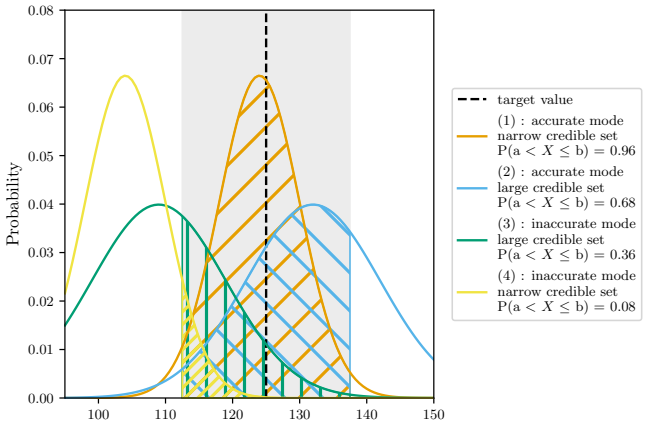


Figure A.4 – Illustration de 4 catégories de distributions postérieures (données synthétiques purement illustratives). Les bornes a et b, par exemple à ±10 % de la valeur cible, définissent une aire d'intérêt (en gris) autour de la valeur cible (en pointillés noirs). La probabilité de chaque distribution d'être dans ces bornes est alors identique à l'aire sous la courbe (hachures) et définit un indicateur quantitatif qui prend valeur entre 0 et 1.

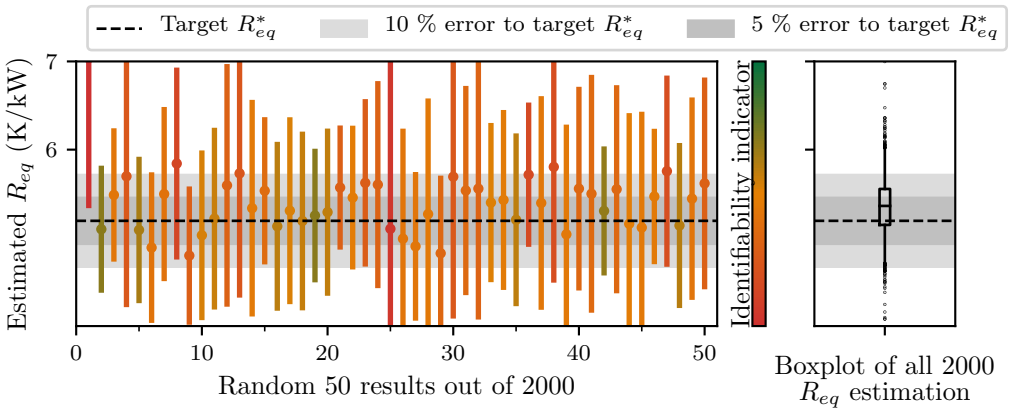


Figure A.5 – Variabilité des estimations de R_{eq} sur des jeux de données de 2 jours.

dans les fichiers sur la variance totale sont aussi visibles dans la figure A.6. La température extérieure et la vitesse de vent semblent être les principales sources de variance des estimations de R_{eq}, quelle que soit la durée d'essai. Cette influence est majeure pour des durées de mesures courtes, vue la forte variance totale. L'irradiation solaire directe a un effet significatif sur des durées inférieures ou égales à 8 jours, puis cet effet s'estompe.

L'influence des variables de température extérieure et vitesse de vent pouvait être attendue dans la mesure où le modèle de référence est basé sur un cas d'étude dont le renouvellement d'air, modélisé comme dépendant de la vitesse de vent et de la différence de température entre l'intérieur et l'extérieur justement, est précisément élevé (autour de $1 vol/h$). Ainsi, l'estimation de la résistance thermique globale de l'enveloppe, nécessairement dépendante du renouvellement d'air, montre une dépendance à la variabilité de ces deux variables météorologiques, vu la variabilité imposée de ces entrées. Si le modèle de référence avait un renouvellement d'air plus faible, on pourrait s'attendre à une robustesse des estimations à partir d'une durée de mesures plus courte, peut-être à partir de 8 jours. Pour autant, la durée minimale pour garantir une robustesse vis à vis des conditions météorologiques devrait tout de même se situer dans l'ordre de grandeur 8-11 jours, quelle que soit la performance du bâtiment.

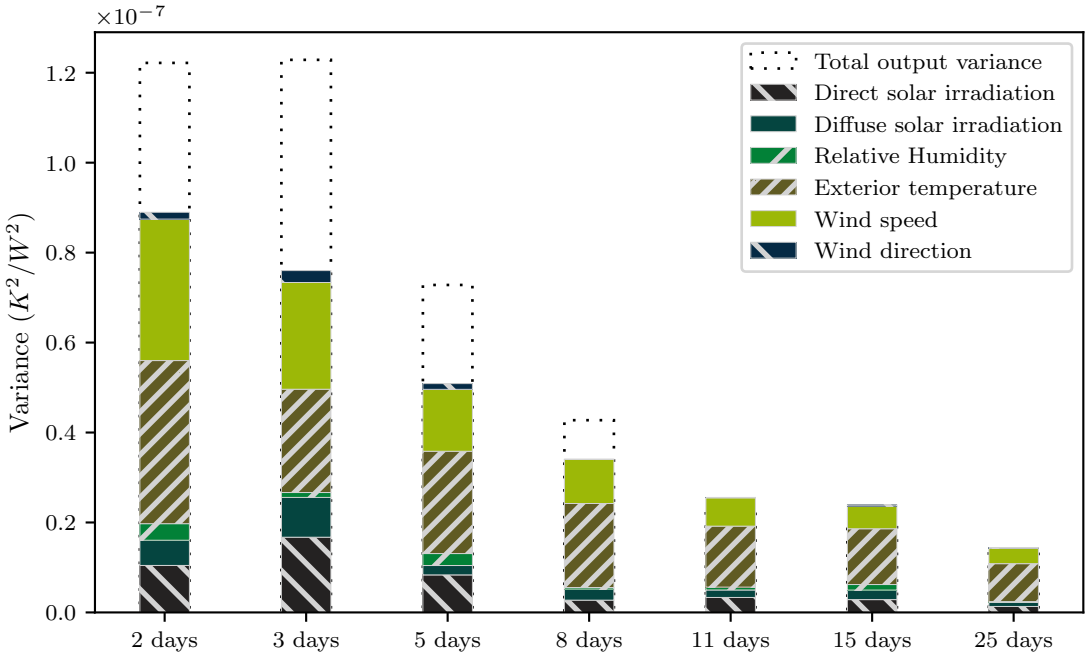


Figure A.6 – Diminution de la variance totale à mesure que la durée de l'essai augmente

A.3.3 Conclusion sur la répétabilité de l'estimation de R_{eq} en conditions météorologiques variables

La méthodologie numérique développée pour évaluer la robustesse d'une estimation de la résistance thermique globale de l'enveloppe R_{eq} a permis de déterminer qu'en conditions de mesures non intrusives, des durées de mesures trop courtes produisent des estimations fortement variables et dépendantes des conditions météorologiques particulières de l'essai. A partir de 11 jours, cette variabilité s'estompe significativement. La variabilité résiduelle observée dans le cas d'étude utilisé ne tient qu'au renouvellement d'air particulièrement élevé. Il est alors probable que pour d'autres cas d'études, l'estimation de R_{eq} puisse être faite de manière robuste avec une durée un peu moindre, bien qu'il soit très probable que l'ordre de grandeur de la durée, environ 10 jours, se maintienne. Ce point serait sujet à vérification à l'avenir.

Dans le même temps, ces résultats mettent en avant le besoin de critères de convergence pour l'estimation de R_{eq} par modèles RC stochastiques, à l'instar de l'ISO 9869-1 (ISO 9869-1, 2014) dont des critères guident l'arrêt de mesures pour l'estimation du $U - value$ d'une paroi. En particulier, ces critères devraient pouvoir mesurer l'évolution de l'apprentissage du modèle par les données récoltées, en lien avec les conditions météorologiques. Déterminer de tels critères de convergence permettrait d'établir, en conditions réelles quand la valeur cible est inconnue, une durée de mesures pour une estimation fiable et robuste.

