Tuning Performance of Grey-Box Models for Thermal Building Applications

Savvas Panagi*†, Lazaros Aresti†, Paul Christodoulides‡, Chrysovalantis Spanias*, and Petros Aristidou†
* Distribution System Operator, Electricity Authority of Cyprus, Nicosia, Cyprus.
† Dept. of Electrical & Computer Engineering, & Informatics, Cyprus University of Technology, Limassol, Cyprus
‡ Faculty of Engineering and Technology, Cyprus University of Technology, Limassol, Cyprus
Emails: {cspanias}@eac.com.cy, {savvas.panagi, paul.christodoulides, petros.aristidou}@cut.ac.cy, lazaros.aresti@edu.cut.ac.cy

***Abstract*—The electrification of heating and cooling systems transforms energy consumption patterns, creating challenges and opportunities for power system operation. To this end, RC (or grey-box) models have emerged as promising modeling structures for forecasting the thermal needs of buildings. This paper presents a tuning algorithm and performance evaluation of the grey-box model using realistic measurement data. It provides insights into the sensitivity of the learning process to sub-optimal solutions and the computational burden. Four different grey-box structures with varying complexity are evaluated. The proposed methodology is critical in integrating heating and cooling systems into future power systems. The results reveal that the 4R3C model achieves the highest accuracy among the evaluated structures; however, it requires significantly higher computational time in certain scenarios. Moreover, the physical representation of the building through this training structure can be challenging, as the optimization process does not consistently converge to a unique set of parameter values, indicating the presence of multiple local optima. The tuning framework is provided as an open-source modeling tool, aiming to support further research on grey-box model.**

***Index Terms*—optimization, buildings, thermal modeling, grey-box models**

I. Introduction

A. Motivation

The urgent need to address climate change is accelerating the transition towards clean energy technologies, particularly in the heating sector. Conventional heating systems, which are still largely dependent on fossil fuels, are increasingly being replaced by electricity-based alternatives such as heat pumps. In the European Union, heating and cooling account for approximately 50% of total energy consumption, with over 70% still relying on fossil fuels, while around 80% of residential energy use is dedicated to space and water heating [1], [2]. As a result, there is a growing need for accurate and adaptable modeling approaches capable of capturing the thermal behavior of buildings and predicting their future heating and cooling demands.

B. Building Thermal Needs

A notable contribution towards the modeling of thermal needs and demand-side flexibility has been carried out in the context of the AmBIENCe project [3], which explores the activation of energy-efficient services in buildings via innovative business models and control mechanisms for active building energy performance contracting. A significant challenge in this domain lies in the limited availability of automated methodologies for generating detailed thermodynamic models and the lack of representative building stock data [4], [5]. Projects such as TABULA [6], Hotmaps [7], and BuiltHub [8] have been developed to address the challenge of limited building stock data by providing comprehensive, open-access datasets. These datasets support various applications, from national energy planning to local building-level simulations, and serve as foundational elements for developing and testing energy management methodologies.

Building thermal modeling methodologies is broadly categorized into three high-level approaches: i) white-box, ii) grey-box, and iii) black-box models [4], each with distinct advantages and limitations. White-box models, based on detailed physical principles, are primarily utilized by simulation tools such as TRNSYS [9], EnergyPlus [10], and IDA ICE [11] to perform comprehensive multizone simulations of building energy flows. These models require extensive inputs, including precise specifications of building envelope materials, and detailed definitions of construction elements (e.g., windows, doors, roofs) considering their area and orientation [12]. While white-box models offer high accuracy due to their detailed physical foundations, they need a significant amount of data and computational resources, making them complex, building-specific, and time-intensive to develop [5], [13]. Moreover, they lack adaptability, as generating a white-box model typically requires deep knowledge of the building's physical structure, making it impractical to derive such models for each building.

Black-box models leverage large datasets collected from a representative sample of buildings to infer thermal demand profiles. These models rely on statistical or machine learning (ML) techniques to capture complex patterns and dependencies among multiple input features, such as the approach used in [14]. A significant limitation of such models is their inability to describe the physical characteristics of a given building, such as construction materials, thermal mass, or specific usage and occupancy schedules. As a result, they provide generalized approximations that may not reflect the variability

This project has received funding from the EU's Horizon Europe Framework Programme (HORIZON) under the GA n. 101120278 - DENSE.

of real-world scenarios at the individual level. Moreover, the predictive performance is sensitive to training data availability and quality. In many practical cases, especially for residential sectors, the necessary volume of historical data is either limited or unavailable, thereby constraining the applicability of purely black-box solutions. However, this type of method is simple during the operational phase; it does not require prior knowledge of the building's physical structure and can be easily implemented using widely available ML libraries [4].

Grey-box models offer a hybrid modeling approach that combines physical principles with mathematical identification techniques to represent building thermal dynamics. These models analyze a building's thermal behavior by representing its components, such as walls, windows, and air volumes, as thermal resistances and capacitances, analogous to an electrical RC network [15]. This structure allows for the derivation of state-space equations governing temperature evolution, using a relatively small set of input parameters. Positioned between the entirely physical white-box and the black-box approaches, grey-box models utilize physical knowledge to define the model structure, while estimating parameters using measurement data. This duality enables a balance between interpretability, accuracy, and computational efficiency. Grey-box models are an ideal choice for applications that require mathematical modeling, such as model predictive control (MPC) structures [5], [13], [16]. The implementation of it will become easier in the near future, when smart meters are expected to be installed in almost all buildings. Compared to black-box models, grey-box methods require less training data and offer more physical insight. Compared to white-box models, they are simpler to implement, less data-intensive, and more adaptable to different building types [5].

However, grey-box models still present some drawbacks. The individual implementation requires tuning/identifying a new set of parameters for each building. Moreover, for forecasting, the model may be sensitive to the initial values, which are more observable in models with multiple nodes. Finally, the usage of the building or even the building parameters such as its insulation quality may change over time; this could lead to an inaccurate model, which may need to be adjusted/retuned [13].

C. Grey-Box Models

Numerous implementations have been proposed in the literature to identify effective grey-box modeling approaches for building thermal behavior. These studies typically focus on a series of fundamental questions, including: (i) the selection of an appropriate model structure, (ii) the required quantity of training data (duration, frequency), (iii) the origin of this data—either from measurements or simulations, (iv) assumptions about building occupancy during training, (v) the specification of physical bounds for parameters to ensure plausibility and interoperability, and (vi) the best practices for the training process.

The literature has not converged on a standard grey-box configuration that can be applied across different building types and use cases. For instance, Harb et al. [4] evaluated four distinct grey-box structures and concluded that the 4R2C model provided the best trade-off between accuracy and complexity. This model exhibited superior dynamic performance by capturing temperature variations and transient responses during system modulation. About the learning strategy, the GlobalSearch algorithm was adopted to increase the chances of finding the global optimum. During the learning process, if the optimal parameter hits the bounds, the simulation is repeated by relaxing the bound limits only for this parameter. Similar grey-box models have been investigated in [12], with the 6R2C model being considered the most suitable model.

