



HAL
open science

Towards automated stochastic grey-box model calibration for Heat Transfer Coefficient inference

Gaelle Faure, Sarah Juricic, Mickael Rabouille, Simon Rouchier, Arnaud Challansonnex, Arnaud JAY

► To cite this version:

Gaelle Faure, Sarah Juricic, Mickael Rabouille, Simon Rouchier, Arnaud Challansonnex, et al.. Towards automated stochastic grey-box model calibration for Heat Transfer Coefficient inference. 13th nordic symposium on building physics, Jun 2023, Aalborg, Denmark. pp.012048, <10.1088/1742-6596/2654/1/012048>. <cea-04653216>

HAL Id: cea-04653216

<https://cea.hal.science/cea-04653216v1>

Submitted on 18 Jul 2024

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

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



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

Towards automated stochastic Grey-Box Model Calibration for Heat Transfer Coefficient Inference

Gaëlle Faure¹, Sarah Juricic², Mickaël Rabouille¹, Simon Rouchier³,
Arnaud Challansonnex³ and Arnaud Jay⁴

¹ Institut National de l'Énergie Solaire (INES), Plateforme Formation & Évaluation, Savoie Technolac Bât. Helios, Le Bourget-du-Lac, France

² Scientific and Technical Center for Building (CSTB) 84 Avenue Jean Jaurès, Champs sur Marne, France

³ LOCIE UMR 5271 University Savoie Mont-Blanc CNRS, Savoie Technolac Bât. Helios, Le Bourget-du-Lac, France

⁴Univ. Grenoble Alpes, CEA, Liten, INES, 73375 Le Bourget-du-Lac, France

E-mail: sarah.juricic@cstb.fr

Abstract. An accurate on-site measurement of the intrinsic thermal performance of the building envelope is a strong leverage to promote quality management and energy performance contracting. The building sector expects on-site measurements of the Heat Transfer Coefficient to be as short as possible and highly accurate. Perturbation measurement methods analysed by RC models have been shown to be promising with accurate results within 2 days. However, stochastic RC model calibration is highly dependent on expert knowledge. In particular, model initialisation jeopardises the calibration success if not carefully chosen a priori and all the more so with second or higher order models. This paper proposes and assesses a novel automated initialisation procedure for RC models. After having brought to light that the capacitance initialisation is critical to the calibration success, the paper shows how first-order models are almost always highly identifiable. The first-order model estimated parameters then serve as initial values for all second and higher order models. This novel procedure is in 78 % of the 156 cases tested better than or similar to expert knowledge initialisation. The results are very encouraging and suggest a successful integration in a global RC model calibration and selection workflow for a fully automated process.

1. Introduction

According to the latest IPCC report [1], the 2020–2030 decade is critical in removing the constraints to achieve highly efficient buildings. To that effect, accurately measuring the actual thermal performance, for example the Heat Transfer Coefficient (HTC) of the building envelope could be beneficial to the building sector. It would promote quality management and favour the implementation of energy performance contracting in a Measurement and Verification approach for building retrofits.

An on-site measurement of the thermal performance is expected by the building sector to last less than a few days and to be highly accurate [2]. Existing methods rely on on-site measurements and use an appropriate model to infer an estimation of the thermal performance, such as the HTC. Such models include steady-state models, linear dynamics input-output models (i.e. ARX) and grey-box models [3]. The shortest measurement protocols are based on perturbation



Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

methods such as [4] and [5]. Grey-box models are then preferred to exploit the collected data, either with first-order models [5] or RC model selection and calibration [4].

A major constraint of RC model calibration is that they need fine tuning and expert knowledge [3, 6]. In particular, for the RC model selection procedure to succeed at fitting a given dataset, the models in the set need to be robust and the number of models need to be sufficient to cover a wide variety of envelope dynamics and insulation levels [6]. However, the larger the model set, the more tedious it becomes to calibrate all models and then proceed to model selection. With larger model sets, the calibration process encounters a two-sided curse: model calibration may fail because the model in question is not appropriate to the data or it may fail because the initial parameter values were too far from the actual values and the algorithm fell in a local maximum. The result is identical but the latter can be avoided by proper parameter initialisation.