A.4 Identification plus fine des pertes thermiques

Pour établir des scenarii de rénovation thermique pertinents et ciblés, il est nécessaire de diagnostiquer finement la performance de l'enveloppe. Idéalement, les pertes thermiques par le renouvellement d'air sont déterminées séparément des pertes par le reste de l'enveloppe. Puis, déterminer la résistance thermique des parois opaques verticales, des planchers hauts et bas ainsi

que des fenêtres permettrait une connaissance plus fine encore de l'enveloppe.

Des mesures non intrusives peuvent cependant ne pas être assez informatives pour permettre une identification fine des pertes thermiques. En particulier, le caractère agrégé des modèles RC stochastiques pose question quant à l'interprétabilité de chacun de ses paramètres. Pour autant, la nature dynamique des données récoltées pourrait être bénéfique à l'apprentissage.

Cette section étudie donc l'identifiabilité et interprétabilité des paramètres de deux modèles RC, l'un représentant les pertes thermiques par le renouvellement d'air, l'autre les pertes thermiques par le plancher bas. La méthodologie d'étude est présentée pour ensuite décrire les principaux résultats.

A.4.1 Méthodologie

Pour évaluer l'identifiabilité de paramètres de modèles à la propriété physique qu'ils représentent, la méthodologie est, comme à la section précédente, basée sur un modèle de référence numérique, à la différence que les propriétés thermiques de l'enveloppe sont variables. Le principe est de pouvoir calculer les valeurs cibles de résistance thermique des différents éléments de l'enveloppe. Les paramètres du modèle RC calé ad hoc sont donc comparables à une valeur théorique. Autrement dit, si une propriété de l'enveloppe est modifiée, le paramètre du modèle RC est identifiable à celle-ci si son estimation change de valeur en conséquence.

La figure A.7 montre le principe général de la méthodologie. Le même cas d'étude que la section précédente est modélisé et sert de modèle de référence. Ce modèle de référence est alors dupliqué n fois, et la composition de l'enveloppe de chaque copie est modifiée. Les propriétés modifiées sont décrites dans le tableau A.2 et les valeurs sont échantillonnées par hypercube latin. 300 versions du modèle de référence sont ainsi générées.

Chaque copie est simulée avec les mêmes données météorologiques et virtuellement en conditions non intrusives. Des n simulations thermiques dynamiques obtenues sont extraits des jeux de données de 11 jours, conformément aux résultats précédemment obtenus. Trois modèles différents sont calés: $T_w T_i\ R_o R_i\ A_w\ R_b$, $T_w T_i\ R_o R_i\ A_w\ c_v$ et $T_w T_i\ R_o R_i\ A_w$. La procédure de sélection et de validation de modèle désignent les deux premiers comme étant meilleurs que le dernier. Le dernier est donc calé en guise de comparaison entre les deux. Le modèle $T_w T_i\ R_o R_i\ A_w\ R_b$ sert donc à évaluer l'identifiabilité de la résistance thermique par le plancher bas et le modèle $T_w T_i\ R_o R_i\ A_w\ c_v$ l'identifiabilité des pertes thermiques par renouvellement d'air.

A.4.2 Identifiabilité des pertes thermiques par ventilation

Le modèle $T_w T_i\ R_o R_i\ A_w\ c_v$ est calé sur chacun des 300 jeux de données extraits des 300 simulations. L'objectif étant d'évaluer l'identifiabilité des paramètres du modèle pris séparément, la figure A.8 présente la variabilité des estimations au maximum de vraisemblance de chaque paramètre. Idéalement, on devrait observer que les variations du modèle de référence relatives à des résistances thermiques en abscisse entraînent des variations des paramètres résistifs du modèle RC en ordonnées. Parallèlement, des variations liées aux capacités thermiques massiques des matériaux devrait influer sur les capacités du modèle RC. Enfin, les variations du renouvellement d'air devrait influer sur le paramère c_v représentant précisément celui-ci.

La figure A.8 montre, en bleu foncé, les paramètres significativement impactés par les variations des propriétés thermiques du modèle de référence. Tous les paramètres sont infuencés par la variabilité du renouvellement d'air tandis que les paramètres résistifs du modèle RC sont peu ou pas influencés par les variations en épaisseur de l'isolant des parois opaques.

La figure A.9 compare alors les valeurs estimées des résistances thermiques relatives au renouvellement d'air, au reste de l'enveloppe et celle globale de l'enveloppe avec les valeurs cibles calculées à partir du modèle de référence numérique, pour chaque variation de composition. S'il est clair que l'estimation de la résistance thermique globale par le modèle $T_w T_i\ R_o R_i\ A_w\ c_v$ est linéairement très corrélée à la valeur cible, ce n'est pas le cas de la résistance thermique relative

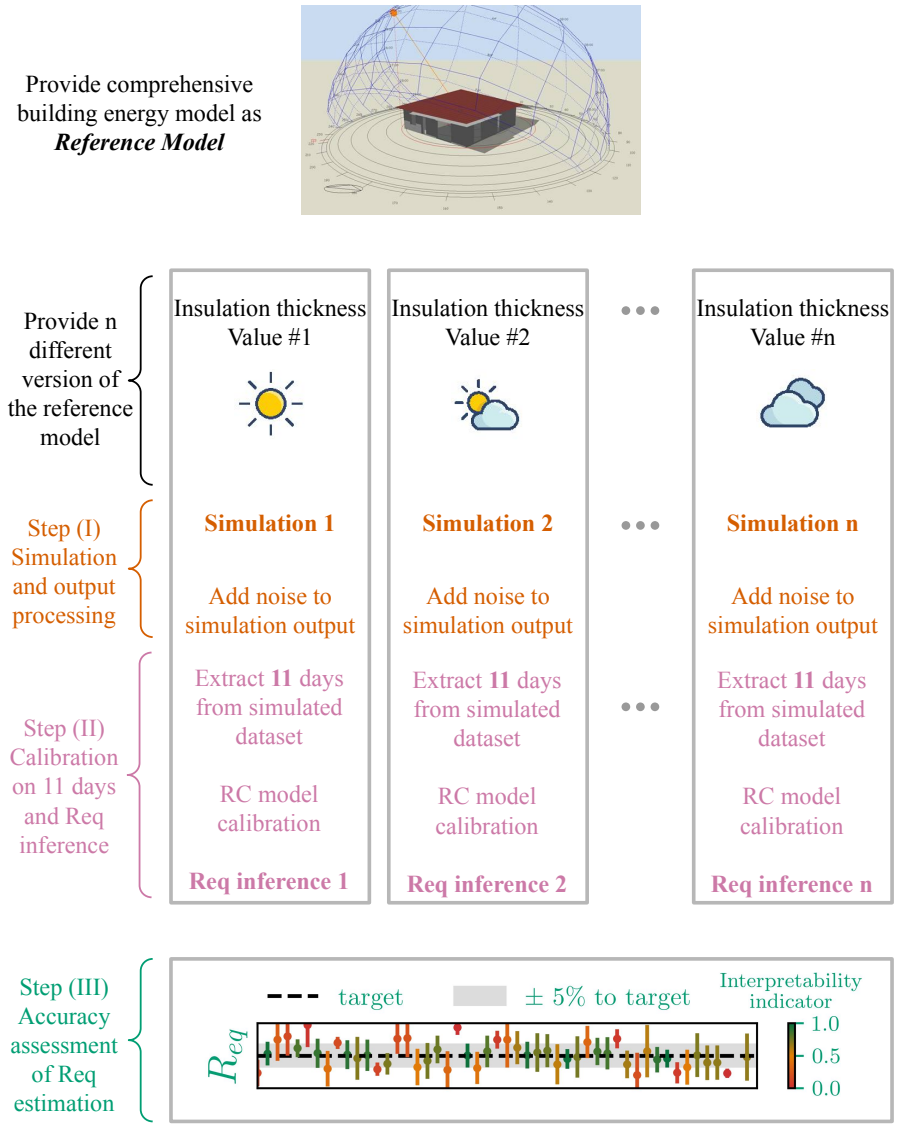


Figure A.7 – Méthodologie pour évaluer l'identifiabilité et interprétabilité de paramètres de modèles RC stochastiques, quand les mesures sont non intrusives. Le modèle de référence est modifié n fois, chaque version présente une enveloppe de différente composition.

au renouvellement d'air ni de celle relative au reste de l'enveloppe. La résistance thermique liée au renouvellement d'air est certes dans un ordre de grandeur plutôt proche des valeurs cibles, notamment pour les renouvellements d'air faibles. Pour autant, l'erreur commise sur l'estimation du renouvellement d'air, à valeur cible égale, est très variable : elle s'étend de -50 à $+90$ %. Une variance aussi élevée n'est pas satisfaisante et rend ces estimations non fiables.