Moreover, in [12], improvements related to occupancy heat gains and variable ventilation modeling were given. Another critical aspect of the study was the investigation of how different training data selection ranges influence forecast performance. Regarding the optimization procedure, an effort was made to guide the solution space by initializing physically meaningful values and applying appropriate bounds to each variable. However, no modeling or learning strategy was employed to ensure that the solution corresponds to a global optimum, meaning that the final results depend strongly on the solver behavior and the local convergence landscape. It is also worth noting that the 4R2C structure proposed in [12] significantly differs from the one presented in [4]. This is a common place when comparing all proposed structures.

In [13], a passive house equipped with a building-integrated photovoltaic facade is modeled using grey-box approaches. The only deviation from conventional models is the separate treatment of one exterior wall, which is thermally decoupled due to elevated temperatures caused by solar gains behind the BIPV structure. In [17], a multi-room grey-box model was developed using a second-order model representation, in which each room was split into distinct thermal zones to capture both short- and long-term temperature dynamics. In [5], a large-scale implementation of grey-box modeling was applied to a comprehensive set of 2,245 representative buildings, including residential and non-residential types. This practical application demonstrates the scalability and flexibility of grey-box models in capturing the dynamic thermal behavior of the entire European building stock.

In [16], the authors focused on evaluating the performance of a parameter estimation algorithm for grey-box building thermal models. Through extensive Monte Carlo simulations, they demonstrated that the identified thermal parameters R and C exhibit high sensitivity to the initial guess and the input dataset. As a result, the parameter estimates show significant variability, undermining their physical interpretability. A systematic methodology for identifying suitable grey-box models for building thermal dynamics is proposed in [18], emphasizing a forward selection strategy incrementally increasing model complexity based on likelihood-ratio tests, ensuring physical interpretability and statistical validation. In [19], the authors investigated the quality and robustness of grey-box building models as a function of the accuracy and type of input and observation signals, using detailed white-box simulations as a controlled reference.

D. Limitations

Despite the progress, grey-box models still face several limitations that hinder their wider adoption. A significant issue is the lack of transparency, particularly regarding parameter estimation procedures, solar gain treatment, and the absence of open-source implementations. Many works omit critical details, such as the tuning algorithm, initialization strategy, and the associated computational cost of the calibration process, which can be substantial depending on the model complexity. The treatment of solar irradiance and heating inputs is often inconsistent across different studies, with each model distributing gains differently across internal zones, the envelope, or internal contents.

Another key challenge lies in the selection and frequency of training data. Existing works vary significantly in the total duration and granularity of the datasets used for identification, and only a few explicitly discuss how often the model needs to be re-trained to remain accurate. These gaps affect both the reliability and generalizability of grey-box models. The most common model structures, data sources, and the quantity of training dataset are summarized in Table I.

E. Novelty and Contributions

This paper addresses these challenges through a grey-box tuning algorithm that balances accuracy and flexibility. Moreover, it provides a sensitivity analysis for all critical decisions. The objective is to deliver tunable models that are good enough to support predictive simulations, planning tools, and energy flexibility services.

The contributions of the paper are threefold:

- A systematic framework to model and forecast the thermal energy needs of individual buildings, based on validated and interpretable grey-box structures.
- To provide insights about the sensitivity of the models to different initial values, and to conduct a computational sensitivity analysis to assess how tuning complexity and parameter initialization affect model performance.
- An open-source implementation that enables stakeholders to efficiently use, adjust, and extend the proposed models for broader studies on electrification and grid integration.

Reference	Models (Best Model)	No Building - Training Data	Data Source
Berthou et al. [12]	2nd–3rd order (6R2C)	1 – 2 weeks and 2 weeks/month	White-box
Harb et al. [4]	1st–3rd order (4R2C)	3 – 26 days	Measurements
Fux et al. [13]	1st, 2rd, 4th order (6R4C)	1 - 12 days	Measurements
Andersen et al. [17]	2nd order	1 - 17 days	Measurements
Jankovic et al. [5]	2nd order	1 - 2 Months	White-box
Brastein et al. [16]	2nd order	1 - 8 Days	Measurements
Madsen et al. [20]	2nd order	1 - 5 Days	Measurements
Bacher et al. [18]	1st - 5th order (4R4C)	1 - 6 Days	Measurements
Reynders et al. [19]	1st - 5th order	2 - 6 Days	White-box

TABLE I: Overview of grey-box models, training durations, and data sources used in related studies.

The remainder of the paper is structured as follows. In Section II, the modeling framework and mathematical modeling of building thermodynamics are given. The tuning process is discussed in Section III. Section IV provides a brief explanation of the case study. Then, in Section V, the simulation results are discussed. Finally, conclusions are drawn in Section VI.

II. Thermal Modeling Framework

A. Electric analogy modeling

In grey-box modeling, the building is represented through a lumped-parameter thermal network inspired by analogies with electrical circuits. In such analogy, temperature corresponds to voltage, heat flow to electric current, thermal capacitance to electric capacitance (representing thermal mass), and thermal resistance to electrical resistance (representing insulation and heat transfer limitations).

The model structure is typically described using the $xRyC$ notation, where x and y denote the number of thermal resistances (R) and thermal capacitances (C), respectively. The number of states in the model defines its order, and it directly influences the trade-off between computational complexity and predictive accuracy. First-order, second-order, or higher-order models can be formulated depending on the desired fidelity and the availability of training data. These models are based on the following set of assumptions that users should be aware of when applying them in practice [4], [12]:

1) Thermal resistance and capacitance values (R, C) are assumed constant over time, representing average building behavior.
2) Building components (walls, windows, etc.) are considered to have uniform material properties.
3) Most of the grey-box models consider a single thermal zone per building model. Also, multi-thermal zone grey box models are proposed; however, they are out of the subject of this paper.
4) Convective heat transfer is treated as a linear process.
5) Occupancy profiles are assumed to be uniform and constant.
6) Air infiltration is assumed constant unless explicitly modeled.

Capturing the thermal response of a building requires knowledge of four main variables: solar irradiance (Q_{irrad}), outdoor air temperature (T_a), heating power input from the HVAC system (Q_h), and the resulting indoor air temperature (T_{in}). In practice, three of these inputs are sufficient to estimate the fourth. Depending on the application, the focus may be on forecasting indoor temperature based on external conditions and heating input or, conversely, estimating the required heating power to maintain thermal comfort (T_{in}).