Brastein et al [7] show indeed how a randomised initial guess of the model parameters is a way of assessing the identifiability of a model with respect to a certain dataset. Each initial guess is followed by a parameter estimation. An identifiable parameter is unambiguously identified regardless of its initial value. This model validation being purely visual, it cannot be automatically applied. Rivera et al [8] used ten different RC models and found that parameter initialisation was critical to the estimation problem, depending on whether the model was identifiable or not. Initial values close to the actual values tended to fasten the estimation. To circumvent this issue, the authors chose initial values within a 50–150% range of an approximated physical value. This solution however relies again on expert knowledge and a prior calculation.

For an HTC measurement method to be widely deployable, automated model calibration and selection becomes a priority. To this end, this paper proposes to tackle the issue of parameter initialisation and moves a step closer towards fully automated RC model selection and validation.

The work presented is part of the Sereine project, that aims at developing and assessing the practical limits of a 24-hour test for the estimation of the Heat Transfer Coefficient of a building envelope. The method relies on a 24- to 48-hour measurement protocol based on a binary heating signal [10]. The collected data are exploited by a set of 26 RC models that range from 1st- to 6th-order models as well as up to 3 boundary conditions.

This paper focuses in particular on the capacitance initialisation as it was found highly critical to the calibration success. It establishes first the relative advantage of having expert knowledge-based parameter initialisation over default values. Then, it shows why and how the estimated parameters of a first-order model are favourably used as initial values for higher-order RC models.

Nomenclature

BFGS Broyden-Fletcher-Goldfarb-Shanno optimisation algorithm

HTC Heat Transfer Coefficient (W/K): overall or whole-building heat transfer of the building envelope towards exterior as defined in the ISO 13789 standard [9]

PRBS Pseudo-Random Binary Signal

2. Methodology

2.1. RC model calibration: 3 strategies for parameter initialisation

As other perturbation methods, the Sereine method is based on the exploitation of the measurement data by a set of RC models. The parameter estimation procedure uses a BFGS algorithm [11]. As any minimisation algorithm, each parameter and state can be initialised as to favour convergence. In literature, parameter initialisation is based either on default values or

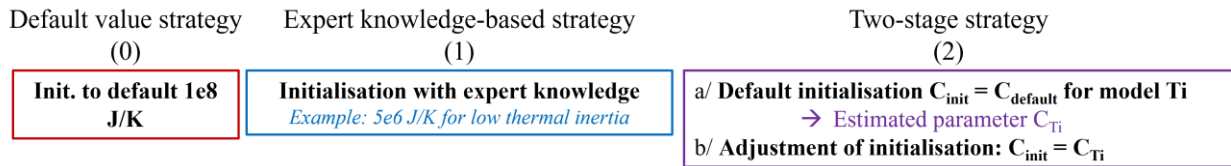


Figure 1. Three strategies for parameter initialisation: strategy (0) serves as control, strategy (1) is expert-knowledge based and strategy (2) uses the estimated parameter values from the calibration of first-order model t_i to initialise all remaining models.

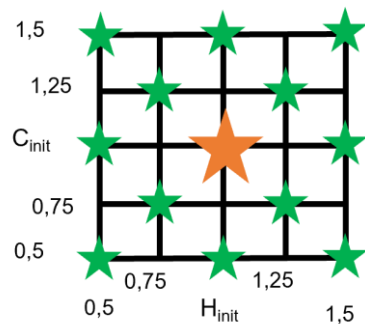


Figure 2. Stability assessment by fractional factorial design: 13 parameter initialisation chosen in total, including the couple prior values (H_{init} , C_{init})

on expert knowledge [7,8], hereafter respectively called strategy (0) and strategy (1), such as in Figure 1.

The novelty brought by this work, strategy (2) in Figure 1, is to use the estimated parameter values of the first order RC model t_i and use this value as an initial values for the calibration of all higher-order models. This strategy is hereafter called the two-stage strategy (2).

Second and higher order models have several capacitance and resistance parameters that need initialisation and a single initial value is concededly insufficient. A ratio is therefore used to split the single initial values C_{init} or H_{init} to all capacitances and conductances.