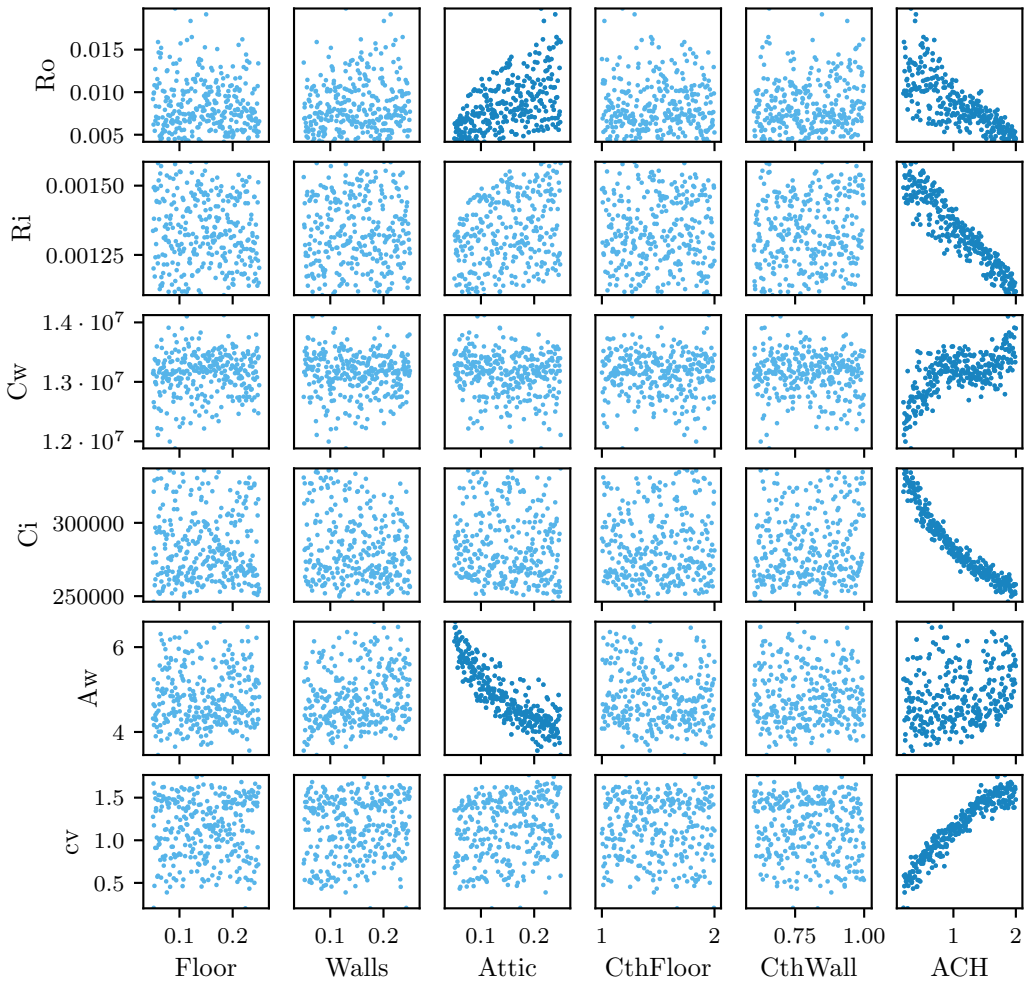


Figure A.8 – Variabilité de chaque paramètre du modèle T_wT_i R_oR_i c_v (axe des ordonnées) en fonction de la variabilité des 6 entrées de l'analyse de sensibilité (axe des abscisses). Les couleurs plus foncées mettent en évidence les influence significatives des entrées sur les estimations de paramètres.

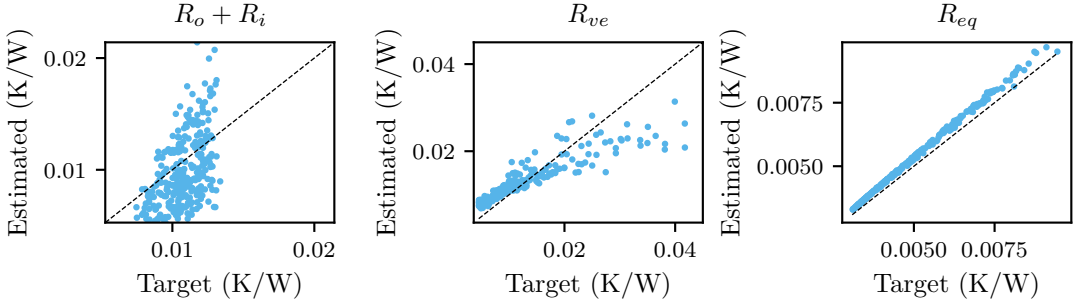


Figure A.9 – Interprétabilité de $R_o + R_i$, R_{ventil} et R_{eq} estimés à partir du modèle T_wT_i R_oR_i c_v

	Object Variable in EnergyPlus	**Bornes**	**Distribution**	**Unités**	**Info. add.**
Walls insulation thickness	(Material) Thickness	$[0.05; 0.25]$	Uniforme	m	$\lambda = 0.032\ W/(m \cdot K)$
Attic insulation thickness	(Material) Thickness	$[0.05; 0.25]$	Uniforme	m	$\lambda = 0.04\ W/(m \cdot K)$
Ground floor insulation thickness	(Material) Thickness	$[0.05; 0.25]$	Uniforme	m	$\lambda = 0.022\ W/(m \cdot K)$
Air Change Rate	(ZoneVentilation: DesignFlowRate) Air changes per Hour	$[0.2; 2.0]$	Uniforme	h^{-1}	(-)
Thermal capacity brick wall	(Material) Specific Heat	$[0.6; 1.0]$	Uniforme	J/kgK	$e = 13$ cm
Thermal capacity floor concrete cast	(Material) Specific Heat	$[1.0; 2.0]$	Uniforme	J/kgK	$e = 20$ cm

Table A.2 – Variations des propriétés thermiques du modèle de référence

A.4.3 Identifiabilité des pertes thermiques au travers du plancher bas

Pour évaluer l'identifiabilité des pertes thermiques au travers du plancher bas, le modèle $T_w T_i\ R_o R_i\ A_w\ R_b$ est calé sur chacun des 300 jeux de données extraits des 300 simulations. De même que pour l'application précédente, on s'attend idéalement à une dépendance linéaire entre les variations d'épaisseur d'isolant du plancher bas vers le vide sanitaire et le paramètres R_b, supposé représenter la résistance thermique du plancher bas. En effet, le modèle $T_w T_i\ R_o R_i\ A_w\ R_b$ fait intervenir une mesure de la température dans le vide sanitaire, ce qui laisse à penser que grâce à l'information complémentaire gagnée par cette mesure, la décomposition peut être faisable.

La figure A.10 montre au contraire que le paramètre R_b n'est influencé que par les variations du l'épaisseur de l'isolant dans les combles et par le renouvellement d'air.

Pour compléter l'analyse, la comparaison des estimations de résistances thermiques du plancher bas, du reste de l'enveloppe et de la résistance thermique globale à leurs valeurs cibles sont montrées en figure A.11. Les graphes montrent que l'estimation de la résistance thermique globale est très corrélée à la valeur cible et est a priori fiable. Au contraire, les résistances estimées permettant a priori la décomposition ne montrent aucune corrélation et ne peuvent en l'état être interprétées physiquement.