B. Mathematical modeling

Building upon the electrical–thermal analogy (See Figure 6 in the Appendix for illustrative models), the thermal dynamics

of a building zone are formulated using the principle of energy conservation [17]. The heat balance at each thermal node can be expressed as:

$$C_i \frac{\mathrm{d}T_i}{\mathrm{d}t} = \sum Q_{\text{in}} - \sum Q_{\text{out}} + d\omega_i$$

where C_i [J/K] is the thermal capacitance of node i, T_i [°C] is the temperature at node i, and Q [W] denotes thermal power exchange due to either active sources (e.g., heating systems, solar gains) or passive heat transfer through conduction. The conductive heat transfer between two thermal nodes is typically modeled by:

$$Q_w = \frac{1}{R_w}(T_j - T_i), \tag{1}$$

where R_w [K/W] is the thermal resistance of element w, and T_i, T_j are the temperatures on either side of the surface. In practice, building thermal behavior is affected by various uncertainties, including unknown internal heat gains, occupant-driven variability, and measurement noise. To account for these effects, a disturbance term $d\omega_i$ [W] is added to the heat balance. This stochastic representation, as discussed in [20], helps capture:

- Model simplifications and neglected physical phenomena,
- Unmeasured or time-varying internal gains (e.g., from occupants or appliances),
- Sensor and measurement noise.

The full ordinary differential equations (ODEs) are derived from the selected RC network structure, where each node corresponds to a thermal element in the building (e.g., indoor air, wall mass). These equations constitute the foundation for simulation and parameter estimation in the grey-box modeling framework. The continuous-time dynamics can be expressed in compact form as:

$$\frac{d\mathbf{T}}{dt} = A(\theta) \cdot \mathbf{T}(t) + B(\theta) \cdot \mathbf{U}(t) \tag{2}$$

where $\mathbf{T}$ is the state vector and $\mathbf{U}$ is the input vector. A and B represent the matrices of parameters, that are dependent on the grey-box model structure. For instance, in the 1R1C model, the state, input, and parameters vectors are defined as:

$$\mathbf{T}(t) = \begin{bmatrix} T_{in}(t) \end{bmatrix}, \quad \mathbf{U}(t) = \begin{bmatrix} Q_h(t) & Q_s(t) & T_a(t) \end{bmatrix}^T \tag{3}$$

$$A(\theta) = \begin{bmatrix} -\frac{1}{R_{in,a} \cdot C_{in}} \end{bmatrix}, B(\theta) = \begin{bmatrix} \frac{1}{C_{in}} & \frac{A_{in}}{C_{in}} & \frac{1}{R_{in,a} \cdot C_{in}} \end{bmatrix} \tag{4}$$

To enable discrete-time simulation and optimization, the system is discretized using the forward Euler discretization:

$$\mathbf{T}(k+1) = A_d(\theta) \cdot \mathbf{T}(k) + B_d(\theta) \cdot \mathbf{U}(k) \tag{5}$$

with $A_d(\theta) = I + \Delta T \cdot A(\theta)$ and $B_d(\theta) = \Delta T \cdot B(\theta)$, where I is the identity matrix of appropriate dimension (matching the size of $A(\theta)$), and ΔT is the discrete time step used for the numerical simulation and control horizon. Appendix A presents a compact formulation of the ODEs used in all grey-box models, alongside a detailed description of their structural components and parameter definitions.

The heat input from solar irradiance and the heating system (Qh, Q_s) constitutes the most critical heat transfer affecting the model behavior. In the modeling structure presented in this study, the solar irradiance heat flux is expressed as $A_i Q_s$, and the heating system heat flux as $f_i Q_h$, where i denotes node i of the grey-box model. The heating system coefficients must satisfy the $\sum f_i = 1$, as the power inputs are distributed across the nodes of the modeling structure. The solar heat gain Q_s is defined as $Q_s = Irr$ [W/m^2] $\cdot B_w$ [m^2], where B_w is the effective window area and Irr is the measured solar irradiance. In most studies, the effective window area is considered to be approximately 60% of the measured window size [17]. However, in our study, B_w is treated as a tunable parameter, constrained only by logical bounds.

III. Tuning Methodology

A. Optimization Framework

The identification of the unknown physical parameters $\theta \in \Theta$ in the grey-box model is formulated as a nonlinear optimization problem because of bilinear equations. The objective is to minimize the deviation between the predicted indoor temperature $\hat{T}_{in}(k)$ and the corresponding measured values $T_{in}^{\text{meas}}(k)$ over a time horizon of N steps which may correspond from a few days up to some weeks:

$$\min_{\theta \in \Theta} \quad \frac{1}{N+1} \sum_{k=0}^{N} \left(T_{\text{in}}(k) - T_{\text{in}}^{\text{meas}}(k)\right)^2 \tag{6}$$

$$\text{s.t.} \quad \mathbf{T}(k+1) = A_d(\theta) \cdot \mathbf{T}(k) + B_d(\theta) \cdot \mathbf{U}(k), \quad \forall k \tag{7}$$

$$\theta_{min} \leq \theta \leq \theta_{max}, \quad \forall \theta \in \Theta, \quad \forall k \tag{8}$$

where:

- Θ is the vector of parameters to be tuned (e.g., thermal resistances and capacities),
- $[\theta_{min}, \theta_{max}]$ denotes the set of admissible physical bounds,
- $T_{in}(k) \in \mathbf{T}$ is the simulated temperature at time step k,
- $T_{in}^{\text{meas}}(k)$ is the measured indoor air temperature.

The θ limits were selected based on building physics to ensure the chosen model correlates directly with the modeled building's physical attributes.

B. Tuning Framework

The overall identification and validation process is illustrated in Fig. 1, which involves three main stages:

1) **Data Acquisition and Filtering:** Generate synthetic data through detailed simulation or gain on-site data. These raw data are preprocessed (testing and validation) and resampled to a common frequency interval.
2) **Parameter tuning:** Optimal parameters θ^* are found by solving the proposed optimization problem. In each iteration, the initial values are randomly selected within the bounds,

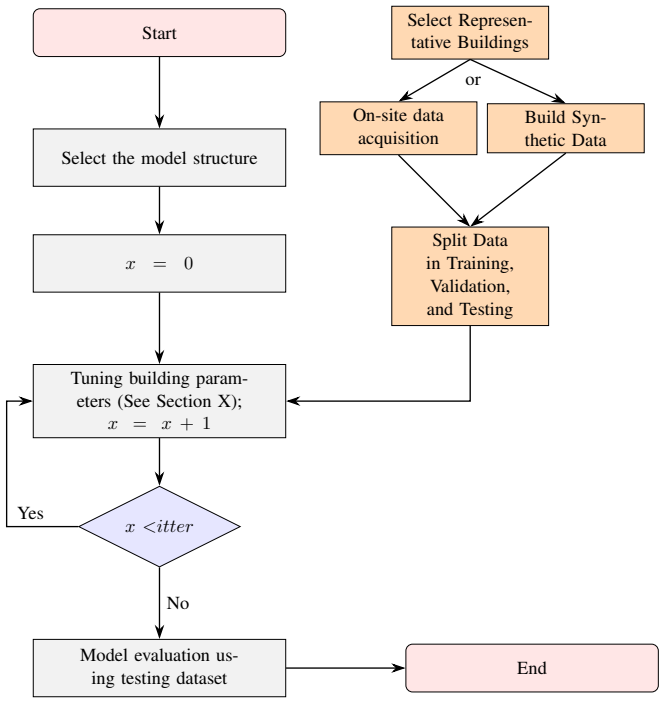


Fig. 1: Grey-Box Model Identification, Tuning and Validation Framework

3) **Validation:** The identified model is tested on unseen datasets to verify the performance. From the entire set of tests, the parameters of the best one, those with the highest accuracy, are selected.