Let us again underline that this paper does not assess initialisation of conductances but focuses on capacitance initialisation. The latter has indeed been found more critical to the success of the calibration process.

As for the RC model denomination, visible in the results section, they imply the ad hoc RC model structure: each t - represents a temperature node; node t_i for the indoor air, t_w for a wall, t_m for an internal thermal mass and t_b for a boundary temperature. Each node has a capacitance and is separated of other nodes by a resistance.

2.2. Assessing the impact of the two-stage initialiation procedure

The final objective of the paper is to assess the performance of the two-stage parameter initialisation strategy (2) in terms of stability of the estimation. By stability is meant that the calibration algorithm yields similar values regardless of the parameter initialisation. It is somehow a measure of the practical identifiability of the model, except only the maximum likelihood values are looked at.

First, the initial overall capacitance C_{init} is adjusted according to the two-stage strategy described in Section 2.1. Then, the actual parameter initial values are shuffled around the C_{init} according to a fractional factorial design, as illustrated in Figure 2. Ranging 50–150 % around C_{init} and H_{init} , 13 parameter combinations are tested, resulting in 13 model calibrations and therefore 13 parameter estimation.

A stable and highly identifiable model will show equal estimations for all capacitance and

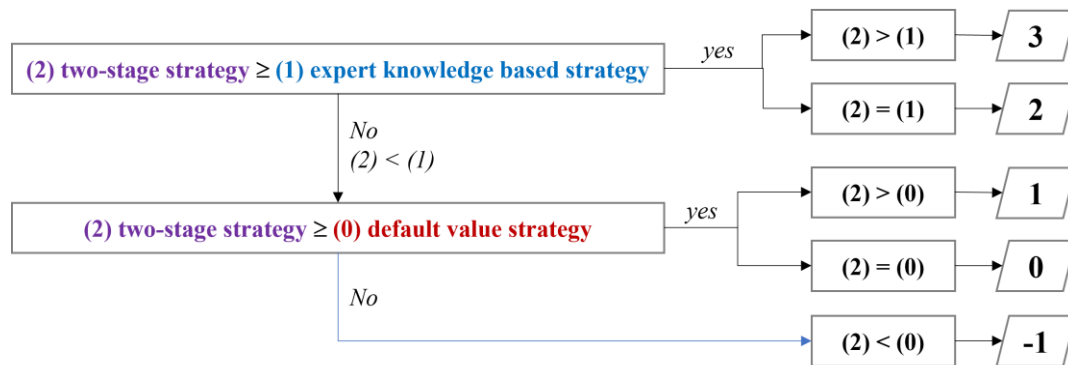


Figure 3. Flowchart of the performance assessment of the two-stage adjustment strategy (2)

resistance parameters regardless of the initialisation, e.g. regardless of the initial values from the factorial design. Some models however will show a higher sensitivity to the initial value. Among the 13 parameter estimations, some model calibrations will therefore produce more or less estimation outliers, depending on the initialisation strategy. A strategy is therefore deemed:

- more stable than another if lower number of outliers;
- equally stable than another if identical number of outliers;
- less stable than another if higher number of outliers.

As illustrated in Figure 3, to assess the performance of the two-stage initialisation strategy (2), the grading approach compares sequentially the respective performances of strategies (2) compared to (1) and then (2) compared to (0). The grading approach is applied on all RC models and on all 6 datasets, reaching a total of 156 grades.

The same outlier counting grading principle is applied to the first strategy comparison between the default and the expert knowledge-based strategies, except the results are illustrated by the self-explanatory +, = or -- symbols.

2.3. Case study: description of the 6 datasets

To assess the 3 initialisation strategies, the convergence of parameter estimations is analysed over 6 different datasets. The datasets come from measurement campaigns in 4 houses as to cover a variety of insulation levels and thermal inertia. Three different heating signals were applied as to assess the convergence of the estimation procedure regardless of the measurement procedure: two are pseudo-random binary signals from [10] and the third is an ISABELE signal [4]. All datasets can be considered highly informative and favour practical identifiability.