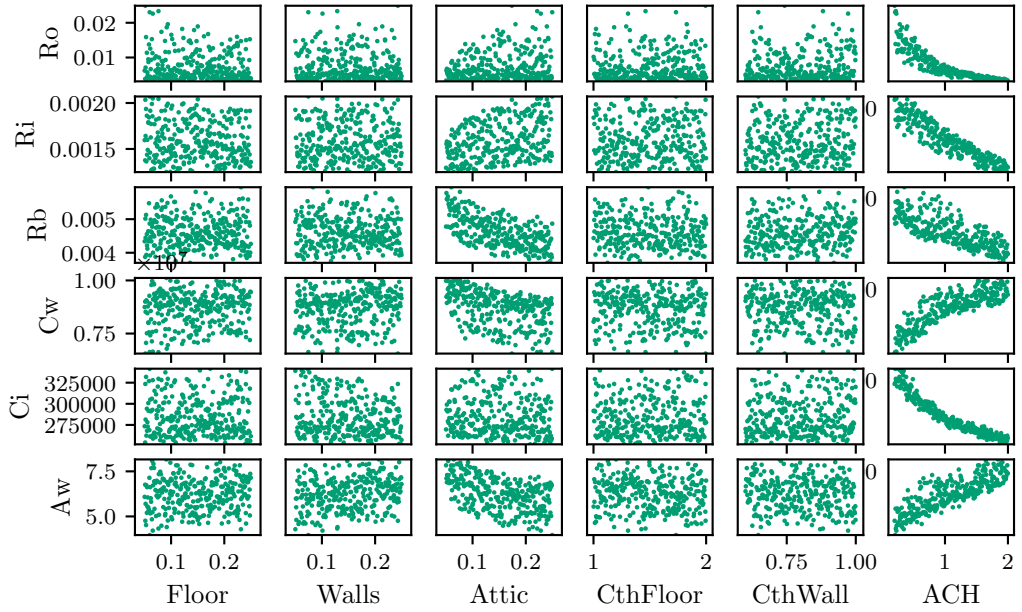


Figure A.10 – Variabilité des paramètres R_o, R_i, R_b, C_w, C_i et A_w du modèle T_wT_i $R_oR_iR_b$

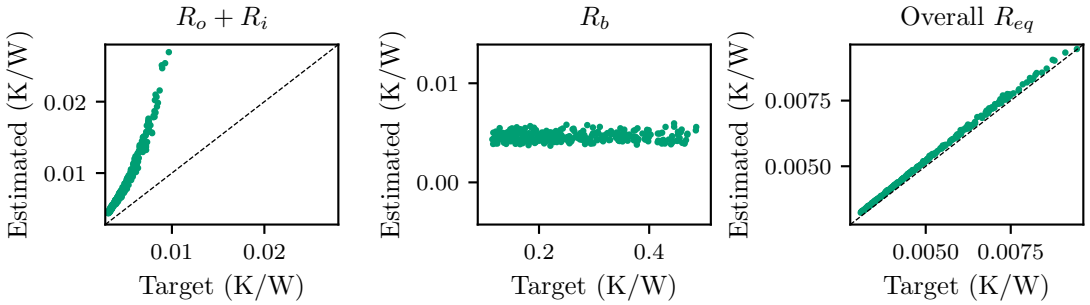


Figure A.11 – Interprétabilité physique des résistances thermiques du modèle T_wT_i $R_oR_iR_b$

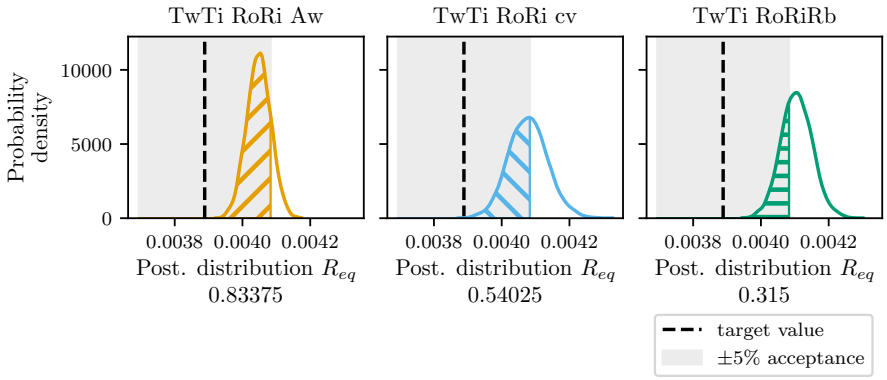
A.4.4 Conclusion sur l'opportunité de la décomposition des pertes thermiques

L'objectif de cette section a été d'évaluer l'identifiabilité et l'interprétabilité physique de paramètres de modèles RC aux propriétés physiques qu'ils représentaient physiquement a priori. Identifiabilité aurait alors signifié possibilité de décomposer les pertes thermiques via le renouvellement d'air ou au travers de parois donnant sur des locaux non chauffés mais dont la température d'air est mesurée.

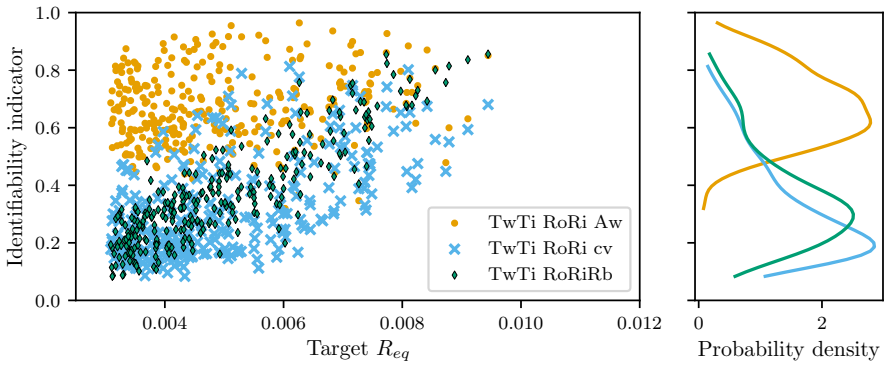
La méthodologie appliquée a pu évaluer cette opportunité par l'évaluation de l'identifiabilité de deux paramètres résistifs: l'un du modèle T_wT_i R_oR_i c_v pour estimer le renouvellement d'air, et l'autre du modèle T_wT_i $R_oR_iR_b$. Les résultats ont montré que ni l'un ni l'autre n'est identifiable, si bien que la décomposition n'est pas fiable, tout au moins quand les mesures sont faites en conditions non intrusives.

Pour autant, les deux modèles testés ont montré une estimation de la résistance thermique globale de prime abord très satisfaisante. De même que dans Deconinck and Roels (2017), il est donc possible que la résistance thermique globale soit estimée de manière satisfaisante par ces deux modèles même si leurs paramètres n'ont individuellement pas de signification. La figure A.12b montre enfin une comparaison des estimations de R_{eq} par ces 2 modèles ainsi que par le modèle de base T_wT_i R_oR_i A_w. La comparaison se fait par le calcul de l'indicateur d'interprétabilité.

La figure A.12b montre que les estimations de R_{eq} par le modèle de base, le plus simple,



(a) Calcul de l'indicateur d'interprétabilité pour les 3 modèles calés (l'aire grise sous la courbe de chaque distribution postérieure).



(b) Comparaison des 3 modèles pour les 300 variations du modèle de référence: la comparaison est faite par leurs indicateurs d'interprétabilité en fonction de la valeur cible R^*_{eq}. En orange, les scores obtenus par le modèle RC de base $T_wT_i\,R_oR_i\,A_w$ sont globalement bien meilleurs que les scores obtenus par les deux autres modèles.

Figure A.12 – Comparaison des modèles par l'indicateur d'identifiabilité

sont celles qui sont les plus fiables et les plus proches de la valeur cible, quelle que soit la valeur cible. Les deux modèles désignés par le processus de sélection et de validation sont un peu moins satisfaisants que le modèle de base.

Ce dernier point ne devrait pas être un étonnement. La sélection et validation de modèle telles que décrites dans le processus de bonnes pratiques sont fondées sur une comparaison de la vraisemblance, elle-même indicatrice de la capacité de prédiction d'un comportement thermique global du bâtiment par le modèle RC. Sans étude de la covariance des estimations des paramètres, le surapprentissage et ainsi la perte en interprétabilité peut ne pas être détecté.

Conclusion générale et perspectives

Résumé des résultats essentiels

Estimer les propriétés thermiques d'une enveloppe de bâtiment in situ en conditions non intrusives pour l'occupant soulève deux défis scientifiques majeurs : les données récoltées sont faiblement informatives et les occupants peuvent considérablement biaiser la qualité des données. Le travail développé dans cette thèse a traité le premier point et s'est attaché à évaluer la faisabilité

d'une estimation des propriétés thermiques d'une enveloppe de bâtiment à partir de mesures non intrusives.