Initially, the user selects the model structures for tuning. Then a repetitive tuning process is executed for a predefined number of iterations, denoted as $itter$. The model is initialized in each iteration with randomly selected parameter values within the specified bounds. After completing all iterations, the set of parameters that yielded the lowest mean squared error (MSE) in the training dataset is retained as the optimal configuration for that model.

C. Metric

Performance metric are used to compare the model's performance and assess its accuracy. The metric concerns the model's ability to predict indoor air temperature. This is quantified by the RMSE, which measures the disparity between simulated indoor temperatures and actual measurements as defined in equation (9).

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(T_{in}[t] - \hat{T}_{\text{in}}[t]\right)^2} \tag{9}$$

IV. Case Study Buildings and Input Data

Measurement data collected from four buildings with varying usage types were initially used to validate and assess the robustness of the proposed modeling framework under real data. To ensure a comprehensive evaluation, the buildings were selected to represent a range of structural characteristics, including building size (57–2100 m^2), and diverse climate conditions across different locations (Norway, Canada, United States). The measurement data used in this study are online available at Mendelay data [21]. A short description of the building characteristics is given in the Table. II.

A. Building

Case Study 1 (CS1) is a zero-emission single-family house, located on the NTNU campus in Trondheim, which is a 100 m² single-family house. Constructed with a wooden frame, the building is highly insulated, featuring 35–40 cm of mineral wool, and maintains a window-to-wall ratio of about 20%. The layout is divided into four zones: two bedrooms, a bathroom, and a living area.

Case Study 2 (CS2) is the first Net-Zero Energy Library in Canada, built in 2014. Situated in Varennes, 30 km northeast of Montreal within ASHRAE Climate Zone 6, this two-story building encompasses approximately 2100 m² of floor space. It features a sizable basement housing the mechanical room and a mezzanine dividing the two levels. The two floors are mainly open-plan areas designated for library use, with a few closed office spaces on the first floor. A 110.5 kWp photovoltaic array, integrated into the building's architecture, is mounted on the south-facing roof.

Case Studies 3 and 4 (CS3, CS4), named FLEXLAB, located at Lawrence Berkeley National Laboratory in Berkeley, California, are well-equipped experimental test centers. The facility comprises four testbeds, each with two identical rooms to enable side-by-side comparisons of a new technology versus a baseline scenario. The data was gathered from the testbed "XR" which includes two identical adjoining rooms, Room Cell A and Room Cell B, each simulating a small commercial office space with a floor area of 57 m² and featuring a large south-facing window.

TABLE II: Overview of case study buildings.

Case Study	Location	Type	Floor area	Available Data
1	Trondheim	Residential	100 m^2	22 Days
2	Canada	Library	2100 m^2	60 Days
3 & 4	California	Office	57 m^2	20 Days

B. Thermal Data

The heating requirements are relative to all zones, with the building's average air temperature calculated by weighting each zone's temperature against its volume. Moreover, each simulation study case obtained heat consumption, outdoor air temperature, and global solar irradiance on a horizontal surface. Obtaining solar irradiance data is challenging. To simplify model implementation, it is best to avoid on-site measurements and instead use global horizontal irradiance data from nearby meteorological stations. Bathrooms are excluded from the calculations due to the very high temperature fluctuations caused by the use of hot water.

V. Results and Discussions

The optimization problem was solved using the IPOPT solver, which is well-suited for large-scale nonlinear programming. The implementation used a MacBook Pro equipped with an M4 Pro Max chip, featuring a 12-core CPU and 36 GB of unified memory. The solver successfully converged to a feasible solution in almost all scenarios tested. A temporal resolution of 30 minutes is used for all studies. The modeling and optimization framework is implemented in an open-source Python-based toolkit, publicly available on GitHub.[1]

A. Tuning Performance Evaluation

To assess the accuracy of different models and the computational time, real-world data from case studies was utilized. The first 12 days of each dataset are used for model training, the next 5 days for validation, and the last 3 days are used for testing purposes. As the methodology suggests that repeated runs are required to ensure convergence of the optimization process, 30 iterations were chosen for each test case in this analysis. Figure 2 shows a comparison of measured and simulated indoor temperatures across all grey-box model structures and case studies for the training dataset. The parameters after the tuning are presented in Table IV. The performance metric, RMSE, is summarized in Table III, for training, validation, and testing phases.

The 1R1C model shows clear limitations in capturing indoor temperature dynamics in some test cases. While it performs reasonably well under stable conditions (e.g., case studies 1 and 2), it struggles in more dynamic cases (e.g., case studies 3 and 4), where it underestimates peak values and produces overly damped responses. This is due to its simple structure, which includes only one thermal resistance and one capacitance, capturing just a single dominant time constant. As a result, it cannot differentiate between fast internal gains and slower thermal effects through the building envelope, leading to poor alignment with measured data.

In contrast, the model's performance increases as the complexity of the model rises. The 3R2C model shows some improvement, which is mainly caused by the extra ventilation parameter, making the 3R2C model slightly better than the 2R2C. The 4R3C structure demonstrates the highest performance among all models evaluated during the training and validation phase. However, in the testing phase, it does not present significant improvement over the other models, and it was either worse in CS2 and CS4.

Therefore, we cannot clearly distinguish the territorial model, as each dataset exhibits different sensitivities under the various models. However, we can confidently conclude that the most stable model is the 3R2C. It consistently outperformed the others, most likely due to its higher complexity compared to the 1R1C and 2R2C models—allowing it to capture both fast and slow thermal dynamics—while still maintaining lower complexity than the 4R3C model. The latter's higher degrees of freedom make it more difficult to tune effectively.