- House IOB of the INCAS platform: 2 datasets based on both PSA. This house is highly insulated ($HTC = 64 \text{ W/K}$) and has low thermal inertia;
- House IMA of the INCAS platform: repeated PSA2 tests. This house is highly insulated ($HTC = 94 \text{ W/K}$) and has an insulating brick structure. Its thermal inertia is medium;
- One-storey house in Mozé ($HTC_{target} = 141 \text{ W/K}$) and two-storey house in Brest ($HTC_{target} = 124 \text{ W/K}$) are both recently built and are insulated on the inside. They have higher thermal inertia than the houses of the INCAS platform. The Brest house has a significantly higher thermal inertia. The measurement protocol in these houses was the ISABELE procedure [4].

The data itself is collected via the Sereine measurement kit. Basically, it relies on the measurement of an equivalent outdoor temperature to serve as boundary condition and relies on heating power and indoor temperature sensors. Fans ensure homogeneity of the indoor temperatures. A detailed description of the Sereine kit can be found in [10].

3. Results

The first tested strategy was to assess whether expert knowledge enhanced the convergence and interpretability of the estimated parameters through enhanced estimation stability. Results are reported in Table 1.

Overall, 98% of all tests show identical or enhanced stability, among which 83% have a more stable convergence (= in Table 1). It can be concluded that adapting the prior initialisation to a physical expert knowledge-based value is beneficial. As for the only 3 less stable results, reported as – in Table 1, they might be explained by (1) either a random failure of the algorithm, since these failures are not consistent with the neighboring models (most probable hypothesis),

Table 1. Stability assessment when expert knowledge is used for parameter initialisation (strategy (1)): – when less stable, = when behaviour identical and + when enhanced stability

Model order	Model name	IOB PSA1	IOB PSA2	IMA PSA2	IMA PSA2	Mozé ISA	Brest ISA	
1	ti	=	=	=	=	=	=	
	tw	=	=	=	=	+	=	
	twti	+	+	+	+	+	+	
	2	twbti	+	+	+	+	+	+
		twtm	–	+	+	+	=	+
		twtw2	+	=	+	+	+	+
		twtitm	+	+	+	+	+	=
	3	twtw2ti	+	+	+	+	+	+
		twtw2bti	+	+	+	+	+	+
twtw2tm		+	+	+	+	+	+	
2	twtb	=	–	=	=	+	+	
	twtbti	+	+	+	+	+	–	
3	twtbtm	+	+	+	+	+	+	
	twtw2tb	+	+	+	+	+	+	
4	twtbtitm	+	+	+	+	+	+	
	twtw2tbtm	+	+	+	+	+	+	
5	twtw2tbtitm	+	+	+	+	+	+	
3	twtb+tb2	+	+	+	+	+	=	
	twtb+tb2ti	+	+	+	+	+	+	
4	twtb+tb2tm	+	+	+	=	+	=	
	twtw2tb+tb2	+	+	+	+	+	=	
5	twtb+tb2titm	+	+	+	+	+	=	
	twtw2tb+tb2ti	+	+	+	+	+	+	
6	twtw2tb+tb2tm	+	+	+	+	+	+	
	twtw2tb+tb2titm	+	+	+	+	+	=	

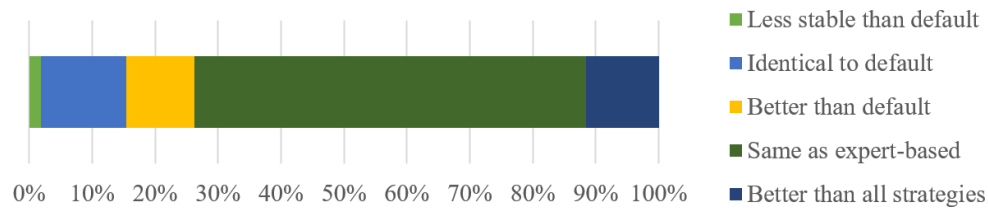


Figure 4. Strategy (2) overall assessment: 74 % of all 156 tests show identical or better stability than expert-based initialisation

or (2) an unfit model for the given data, probably poorly identifiable.