Pour effectuer une estimation de propriétés thermiques d'enveloppe à partir de peu de données, la revue de littérature a identifié les modèles RC stochastiques comme de bons candidats à l'exploitation des données récoltées. Ces modèles sont basés sur la physique donc possiblement interprétables et la littérature suggère que leur calage converge plus rapidement que des modèles de type régression linéaire. Le caractère stochastique de ces modèles permet de plus de prendre en compte dans l'incertitude de l'estimation de l'erreur de modèle intrinsèque à tout modèle simplifié. Cela rend l'incertitude réaliste et représentative de l'état réel de l'information gagnée dans les données.

L'existence et l'unicité de l'estimation, c'est-à-dire son identifiabilité, ont été étudiées dans un deuxième temps. L'identifiabilité dépend d'abord de la formulation mathématique du modèle RC, qui peut aisément être vérifiée, et dépend ensuite de la qualité et de la quantité des données récoltées. L'identifiabilité pratique aussi peut être vérifiée, en observant la variation de la vraisemblance en fonction des valeurs prises par les paramètres. En sus, l'erreur de modèle du fait d'utiliser un modèle simplifié peut être réduite en sélectionnant le modèle RC le plus adéquat, c'est-à-dire celui dont la prédiction est la plus proche des données mesurées. Tout ces éléments constituent ainsi un processus de bonnes pratiques nécessaires à l'estimation de propriétés thermiques. Pourtant, en conditions non intrusives, la faible informativité des données récoltées combinée à l'erreur intrinsèque des modèles RC constituent encore une menace à l'interprétabilité des paramètres estimés.

Pour évaluer la faisabilité de l'estimation de la résistance thermique par l'exploitation de mesures non intrusives par des modèles RC, la suite de la thèse a consisté à développer une méthodologie numérique pour pouvoir tester l'interprétabilité des paramètres des modèles calés utilisés. Ce banc d'essais virtuels a été appliqué à deux évaluations: celle de la répétabilité de l'estimation vis à vis de conditions météorologiques variables et celle d'identifier séparément les pertes thermiques d'une enveloppe, soit par renouvellement d'air, soit au travers du plancher bas donnant sur un vide sanitaire.

La première application a évalué l'influence des conditions météorologiques, naturellement variables, sur l'estimation d'une résistance thermique globale. Les résultats ont montré que des durées de mesures trop courtes provoquent une grande variabilité dans les estimations, dues pour le cas d'étude utilisé, à l'irradiation solaire directe, la température extérieure et la vitesse de vent. Au delà de 11 jours, la variabilité des estimations est 4 fois moins élevée que pour 2 jours de durée et peut être considérée comme satisfaisante.

L'identifiabilité des paramètres calés représentant le renouvellement d'air et la résistance thermique d'un plancher bas a également été évaluée sur le même principe. Les résultats montrent que les paramètres calés des modèles testés ne sont pas interprétables, même si les deux modèles calés présentaient des vraisemblances satisfaisantes. Les covariances élevées des paramètres estimés semblent pouvoir être un signe de non interprétabilité.

En quelques mots, en conditions de mesures non intrusives, les modèles RC stochastiques semblent donner des estimations de la résistance thermique globale d'une enveloppe de bâtiment très satisfaisantes, pourvu que la durée de mesures soit suffisamment longue, donc un peu plus d'une dizaine de jours. Pour autant, décomposer les sources des pertes thermiques par ces mêmes modèles à partir de mesures si peu informatives s'est révélé infructueux. Quoi qu'il en soit, ces travaux ont souligné l'importance et la nécessité d'appliquer le processus de bonnes pratiques pour espérer obtenir des estimations fiables des propriétés d'intérêt.

Perspectives

Les travaux présentés supra suggèrent plusieurs pistes de travail pour assurer la faisabilité de l'estimation de la performance thermique d'une enveloppe par des modèles RC stochastiques.

La première perspective revient à évaluer l'influence de la non représentativité des mesures relevées sur la fiabilité du résultat de l'estimation. En effet, une hypothèse majeure dans ce travail a été de considérer les mesures de température intérieure ou de puissance de chauffage comme sans erreur. Pourtant, les températures présentent toujours des différences entre différentes pièces, et de surcroît y compris au sein d'une même pièce. La puissance de chauffage délivrée dans le bâtiment, en conditions non intrusives, n'est de même pas si facile à obtenir.

Ensuite, la reproductibilité d'une telle estimation dans d'autres bâtiments, avec notamment d'autres typologies et d'autres géométries, devra être démontrée et n'est pas acquise. La typologie aura un effet sur la durée minimale nécessaire à une estimation robuste. La géométrie pourra avoir une influence sur l'incertitude. Le cas de locaux mitoyens chauffés ou non chauffés devra être particulièrement étudié.

Pour la suite, les perspectives de travaux portent sur les verrous scientifiques majeurs soulevés par la présence d'occupants dans les locaux. Désagréger les données de consommations d'énergie, sachant que partie de cette énergie est dédiée à la consommation par l'occupant, constitue un premier frein à traiter. Ensuite, l'impact de l'occupation sur les flux thermo-hygro-aérauliques est une source majeure de biais et d'incertitude. Ces points questionnent la répétabilité d'une estimation sous scenarii d'occupation variable et appellent à une liste de critères pour établir la convergence d'une estimation. Le besoin d'évaluer de façon réaliste les incertitudes du résultat de l'estimation, particulièrement en scenarii occupés, pousse d'autant plus à envisager l'approche Bayésienne, pour proposer une prise en compte systémique des incertitudes liées à une mesure en conditions ni maîtrisées ni contrôlables.

B

Code for structural identifiability derivations

Structural identifiability derivation can be automated in symbolic languages. Chapter 2 has presented the principles, some illustrative application and as a result a set of structurally identifiable RC models. For the interested reader, this annex plainly lays out the code used in Maple and Reduce that has been used in this thesis.

B.1 Introductory note

Section 2.2.3 has presented several methods to determine the structural identifiability of state space models. This annex plainly gives the codes written for the illustrations of each method as well as for all RC models tested.

The code corresponding to the following derivations can also be found on the github page of the BAYREB project.

B.2 Structural identifiability methods: code for the illustrative models

B.2.1 Laplace transform method

The Laplace transform method has been written with Maple, but could be applied with any language that has efficient capabilities in symbolic programming.

The transfer function of model 2.2, T_iRA model in free floating temperature conditions, is calculated with:

```
with(LinearAlgebra);

A := Matrix([-1/(R*C)]);
B := Matrix([1/(R*C), gA/C]);
Id := Matrix([1]);

Htransfert := MatrixInverse(Id*s-A) . B;
```

The transfer function of model 2.3, T_iRA model in controllable temperature conditions, is calculated with:

```
with(LinearAlgebra);

A := Matrix([-1/(R*C)]);
B := Matrix([1/(R*C), gA/C, 1/C]);
Id := Matrix([1]);

Htransfer := MatrixInverse(Id*s-A) . B;
```

As a reminder, once the transfer function is obtained, it needs to be written in its canonical form which then provides the coefficients to use to solve the system, in application of the method described in Section 2.2.3. Although feasible, this last part of the proof has not been automatized in this work.

The use of the functions *simplify* and *RationalCanonicalForm* help determine the canonical form, especially for larger order models. Finally, when the system of coefficients has been established, the function *solve(system,list of parameters)* can be used to verify the uncitity of the solution.

B.2.2 Markov parameters

The Taylor series/Markov parameters method has also been written with Maple, but could also be applied with any language that has efficient capabilities in symbolic programming.

The Markov parameters and following Jacobian of the model 2.2, T_iRA model in free floating temperature conditions, are determined by the following code:

```
with(LinearAlgebra);
with(ArrayTools);

A := Matrix([[-1/(C*R)]]);
B := Matrix([[1/(C*R), gA/C]]);
Cmat := Matrix([1]);

s1 := Cmat.B;
s2 := simplify(Cmat.A.B);
s3 := simplify(Cmat.A.A.B);

S := Concatenate(2, s1, s2, s3)[1];

Jac := Matrix([[diff~(S, C)], [diff~(S, R)], [diff~(S, gA)]]);

Rank(Jac);
```

The Markov parameters and following Jacobian of the model 2.3, T_iRA model in controllable temperature conditions, are determined by the following code (only matrix B changes):

```
with(LinearAlgebra);
with(ArrayTools);

A := Matrix([[-1/(C*R)]]);
B := Matrix([[1/(C*R), gA/C, 1/C]]);
Cmat := Matrix([1]);

s1 := Cmat.B;
s2 := simplify(Cmat.A.B);
s3 := simplify(Cmat.A.A.B);

S := Concatenate(2, s1, s2, s3)[1];

Jac := Matrix([[diff~(S, C)], [diff~(S, R)], [diff~(S, gA)]]);

Rank(Jac);
```

B.2.3 DAISY algorithm: code of models

To apply the DAISY algorithm, the Models 2.2 and 2.3 need to be written in REDUCE for the DAISY package to process them. As suggested by the package documentation, one may either write each command separately directly in the REDUCE software or write all commands in a separate text file and simply call the file to be executed in the REDUCE environment. The second option is applied here. Model 2.2 is called as following :

- the description in line 1 is written as reference,
- line 2 and 3 declare the time dependant variables,
- line 4 declares the parameters to identify,
- lines 5, 6 and 7 the number of inputs, outputs and states of the model,
- line 8 and 9 describe the model,
- line 10 calls the DAISY algorithm,

- and line 11 ends the call to the text file.