[1]https://github.com/SPS-L/HeatPump-FlexModeling-Toolkit

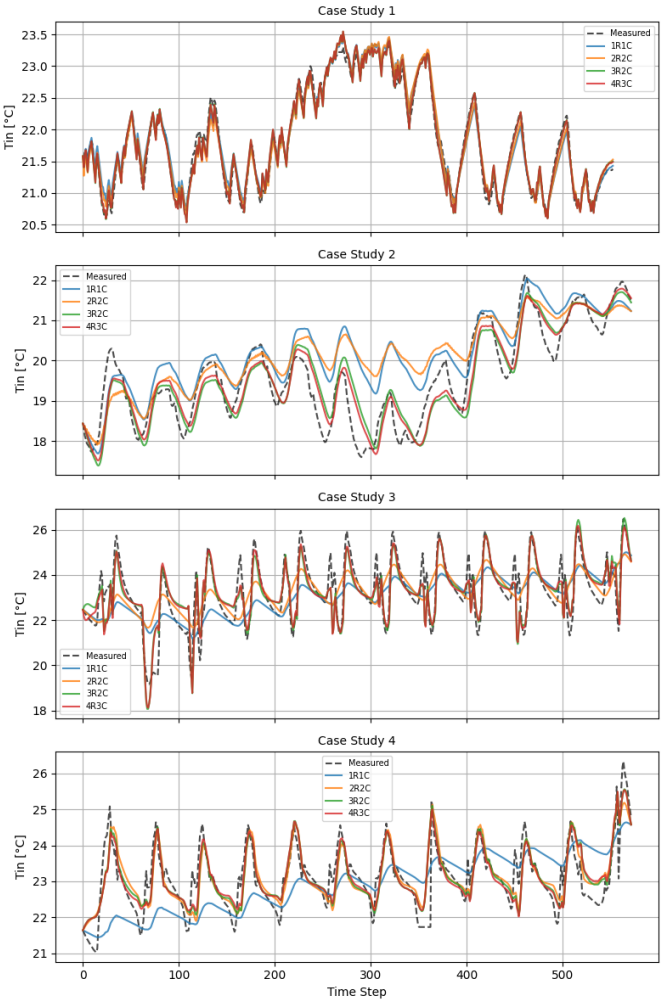


Fig. 2: Comparison of measured and simulated indoor temperatures for all grey-box models across the four case studies.

B. Computational Analysis

The boxplot results depicted in Figure 3 illustrate the expected trend in solve times across the different thermal models. The reported values correspond to the solve time for each iteration. The increasing model complexity (from 1R1C to 4R3C) generally leads to longer computation times. This is evident in most case studies, especially in CS3 and CS4, where the 4R3C model exhibits significantly higher solve times compared to simpler models, as expected due to its increased number of parameters and degrees of freedom.

High solving time is presented for some iterations (longer outliers), and this becomes more evident as the complexity of the model increases. This behavior can be attributed to the fact that some repetitions required significantly more iterations to converge because of poor initialization, which delays convergence, distorting the expected performance. As a result, the solve time for a specific iteration can appear higher than the median value, even though its typical convergence time remains lower.

TABLE III: RMSE metrics for different grey-box thermal models during training and testing phases. "CS" denotes the case study identifier used for each building scenario.

	Training RMSE [°C]				**Validation RMSE [°C]**				**Testing RMSE [°C]**			
	1R1C	2R2C	3R2C	4R3C	1R1C	2R2C	3R2C	4R3C	1R1C	2R2C	3R2C	4R3C
CS 1	0.15	0.14	0.10	0.10	0.28	0.22	0.18	0.18	0.36	0.32	0.33	0.31
CS 2	0.87	0.93	0.51	0.44	1.34	1.42	0.86	0.48	0.34	0.37	0.36	0.74
CS 3	1.16	1.12	0.65	0.62	1.06	0.97	0.86	0.76	0.67	0.58	0.49	0.48
CS 4	0.93	0.51	0.48	0.48	1.35	0.60	0.52	0.53	0.85	0.64	0.56	0.69

TABLE IV: Summary of results across varying grey-box models and buildings for the training dataset. Different buildings denoted with Case Study 1 / Case Study 2 / Case Study 3 / Case Study 4

Metric	**1R1C**	**2R2C**	**3R2C**	**4R3C**
$R_{int,in}$	-	-	-	0.016 / 0.025 / 0.0043 / 0.0031
$R_{in,e}$	-	0.048 / 0.0005 / 0.0016 / 0.023	0.0066 / 0.015 / 0.0064 / 0.0047	0.06 / 0.002 / 0.093 / 0.0027
$R_{in,a}$	0.013 / 0.00043 / 0.012 / 0.072	-	0.049 / 0.0015 / 0.039 / 0.035	0.069 / 0.002 / 1 / 0.88
$R_{e,a}$	-	0.0014 / 0.0005 / 0.001 / 0.0051	0.02 / 1 / 0.0037 / 1	0.0029 / 1 / 0.0097 / 0.019
C_{int}	-	-	-	4.2e+6 / 2.7e+8 / 1.2e+6 / 2.6e+6
C_{in}	2.3e+7 / 2.5e+9 / 5.1e+7 / 2.5e+7	6.3e+6 / 1.3e+9 / 1.1e+8 / 1.1e+6	4.5e+6 / 5e+8 / 1e+6 / 3.2e+6	4.2e+6 / 2.7e+8 / 1.2e+6 / 2.6e+6
C_e	-	1e+6 / 3.9e+7 / 4.5e+8 / 3.3e+8	2e+7 / 5e+8 / 3.4e+8 / 3.8e+7	2.7e+7 / 5.0e+8 / 1.2e+6 / 2.8e+6
A_{int}	-	-	-	0.00 / 0.00 / 1.60 / 2.35
A_{in}	3.47 / 630.00 / 5.35 / 3.06	0.39 / 226.05 / 14.51 / 0.37	0.93 / 136.12 / 0.74 / 1.02	0.86 / 76.07 / 1.02 / 1.60
A_e	-	12.37 / 0.00 / 0.00 / 8.55	1.25 / 0.00 / 13.83 / 2.45	3.59 / 0.00 / 3.87 / 0.00
$f_{h,int}$	-	-	-	0.17 / 0.81 / 0.54 / 0.35
$f_{h,in}$	1 / 1 / 1 / 1	0.26 / 0.00 / 1.00 / 0.13	0.25 / 0.32 / 0.36 / 0.43	0.22 / 0.13 / 0.46 / 0.49
$f_{h,e}$	-	0.74/1/0/0.87	0.75/0.68/0.64/0.57	0.61 / 0.06 / 0.00 / 0.16

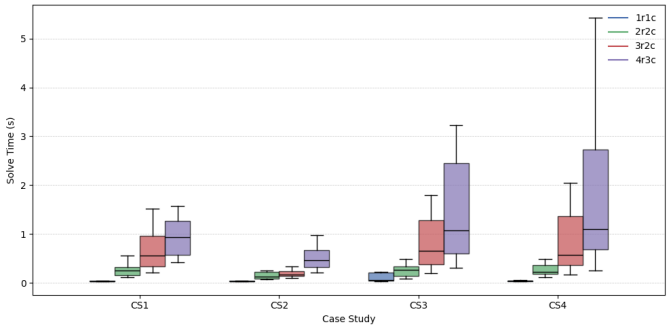


Fig. 3: Solve time for each thermal model across all case studies. Each boxplot represents the optimization time over 30 iterations per tuning process.