Interestingly, the Brest dataset gives more often identical behaviours (output = in Table 1) between default (10^8 W/K) and expert-knowledge (1.5×10^7 W/K) initialisation values. This is very likely due to the fact that the estimated final capacitance values are fairly close to the expert initial value and to the default value. There is therefore in the case of Brest almost no practical identifiability issue. This also suggests that default initialisation of capacitance parameters is more likely to fail when it differs by more than a factor of 10.

Noteworthy is the behaviour of the first order models. Whether the prior value of the capacitance is given by default or priorly estimated by expert-knowledge, the algorithm almost certainly converges towards the same final values. This indicates that first order models are highly identifiable with the given datasets.

That the first order models have a high practical identifiability regardless of the measurement protocol and regardless of the initialisation values, motivates the two-stage initialisation procedure of strategy (2). Model t_i estimated capacitance and conductance are used as output of the first stage of the initialisation procedure of strategy (2).

Figure 4 shows how the t_i adjusted strategy influences the stability of the parameter estimation:

- stability is enhanced or similar to using the default initialisation in 98 % of all tests;
- stability is better or similar to expert-knowledge based initialisation in 74 % of all tests, among which 11.5 % show a better performance with strategy (2) than strategy (1);
- inversely, 15.5 % of all tests perform equally or worse than the default strategy.

Table 2 shows the results in detail. There is a striking consistency with results in Table 1, especially regarding less or equally stable outcomes: tests where the outcome is equally or less stable than the default strategy were previously already equally or less stable. This tends to support the hypothesis that the issue is model related, meaning that the model calibrated is unfit for explaining the data.

Similarly to Table 1, first-order models do not gain in stability with strategy (2) because there are highly identifiable with the given datasets. As both Figure 4 and Table 2 suggest, the higher-order models gain from the implementation of strategy (2).

4. Discussion

This paper proposed an assessment of two parameter initialisation strategies: (1) either expert knowledge-based (2) or a two-stage initialisation where a first-order model is first calibrated and its capacitance estimation used for initialisation of all other models.

The results were found to be consistent with what the literature suggested. When models are highly identifiable, here the first-order models, parameter initialisation is not an issue and parameter estimation is highly stable. This means that regardless of the initial values, the final

Table 2. Detailed stability assessment of strategy (2): **-1** less stable than default strategy, **0** stability identical to default strategy, **1** better than with default **2** same stability as expert knowledge based strategy **3** better stability than all other strategies

Model order	Model name	IOB 1	IOB 2	IMA 1	IMA 2	Mozé	Brest
1	ti	0	0	0	0	0	0
	tw	0	0	0	0	2	0
	twti	2	2	2	2	2	2
	twbti	2	2	2	2	2	2
	twtm	1	2	2	3	0	2
	twtw2	0	-1	2	0	3	2
	twtitm	2	2	2	3	2	0
	twtw2ti	2	2	2	2	3	2
	twtw2bti	2	2	2	2	2	1
twtw2tm	2	1	2	3	2	1	
2	twtb	0	2	0	0	2	2
	twtbti	2	1	2	2	2	-1
3	twtbtm	2	2	2	3	2	1
	twtw2tb	2	2	2	2	2	3
	twtbtitm	2	2	2	2	2	2
4	twtw2tbti	2	2	2	2	2	2
	twtw2tbtm	2	2	2	2	2	2
5	twtw2tbtitm	2	2	1	2	1	2
3	twtb1tb2	3	2	2	2	1	0
	twtb1tb2ti	3	2	1	2	3	2
4	twtb1tb2tm	2	2	3	2	3	0
	twtw2tb1tb2	2	2	1	2	1	-1
	twtb1tb2titm	1	1	3	2	2	0
5	twtw2tb1tb2ti	2	2	3	1	3	2
	twtw2tb1tb2tm	2	2	2	3	2	3
6	twtw2tb1tb2titm	3	1	1	2	2	2

estimation are almost always identical. In addition, as suggested in Rivera et al [8], expert-knowledge based initialisation favours indeed a stable parameter estimation.