Note : as Reduce works intrinsically with small caps only, the single temperature state T is named Ti in the following code as not to confound it with the time variable t.

```
WRITE "MODEL 1R1C1A free floating conditions"$
B_:={Text,Isol,y,Ti}$
FOR EACH EL_ IN B_ DO DEPEND EL_,t$
B1_:={C,R,gA}$
NU_:=2$
NY_:=1$
NX_:=1$
C_:={df(Ti, t) = -1/(C*R)*Ti + 1/(C*R)*Text + gA/C*Isol,
y=Ti}$
FLAG_:=1$
DAISY()$
END$
```

As for Model 2.3, it is written as following :

```
WRITE "MODEL 1R1C1A with heating input"$

B_:={Text,Isol,Ph,y,Ti}$

FOR EACH EL_ IN B_ DO DEPEND EL_,t$
B1_:={C,R,gA}$

NU_:=3$
NY_:=1$
NX_:=1$

C_:={df(Ti, t)=-1/(C*R)*Ti+1/(C*R)*Text+gA/C*Isol+1/C*Ph,
y=Ti}$

FLAG_:=1$
DAISY()$
END$
```

B.3 Structural non-identifiability: code for proof

Section 2.2.4 discussed the non identifiability of a few RC models. Code to test them with the Daisy algorithm are available in this section.

B.3.1 Code for Model $T_iR\ g \times A$

Model $T_iR\ g.A$ from model description 2.4 is described in Reduce by the following lines:

```
WRITE "MODEL 1R1C g A"$
B_:={Text,Isol,Ph,y,Ti}$
FOR EACH EL_ IN B_ DO DEPEND EL_,T$
```

```
B1_:={C,R,g,A}$
NU_:=3$
NY_:=1$
NX_:=1$
C_:={df(Ti, t)=1/(C*R)*Text-1/(C*R)*Ti+g*A/C*Isol+1/C*Ph,
y=Ti}$
SEED_:=100000$
DAISY()$
END$
```

B.3.2 Code for Model $T_iRA\ \eta_{hea}$

Model $T_iRA\ \eta_{hea}$ from model description 2.5 is described in Reduce by the following lines:

```
WRITE "MODEL 1R1C1A eta"$
B_:={Text,Isol,Ph,y,Ti}$
FOR EACH EL_ IN B_ DO DEPEND EL_,T$
B1_:={C,R,A,eta}$
NU_:=3$
NY_:=1$
NX_:=1$
C_:={df(Ti, t)=1/(C*R)*Text-1/(C*R)*Ti+g*A/C*Isol+eta/C*Ph,
y=Ti}$
SEED_:=100000$
DAISY()$
END$
```

B.3.3 Model equations and code for $T_wT_iR_oR_iA_eA_i\ R_{hext}R_{hint}$

Model $T_wT_iR_oR_iA_eA_i\ R_{hext}R_{hint}$ from model description 2.6 is governed by the following equations:

$$\begin{cases} \begin{bmatrix} \dot{T}_w \\ \dot{T}_i \end{bmatrix} = \begin{bmatrix} -\frac{1}{C_w}\left(\frac{1}{R_o}+\frac{1}{R_i}-\frac{R_{hint}}{R_i(R_i+R_{hint})}-\frac{R_{hext}}{R_o(R_o+R_{hext})}\right) & \frac{1}{C_w(R_i+R_{hint})} \\ \frac{1}{C_i(R_i+R_{hint})} & -\frac{1}{C_i(Ri+R_{hint})} \end{bmatrix} \begin{bmatrix} T_w \\ T_i \end{bmatrix} \\ \qquad + \begin{bmatrix} \frac{1}{C_w(R_o+R_{hext})} & \frac{A_wR_{hext}}{C_w(R_o+R_{hext})}+\frac{A_iR_{hint}}{C_w(R_i+R_{hint})} & 0 \\ 0 & \frac{A_iR_i}{C_i(R_i+R_{hint})} & \frac{1}{C_i} \end{bmatrix} \begin{bmatrix} T_e \\ I_{\text{sol}} \\ \Phi_{hea} \end{bmatrix} \\ y = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} T_w \\ T_i \end{bmatrix} + \varepsilon \end{cases} \tag{B.1}$$

Code for establishing structural non identifiability of Model $T_wT_iR_oR_iA_eA_i\ R_{hext}R_{hint}$ (2.6) is described in Reduce by the following lines:

```
WRITE "MODEL TwTi RoRi Rhe Rhi AwAi"$
B_:={Text,Isol,Ph,y,Ti}$
FOR EACH EL_ IN B_ DO DEPEND EL_,T$
B1_:={Cw,Ci,Ro,Ri,Aw,Ai,Rhe,Rhi}$
NU_:=3$
NY_:=1$
NX_:=2$
```

```
C_:={df(Tw, t) = -1/Cw*(1/Ro+1/Ri-Rhi/(Ri*(Ri+Rhi)) - Rhe/(Ro*(
    Ro+Rhe)))*Tw + 1/(Cw*(Ri+Rhi))*Ti + 1/(Cw*(Ro+Rhe))*Text + (
    Aw*Rhe/(Cw*(Ro+Rhe))+Ai*Rhi/(Cw*(Ri+Rhi)))*Isol,
df(Ti, t) = 1/(Ci*(Ri+Rhi))*Tw - 1/(Ci*(Ri+Rhi)*Ti + Ai*Ri/(Ci*(
    Ri+Rhi))*Isol + Ph/Ci,
y=Ti}$
SEED_:=100000$
DAISY()$
END$
```

B.4 A set of structurally identifiable models: available code

Section 2.2.5 has established a list of structurally identifiable models. This sections lists the code in Reduce to apply the DAISY algorithms directly.

On the BAYREB github page, these codes are directly available in txt formats. They can be used straight away by running the following lines in the Reduce environment, which creates a 'EXAMPLERESULTS.txt' file with the results in the current directory 'CD':

```
LOAD DAISY$
IN "CD\EXAMPLE.TXT"$
OUT "CD\EXAMPLERESULTS.TXT"$
SHUT "CD\EXAMPLERESULTS.TXT"$
```

B.4.1 Code for Model $T_wT_i\ R_oR_i\ A_wA_i$

Model $T_wT_i\ R_oR_i\ A_wA_i$ from model description 2.7 is described in Reduce by the following lines:

```
WRITE "MODEL TwTi RoRi AwAi"$
B_:={Text,Isol,Ph,y,Tw,Ti}$
FOR EACH EL_ IN B_ DO DEPEND EL_,T$
B1_:={Cw,Ci,Ro,Ri,Aw,Ai}$
NU_:=3$
NY_:=1$
NX_:=2$
C_:={df(Tw, t) = -1/Cw*(1/Ro+1/Ri)*Tw + 1/(Cw*Ri)*Ti + 1/(Cw*Ro)
    *Text + Aw/Cw*Isol,
df(Ti, t) = 1/(Ci*Ri)*Tw - 1/(Ci*Ri)*Ti + Ai/Ci*Isol + 1/Ci*Ph,
y=Ti}$
SEED_:=100000$
DAISY()$
END$
```