C. Tuning Sensitivity to Initial Parameter Values

Case study 1 is considered for this sensitivity analysis; however, the results were very similar for all test cases. As before, 17 days are used for model training and evaluation. The proposed framework executes the optimization procedure 100 times. We assess the impact of initialization variability on the final model performance. The RMSE and the corresponding optimized parameter values are recorded for every trial. The summary key results presented in Table V underscore the critical influence of initialization on the training of grey-box models.

The 1R1C model includes three degrees of freedom: $R_{\text{in,a}}$, C_{in}, and A_{in}. As shown in Fig. 4, the model exhibits low sensitivity to the initial values of R_{in}, which consistently converge to approximately 0.013. In contrast, A_{in} and C_{in} are identified as the most sensitive parameters. The C_{in} optimized value varying between $2e^7$ and $4e^8$, depending on the initialization. Despite this variation, the optimization process converges toward the same optimal solution across 43% of the trials. In the remaining 57% of the cases, slightly elevated RMSE values were observed. Therefore, even if the optimum solution cannot be reached, good model performance can still be gained.

In contrast, the 2R2C model introduces seven degrees of freedom, significantly increasing the nonlinearity and complexity of the optimization problem. As shown in Fig. 5, this results in a greater tendency to converge to suboptimal solutions, with multiple local minima observed across different initializations. Interestingly, the tuning variables do not converge to a single parameter value, indicating the difficulty in plausibly describing the building characteristics.

The 3R2C model, even though it is more complex than the previous two structures, as shown in Table V, nearly all initializations converge to a similar optimal value. As a result, the standard deviation of the RMSE values is reduced, indicating highly stable and consistent performance across all trials. A similar convergence trend was also observed for the 4R3C model. However, again, some significant variability in the optimal solution is presented. This probably occurs since the number of degrees of freedom is significantly high.

Finally, the number of times reaching the optimal solution is significantly reduced for 2R2C, 3R2C, and 4R3C models compared to the 1R1C model. These results emphasize the necessity of employing multiple training runs with randomized initializations to enhance parameter estimation's reliability and reduce the likelihood of entrapment in local minima.

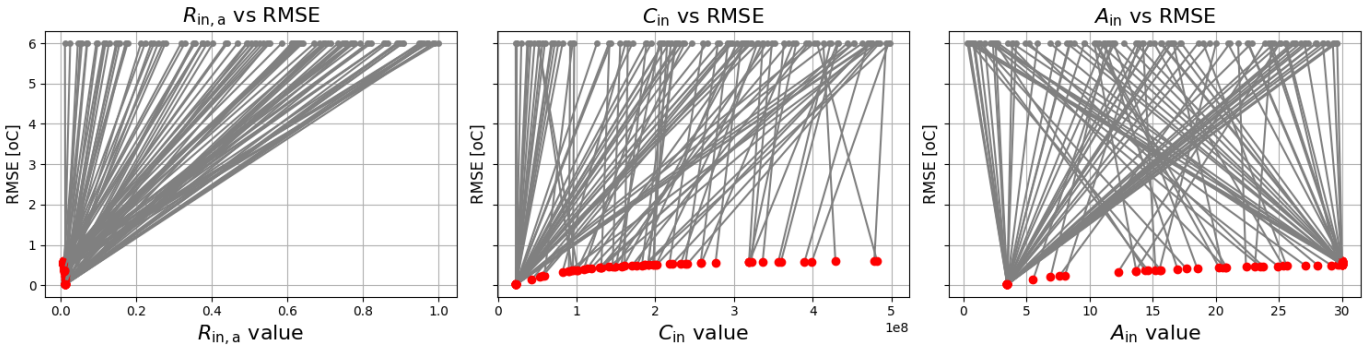


Fig. 4: Parameter trajectories and training RMSE for the 1R1C model. Each line connects the initial randomized value (grey marker) to the final optimized value (red marker).

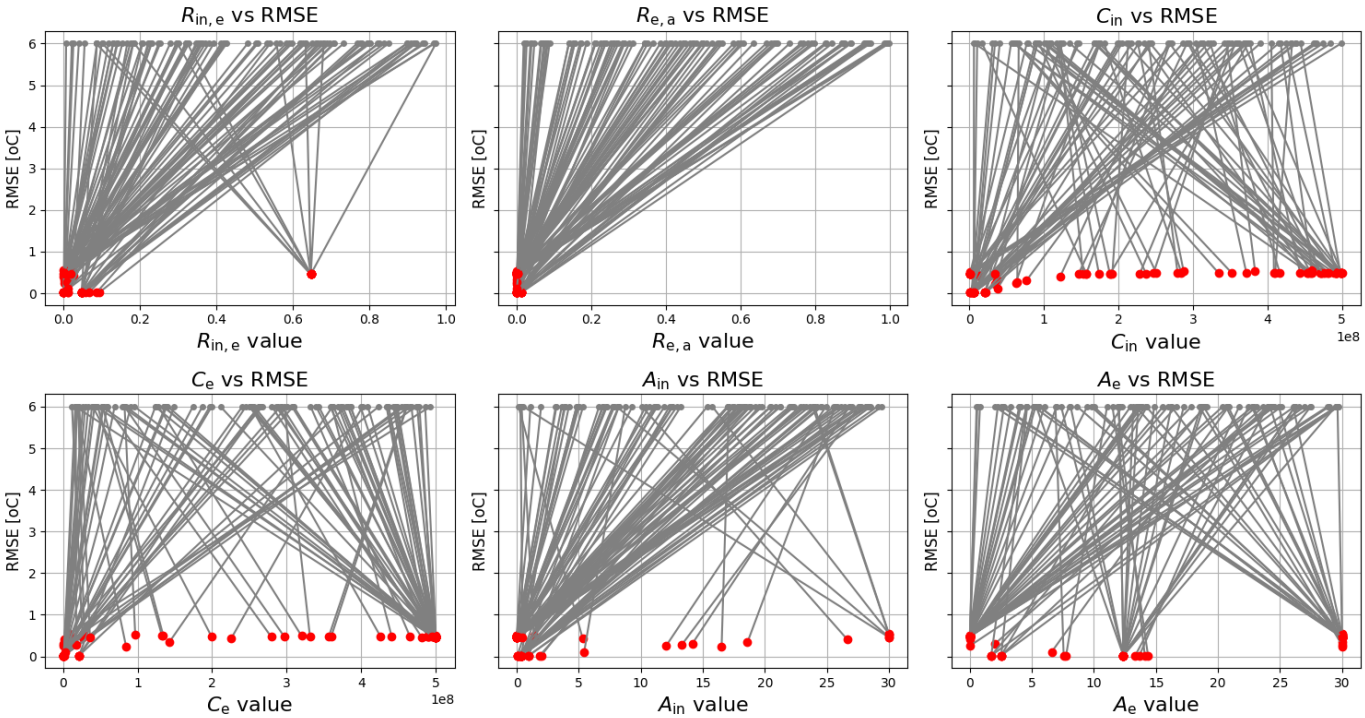


Fig. 5: Parameter trajectories and training RMSE for the 2R2C model. Each line connects the initial randomized value (blue marker) to the final optimized value (red marker).