This work went a step further by proposing an automated capacitance initialisation by a two-step methodology: a first-order model is initialised by a default value, calibrated and its parameter estimation is used to initialise all remaining models. The results showed that this methodology performs in overall better or equally to the expert-based strategy. It means therefore that expert knowledge is no longer needed for stable parameter estimation. The new two-stage initialisation can therefore be considered as a practical step towards automated RC model calibration.

We expect the two-stage initialisation procedure to be useful not only for RC model calibration from Sereine data but actually from any data. The procedure is neither data dependent nor sensor dependent. Whether it works as accurately as shown in this paper on less informative datasets is admittedly not established. However, successful model calibration from lowly informative data cannot be expected in general. If the quality of the data is unknown, the two-stage initialisation can still be helpful in the sense that it rules out, with high confidence,

the possibility that the issue is algorithm-related.

Although not properly tested in this work, the calibration algorithm used in project Sereine makes it possible to change the default initialisation of the first-order model. It might be overkill in most cases given that the first-order models were found highly identifiable. Yet, in the authors experience, setting lower default values comes in useful when applied to small volume dwellings or flats.

As for the failed parameter estimations, although a first analysis suggested that some models remain poorly or non-identifiable with one or more datasets, an additional analysis would help uncover the specific reasons for the enhanced initialisation to have failed. This could lead to new improvements of the calibration code.

Finally, this work studied how initialisation strategies influences the stability of parameter estimation but tackled neither the issue of physical interpretation of the estimations nor the outcome of BIC or likelihood ratio model selection. Some high-order models may show a better stability, but if they are never interpretable or never relevant regarding the data, it makes no sense to use them in an automated model calibration and selection procedure.

5. Conclusion and research prospects

The objective of the work was to tackle the issue of parameter initialisation as to move towards automatic model calibration and selection. Two strategies were tested: (1) expert-based parameter initialisation and (2) a two-stage initialisation methodology where a first-order model estimation serves as parameter initialisation for all other RC models.

The two-stage initialisation method showed in almost all cases an equally or better stability of the parameter estimation. This work showed therefore a path which avoids relying on expert-knowledge and marks a step towards automation of RC model calibration.

Further research includes at least a deeper analysis of the few failed estimations as well as a practical identifiability and interpretability analysis of the higher-order models to simplify the model selection process to the only relevant RC models.

References

- [1] Cabeza L F, Bai Q, Bertoldi P, Kihila JM, Lucena A F P, Mata É, Mirasgedis S, Novikova A and Saheb Y 2022 *IPCC, 2022: Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (Buildings)* ed P R Shukla, J Skea, R Slade, A Al Khourdajie, R van Diemen, D McCollum, M Pathak, S Some, P Vyas, R Fradera, M Belkacemi, A Hasija, G Lisboa, S Luz and J Malley (Cambridge University Press, Cambridge, UK and New York, NY, USA) p995
- [2] Deb C, Gelder L V, Spiekman M, Pandraud G, Jack R and Fitton R 2021 *Renew. Sustain. Energy Rev.* **144** 111008
- [3] Madsen H, Bacher P, Bauwens G, Deconinck A H, Reynders G, Roels S, Himpe E and Lethé G 2016 *IEA EBC Annex 58 – Reliable building energy performance characterisation based on full scale dynamic measurements. In Report of Subtask 3, Part 2* Leuven Belgium
- [4] Thebault S and Bouchié R 2018 *Energy Build.* **178** 182
- [5] Alzetto F, Pandraud G, Fitton R, Heusler I and Sinnesbichler H 2018 *Energy Build.* **174** 124-133
- [6] Baasch G, Westermann P and Evins R 2021 *Energy Build.* **241** 110889
- [7] Brastein O M, Lie B, Pfeiffer C F and Skeie N-O 2019 *Model. Ident. Control* **40** 213
- [8] Rivera A K, Sanchez J and Chen Austin M 2022 *Front. Built Environ.* **8**
- [9] ISO 13789 2017 *Thermal performance of buildings - Transmission and ventilation heat transfer coefficients - Calculation method*
- [10] Juricic S, Goffart J, Rouchier S, Jay A and Oberlé P 2022 *Energy Build.* **269** 112211
- [11] Raillon L, Rouchier S and Juricic S 2019 *In: Congrès français de thermique* (Nantes)