B.4.2 Code for Model $T_wT_i\ R_oR_i\ A_wA_i\ c_v$

Model $T_wT_i\ R_oR_i\ A_wA_i\ c_v$ from model description 2.8 is described in Reduce by the following lines:

```
WRITE "MODEL TwTi RoRi AwAi cv"$
```

```
B_:={Text,Isol,Ph,Pventil,y,Tw,Ti}$
FOR EACH EL_ IN B_ DO DEPEND EL_,T$
B1_:={Cw,Ci,Ro,Ri,Aw,Ai,cv}$
NU_:=4$
NY_:=1$
NX_:=2$
C_:={df(Tw, t) = -1/Cw*(1/Ro+1/Ri)*Tw + 1/(Cw*Ri)*Ti + 1/(Cw*Ro)
    *Text + Aw/Cw*Isol,
df(Ti, t) = 1/(Ci*Ri)*Tw - 1/(Ci*Ri)*Ti + Ai/Ci*Isol + 1/Ci*Ph +
     cv/Ci*Pventil,
y=Ti}$
SEED_:=100000$
DAISY()$
END$
```

B.4.3 Code for Model $T_wT_i\ R_oR_iR_p\ A_wA_i$

Model $T_wT_i\ R_oR_iR_p\ A_wA_i$ from model description 2.9 is described in Reduce by the following lines:

```
WRITE "MODEL TwTi RoRi Rp AwAi"$
B_:={Text,Isol,Ph,y,Tw,Ti}$
FOR EACH EL_ IN B_ DO DEPEND EL_,T$
B1_:={Cw,Ci,Ro,Ri,Rp,Aw,Ai}$
NU_:=3$
NY_:=1$
NX_:=2$
C_:={df(Tw, t) = -1/Cw*(1/Ro+1/Ri)*Tw + 1/(Cw*Ri)*Ti + 1/(Cw*Ro)
    *Text + Aw/Cw*Isol,
df(Ti, t) = 1/(Ci*Ri)*Tw - 1/Ci*(1/Ri+1/Rp)*Ti + 1/(Ci*Rp)*Text
    + Ai/Ci*Isol + 1/Ci*Ph,
y=Ti}$
SEED_:=100000$
DAISY()$
END$
```

B.4.4 Code for Model $T_wT_i\ R_oR_i\ R_b\ A_wA_i$

Model $T_wT_i\ R_oR_i\ R_b\ A_wA_i$ from model description 2.10 is described in Reduce by the following lines:

```
WRITE "MODEL TwTi RoRi Rb AwAi"$
B_:={Text,Isol,Ph,Tcs,y,Tw,Ti}$
FOR EACH EL_ IN B_ DO DEPEND EL_,T$
B1_:={Cw,Ci,Ro,Ri,Rb,Aw,Ai}$
NU_:=4$
NY_:=1$
NX_:=2$
C_:={df(Tw, t) = -1/Cw*(1/Ro+1/Ri)*Tw + 1/(Cw*Ri)*Ti + 1/(Cw*Ro)
    *Text + Aw/Cw*Isol,
df(Ti, t) = 1/(Ci*Ri)*Tw - 1/Ci*(1/Ri+1/Rb)*Ti + Ai/Ci*Isol + 1/
    Ci*Ph + 1/(Ci*Rb)*Tcs,
```

```
y=Ti}$
SEED_:=100000$
DAISY()$
END$
```

C

Practical non identifiability of a 3 nodes RC model

Calibration of some 3 nodes models on several different datasets has yielded proof of practical non identifiability. This annex proposes some insight in the behaviour of one 3 nodes model when trained on poorly informative data.

The model selection procedure could not exploit calibration results of model $T_wT_iT_b\ R_oR_i\ R_{ib}R_{bb}\ A_w$ as it showed inconsistent estimations with no apparent reasons. This appendix is meant as documentation of what is thought to be a form of practical non identifiability.

Variability in accuracy of the other tested models has indeed been linked to the variability of the input variables of the sensitivity analysis. As an illustration, large air change rates would be positively related to higher values of resistance parameters.

However, the estimations of an overall thermal resistance with model $T_wT_iT_b\ R_oR_i\ R_{ib}R_{bb}\ A_w$ yielded the results shown in Figure C.1. There seem to be two cases of calibration results: either extremely erroneous or slightly erroneous. This double line seems to be particular to lower air change rates.

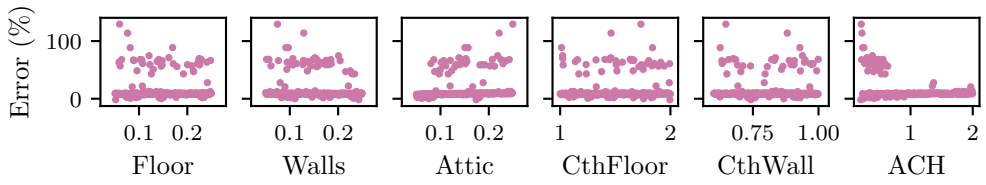


Figure C.1 – Error in estimation of R_{eq} against each estimation: there is a clear double line in estimation, apparently only when the air change rate are below $0.6\ h^{-1}$

One might expect these differences to be failed calibrations, which should be visible in the log-likelihood. Figure C.2 shows precisely that there is no correlation between the error and the log-likelihood. On the contrary, higher log-likelihoods are supposed to show stronger evidence for a good fit on the data but the double line happens exactly for the calibrations with the highest log-likelihoods.

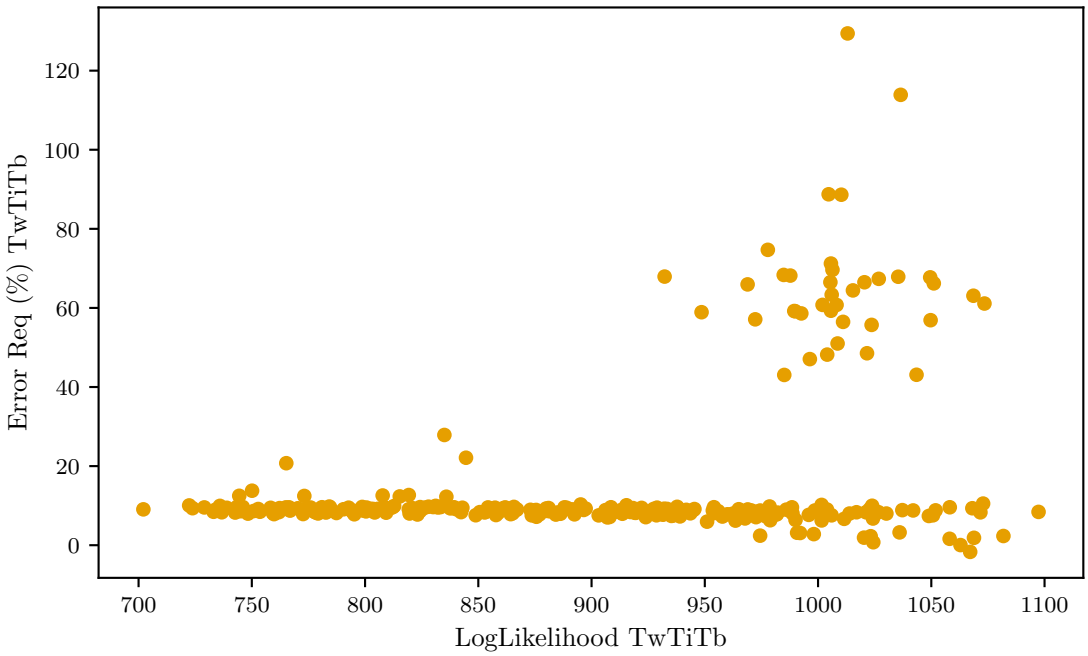


Figure C.2 – Error in estimation of R_{eq} against log-likelihood of the estimation: the double line in estimation is not related to the likelihood

This case seems to be a case of practical non identifiability where there are at least two equally probable values, as though there were a set of parameters for which model $T_wT_iT_b\ R_oR_i\ R_{ib}R_{bb}\ A_w$ overfits the data.

The covariances between all estimated parameters for one of the configurations shows how some parameters are very strongly correlated to one another. This means that some of the parameter values may vary in opposite or identical ways and still provide similar likelihoods. As found in Chapter 5, physical interpretation from correlated estimations should not be performed, and in this case particularly.