TABLE V: Summary of Tuning Sensitivity Analysis in Terms of Random Initializations

Metric	**1R1C**	**2R2C**	**3R2C**	**4R3C**
No. of Optimal Solutions	43	7	9	2
Mean RMSE	0.28	0.32	0.12	0.13
RMSE Std. Dev.	0.24	0.21	0.15	0.15

VI. Conclusions

In conclusion, this work presented a systematic methodology for the identification and tuning of grey-box thermal models. While the 4R3C model achieved the highest accuracy during the training and validation phases, its performance during the testing phase was less encouraging. Additionally, its computational cost increases exponentially, making it unsuitable for large-scale implementations. Simpler models, such as the 1R1C, demonstrated high computational efficiency but failed to capture the dynamic thermal behavior under varying external conditions. Overall, the 3R2C model emerged as the most suitable option, offering a balanced trade-off between accuracy and computation time.

Moreover, the tuning process was found to be sensitive to the initial parameter values, indicating the presence of multiple local minima in the solution space. Repeating the optimization with randomized initializations proved essential to ensure robustness and to avoid convergence to suboptimal solutions. While the primary advantage of grey-box models lies in maintaining the physical interpretability of model parameters, the experimental results indicate that this is not always guaranteed. The final solution depends heavily on

the solver's algorithm and the input data, which may limit interpretability when convergence to the global optimum is not achieved.

Future work will adopt the proposed framework to optimal operate heat pumps in active distribution grids. Therefore, this linear model can be easily integrated into optimization frameworks with the corresponding network elements and used to take operation decisions.

REFERENCES

[1] A. S. Gaur, D. Z. Fitiwi, and J. Curtis, "Heat pumps and our low-carbon future: A comprehensive review," *Energy Research & Social Science*, vol. 71, p. 101764, 2021.

[2] European Commission, "Heat pumps," https://energy.ec.europa.eu/topics/energy-efficiency/heat-pumps_en, 2023, accessed: 2025-04-11.

[3] AmBIENCe Consortium, "AmBIENCe Project: Active Managed Buildings with Energy Performance Contracting," https://www.ambience-project.eu/, 2019, funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No 847054.

[4] H. Harb, N. Boyanov, L. Hernandez, R. Streblow, and D. Müller, "Development and validation of grey-box models for forecasting the thermal response of occupied buildings," *Energy and Buildings*, vol. 117, pp. 199–207, 2016.

[5] I. Jankovic, X. Fernandez, and J. Diriken, "Database of grey-box model parameter values for eu building typologies," *Ambience EU Project Report*, 2021.

[6] TABULA Project Team, "TABULA: Typology Approach for Building Stock Energy Assessment," https://episcope.eu/welcome/, 2012, funded by the European Union's Intelligent Energy Europe programme.

[7] Hotmaps Project Consortium, "Hotmaps: The Open Source Mapping and Planning Tool for Heating and Cooling," https://www.hotmaps-project.eu/, 2020, funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No 723677.

[8] BuiltHub Consortium, "BuiltHub: Building Stock Data Flow Hub," https://builthub.eu/, 2020, funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No 957026.

[9] University of Wisconsin-Madison, "TRNSYS: Transient System Simulation Tool," https://www.trnsys.com/, 2023, accessed: 2025-04-11.

[10] U.S. Department of Energy, "EnergyPlus Energy Simulation Software," https://energyplus.net/, 2023, accessed: 2025-04-11.

[11] EQUA Simulation AB, "IDA Indoor Climate and Energy (IDA ICE)," https://www.equa.se/en/ida-ice, 2023, accessed: 2025-04-11.

[12] T. Berthou, P. Stabat, R. Salvazet, and D. Marchio, "Development and validation of a gray box model to predict thermal behavior of occupied office buildings," *Energy and Buildings*, vol. 74, pp. 91–100, 2014.

[13] S. F. Fux, A. Ashouri, M. J. Benz, and L. Guzzella, "Ekf based self-adaptive thermal model for a passive house," *Energy and Buildings*, vol. 68, pp. 811–817, 2014.

[14] A. Navarro-Espinosa and P. Mancarella, "Probabilistic modeling and assessment of the impact of electric heat pumps on low voltage distribution networks," *Applied Energy*, vol. 127, pp. 249–266, 2014.

[15] F. Déqué, F. Ollivier, and A. Poblador, "Grey boxes used to represent buildings with a minimum number of geometric and thermal parameters," *Energy and buildings*, vol. 31, no. 1, pp. 29–35, 2000.

[16] O. M. Brastein, D. W. U. Perera, C. Pfeifer, and N.-O. Skeie, "Parameter estimation for grey-box models of building thermal behaviour," *Energy and Buildings*, vol. 169, pp. 58–68, 2018.

[17] K. K. Andersen, H. Madsen, and L. H. Hansen, "Modelling the heat dynamics of a building using stochastic differential equations," *Energy and Buildings*, vol. 31, no. 1, pp. 13–24, 2000.

[18] P. Bacher and H. Madsen, "Identifying suitable models for the heat dynamics of buildings," *Energy and buildings*, vol. 43, no. 7, pp. 1511–1522, 2011.

[19] G. Reynders, J. Diriken, and D. Saelens, "Quality of grey-box models and identified parameters as function of the accuracy of input and observation signals," *Energy and buildings*, vol. 82, pp. 263–274, 2014.

[20] H. Madsen and J. Holst, "Estimation of continuous-time models for the heat dynamics of a building," *Energy and buildings*, vol. 22, no. 1, pp. 67–79, 1995.

[21] I. Sartori, H. T. Walnum, K. S. Skeie, L. Georges, M. D. Knudsen, P. Bacher, J. Candanedo, A.-M. Sigounis, A. K. Prakash, M. Pritoni *et al.*, "Sub-hourly measurement datasets from 6 real buildings: Energy use and indoor climate," *Data in Brief*, vol. 48, p. 109149, 2023.

APPENDIX A
MODEL STRUCTURES

1) Model 1R1C – Single-Node: The most basic configuration is the 1R1C model, shown in Fig. 6a, and widely used in early-stage estimation studies [4], [5], [13]. It consists of a single thermal node with capacitance C_i representing the thermal capacity of the whole building's indoor environment. This node is connected to the external environment through a single thermal resistance $R_{\text{in,ext}}$; a higher resistance means better insulation. Heat inputs from the heating system ($\dot{Q}_h$) and solar irradiance ($\dot{Q}_{\text{sun}}$) are applied directly to the internal air node. This model provides a lumped estimation of building thermal dynamics, with minimal computational burden. The product $A_{in}Q_s$ is a key component of this model structure. According to [4], A_{in} is a tunable parameter typically ranging from 0 to 0.25, while [5] defined this variable without further explaining how this parameter is estimated, tuned, or bounded.

$$C_{in}\frac{dT_{in}}{dt} = \frac{T_a - T_{in}}{R_{\text{in,a}}} + Q_h + A_{in}Q_s \tag{10}$$

2) Model 2R2C – Envelope Split: A 2R2C model structure—illustrated in Fig. 6b—represents the thermal mass of the building envelope (T_e), distinguishing it from the indoor air temperature T_{in} [5]. The key difference in the model presented here compared with the proposed one by [5] is that the thermal input powers (Q_s, Q_h) were modeled in both indoor and envelope temperatures, with the corresponding tuning factor instead of only indoor input thermal powers.

$$C_{in}\frac{dT_{in}}{dt} = \frac{T_e - T_{in}}{R_{\text{in,e}}} + f_h Q_h + A_{in}Q_s \tag{11}$$

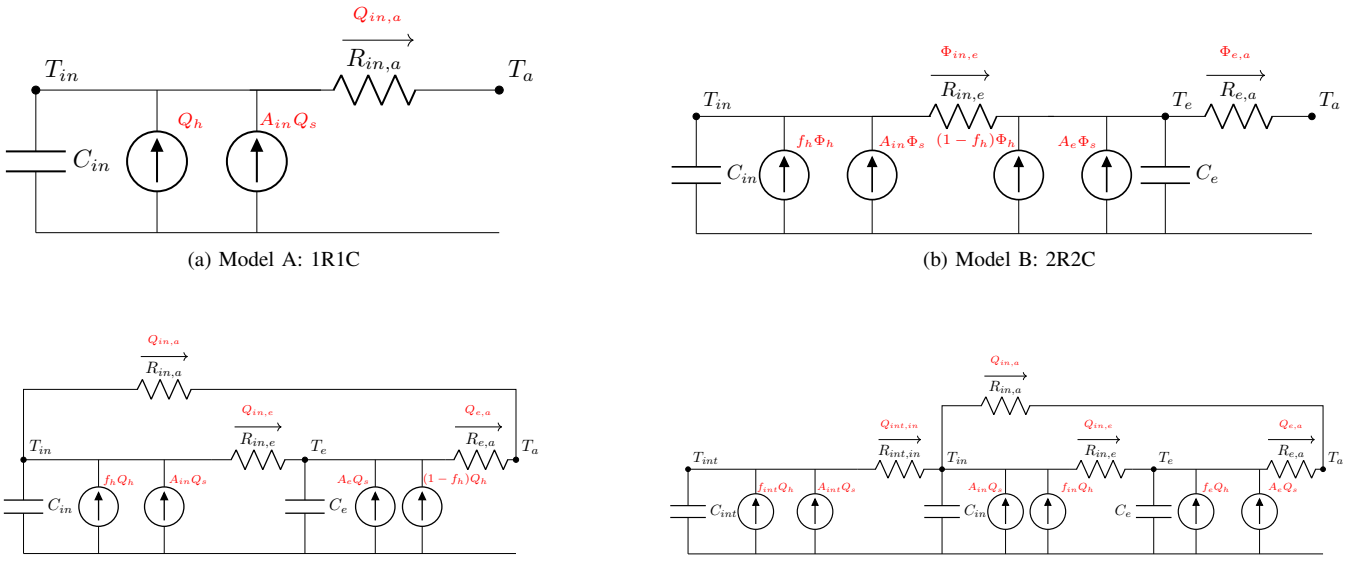
$$C_e\frac{dT_e}{dt} = \frac{T_{in} - T_e}{R_{in,e}} + \frac{T_a - T_e}{R_{a,e}} + (1 - f_h)Q_h + A_eQ_s \tag{12}$$

3) Model 3R2C – Ventilation-Aware: The 3R2C model—depicted in Fig. 6c—introduces an additional resistance to represent heat transfer due to ventilation, distinguishing it from internal conduction ($R_{in,e}$) and external envelope resistance ($R_{e,a}$). This model extends the 2R2C structure, enabling more accurate predictions in scenarios with constant airflow rates. This model is also adopted from [12], [16]. Again, the heat input flux varies between models in the literature. Here, a superset relies on the tuning process to define which is more important.

$$C_{in}\frac{dT_{in}}{dt} = \frac{T_e - T_{in}}{R_{in,e}} + \frac{T_a - T_{in}}{R_{in,a}} + f_h Q_h + A_{in}Q_s \tag{13}$$

$$C_e\frac{dT_e}{dt} = \frac{T_{in} - T_e}{R_{in,e}} + \frac{T_a - T_e}{R_{e,a}} + (1 - f_h)Q_h + A_eQ_s \tag{14}$$

4) Model 4R3C – Four-Node Structure: The 4R3C model illustrated in Fig. 6d extends the previously mentioned 3R2C model by introducing a second thermal node to represent internal contents, such as furniture or indoor walls, improving the accuracy of capturing thermal inertia. The internal components temperature, the indoor air temperature, and the temperature



(a) Model A: 1R1C

(b) Model B: 2R2C

(c) Model C: 3R2C

(d) Model D: 4R3C

Fig. 6: Comparison of different modeling scenarios.

of the envelope are included. The heat flux from the heating system is partly transferred to each node as described in (18).

$$C_{\text{int}}\frac{dT_{\text{int}}}{dt} = \frac{T_{\text{in}} - T_{\text{int}}}{R_{\text{int,in}}} + f_{h,int}Q_{\text{h}} + A_{int}Q_{\text{s}} \tag{15}$$

$$C_{\text{in}}\frac{dT_{\text{in}}}{dt} = \frac{T_{\text{e}} - T_{\text{in}}}{R_{\text{in,e}}} + \frac{T_{\text{a}} - T_{\text{in}}}{R_{\text{in,a}}} + \frac{T_{\text{int}} - T_{\text{in}}}{R_{\text{int,in}}} + f_{h,in}Q_{\text{h}} + A_{in}Q_{\text{s}} \tag{16}$$

$$C_{\text{e}}\frac{dT_{\text{e}}}{dt} = \frac{T_{\text{a}} - T_{\text{e}}}{R_{\text{e,a}}} + \frac{T_{\text{in}} - T_{\text{e}}}{R_{\text{in,e}}} + f_{h,e}Q_{\text{h}} + A_eQ_{\text{s}} \tag{17}$$

$$f_{h,int} + f_{h,in} + f_{h,e} = 1 \tag{18}$$