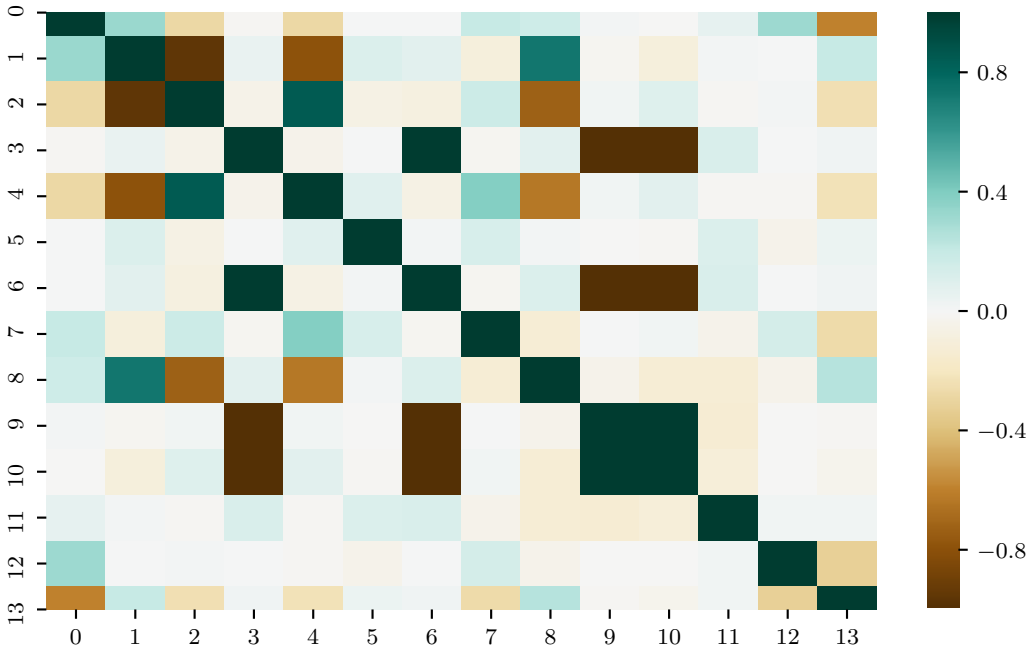


Figure C.3 – Covariance map of all estimated parameters for one deeply erroneous estimation: there are strong covariances between particular parameters, shown by the darker colors

D

Convergence and estimation uncertainty of the sensitivity analysis

This appendix presents the convergence and estimation uncertainty of the sensitivity indices presented in Chapter 5.

Chapter 5 presents results based on a global sensitivity analysis. Such results may only be trusted if convergence is achieved and put in perspective with an estimation of the uncertainty of each sensitivity index (see 3.4.3.3). Figures D.1, D.2 and D.3 show the convergence of all 6 estimated indices, with a bootstrap to show uncertainty of the estimation.

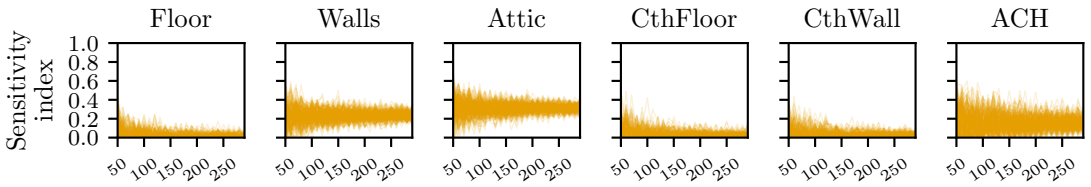


Figure D.1 – Convergence and estimation uncertainty of the sensitivity indices in case of model $T_w T_i\ R_o R_i\ A_w$

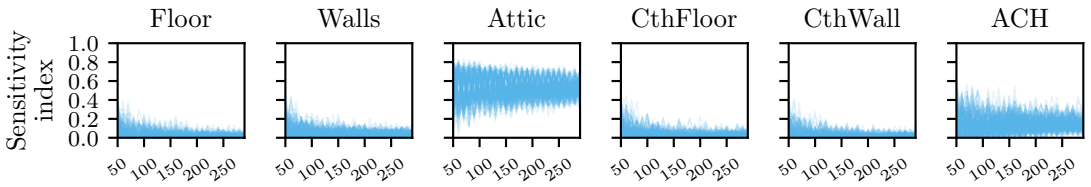


Figure D.2 – Convergence and estimation uncertainty of the sensitivity indices in case of model $T_w T_i\ R_o R_i\ c_v$

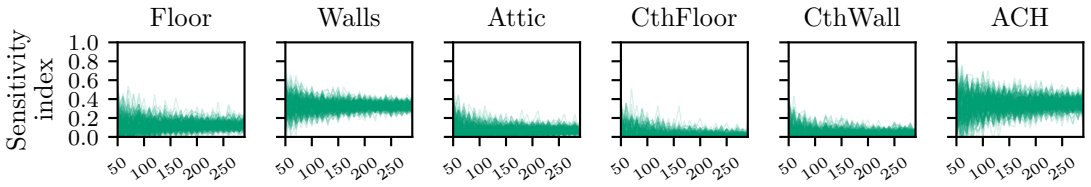


Figure D.3 – Convergence and estimation uncertainty of the sensitivity indices in case of model $T_w T_i\ R_o R_i\ c_v$

Identifiability of the thermal performance of a building envelope from poorly informative data

Résumé

Estimer la performance thermique d'une enveloppe de bâtiment à partir de mesures in-situ non intrusives présente l'intérêt d'être représentatif des conditions réelles de fonctionnement du bâtiment mais induit une qualité des données mesurées faible. Cela peut biaiser l'estimation voire la rendre impossible. Cette thèse s'attache à étudier la faisabilité de la caractérisation thermique d'une enveloppe à partir de données peu informatives.

Si l'identifiabilité théorique et pratique peut être aisément vérifiée en approches fréquentiste ou Bayésienne, elle ne constitue pas une condition suffisante à l'interprétation physique des paramètres estimés.

Une méthodologie originale est donc proposée pour évaluer l'interprétation physique de modèles RC stochastiques. Un modèle de référence délivre des simulations thermiques qui servent à l'apprentissage du modèle étudié. L'originalité est de pouvoir quantifier l'influence des incertitudes de l'environnement du bâtiment sur la caractérisation par une analyse de sensibilité globale. La méthodologie est appliquée d'une part à l'étude de l'influence de la variabilité des conditions météorologiques et d'autre part à l'interprétabilité physique individuelle de paramètres estimés. Les résultats montrent que 11 jours suffisent à obtenir une bonne répétabilité et précision de l'estimation de la résistance thermique. Aller plus loin en identifiant les pertes thermiques par la ventilation ou par les parois donnant sur un local non chauffé n'est en revanche pas possible dans ces conditions.

Les travaux de cette thèse montrent donc d'encourageants résultats pour la caractérisation thermique globale d'une enveloppe de bâtiment et invitent à de futurs travaux vers la faisabilité de cette estimation en situation d'occupation des locaux mesurés.

Mots-clés : Caractérisation thermique; Problème inverse; Modèles RC stochastiques; Calibration fréquentiste et Bayésienne; Analyse de sensibilité globale; Identifiabilité; Interprétabilité physique

Abstract

Accurate thermal characterisation of a building envelope is necessary for efficient energy retrofits of the building sector. Characterisation based on in-situ non intrusive measurements deliver insight on the performance in actual operating conditions but provide poorly informative data, which may bias the estimation or make it impossible. This thesis proposes to tackle the feasibility of thermal performance estimation from poorly informative data.

The identifiability, i.e. existence and unicity of the estimations, in theory and in practice can be verified in both frequentist and Bayesian approaches to model calibration. However, even when following calibration good practices, identifiability does not necessarily imply physical interpretability.

An original methodology is therefore proposed to assess the physical interpretation of calibrated stochastic RC models. Its is based on a numerical model which simulation output serves a training data. The novelty is to quantify the influence of uncertainties in the building environment on the thermal characterisation through a global sensitivity analysis. It is applied to study on the one hand the impact of weather variability on the estimation and on the other hand assess the interpretability of parameter estimates individually. RC models trained from poorly informative data have been found to achieve with good repeatability a satisfactory estimation of the overall thermal resistance. Further decomposition of heat losses through ventilation or heat losses towards an unheated crawl space has however been found non identifiable in these conditions.

The results show therefore promising opportunities for an overall thermal characterisation of the building envelope and engage for further research towards such estimation from data collected under occupancy.

Keywords : Thermal performance estimation; Inverse problem; Stochastic RC models; Frequentist and Bayesian calibration; Global sensitivity analysis; Identifiability; Physical interpretation