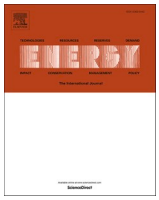
Energy 346 (2026) 140260

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy



RCBldEng - a computationally efficient RC network modeling method and engine for multizone building simulation

Pengyuan Shen

Institute of Future Human Habitats, Shenzhen International Graduate School, Tsinghua University, Shenzhen, 518055, China

ARTICLE INFO

Keywords:
Building simulation
RC model
Energy use
Model calibration
Parametric optimization

ABSTRACT

As there is a lack of resistance-capacitance (RC) modeling method and framework for universal applicability and standardized implementation to multizone building simulation practices, a novel hybrid RC modeling approach that combines physics-based principles (forward modeling) with real-world data calibration (inverse modeling) for building energy simulation is presented in this study. A simulation engine called RCBldEng that combines theoretical building physics with real world calibration capability is developed. Four RC model configurations are proposed and evaluated for different building types and thermal behaviors (4R1C, 6R1C, 7R1C, and 7R2C) with increasing model complexity. The EnergyPlus simulations of two prototype buildings and an educational building case study in the real world are used to test and validate the performance of the simulation engine. The results indicate that the inclusion of interzonal thermal coupling and dual capacitance mechanisms (7R1C and 7R2C) substantially improves prediction accuracy leading to R^2 values of up to 98.78% for office building cooling load prediction against EnergyPlus. For simplified building simulation model, configurations such as the 4R1C model are suitable, but the 7R2C model can better represent the thermal behavior of buildings with high effective thermal mass and multiple zones. Moreover, the differential evolution algorithm turns out to be an effective choice for model calibration for existing buildings, and real-world operational uncertainties due to seasonal variations can be important bias sources. It is shown that the computational efficiency of RCBldEng can be a competitive candidate in performing building simulation for preliminary design stage and optimization applications.

1. Introduction

How architectural design and engineering teams make rapid and effective green low-carbon architectural design decisions and apply energy-saving technologies while ensuring a good built environment has become an indispensable problem in achieving sustainable goal in the building sector [1,2]. Among them, building thermal load and energy consumption simulation calculations are essential methods and means for accurately evaluating energy consumption and carbon emission levels during the operation phase of a building when combined with appropriate emission factors and grid data [3]. Building energy simulation tools like EnergyPlus provide accurate thermal load predictions but require substantial computational resources, particularly when evaluating multiple design alternatives during early design stages [4]. This computational burden limits architects and engineers from conducting the exhaustive parametric studies essential for optimizing green building performance [5]. Considering this, reduced-order resistance--capacitance (RC) network models offer a promising alternative by dramatically reducing simulation time while maintaining acceptable accuracy. However, existing RC modeling approaches face challenges: pure forward models lack calibration capability for existing buildings, while inverse data-driven methods cannot be applied to new designs without operational data. This study addresses this gap by developing RCBldEng - a hybrid RC modeling engine that combines theoretical building physics principles with real-world data calibration, enabling both new building design optimization and existing building retrofit analysis with computational efficiency 10–15 times faster than traditional tools.

A novel hybrid RC building modeling approach is proposed in this study, which combines forward building physics principles (e.g. RC theory) with real world data calibration (inverse modeling). What distinguishes this approach from existing RC modeling methods is the methodological integration that enables dual applicability within a single framework. Traditional RC modeling studies employ either forward methods—calculating parameters directly from known building physics properties—or inverse methods—identifying all parameters

E-mail address: Pengyuan_pub@163.com.

https://doi.org/10.1016/j.energy.2026.140260
Received 5 July 2025; Received in revised form 1 November 2025; Accepted 28 January 2026
Available online 3 February 2026

through data-driven optimization of measured performance. The hybrid approach developed in this study differs fundamentally in three key aspects: (1) Selective parameter treatment: physics-based parameters (envelope thermal properties, geometry) are calculated using forward modeling, while uncertain parameters (effective thermal capacitance, infiltration rates) are calibrated through inverse optimization, rather than treating all parameters uniformly; (2) Universal applicability: the same modeling engine serves both new building design scenarios (where physical properties are specified but no operational data exists) and existing building analysis (where measured performance data enables calibration); (3) Progressive model framework: four distinct RC configurations (4R1C through 7R2C) are systematically developed and validated, with explicit model selection guidance based on building characteristics, interzonal coupling importance, and accuracy requirements. Different from traditional methods solely based on forward physics-based calculations or inverse data-driven parameter identification, the hybrid approach makes the model applicable to both new building design scenarios for which physical properties are known and existing building analysis for which calibration data is available.

2. Literature review

Current research efforts aimed at reducing building energy consumption focus on areas such as green building scheme optimization [6], building energy system optimization control [7], renewable energy application [8], and fault modeling analysis [9]. Typically, building energy simulation models can be categorized into white-box, black-box, and gray-box models. The white-box model is based on the conservation equations of mass, energy, and momentum. A notable feature of the white-box model is that each parameter has solid physical significance, hence it is also referred to as physics-based model. Basic inputs for the white-box building energy consumption model generally include building envelope structure, information on personnel and equipment schedules, HVAC system parameters, building environmental parameters, etc. The downside is its numerous input parameters, and lengthy modeling and calculation times, although the simulation results are relatively accurate. Current mature white-box model software includes EnergyPlus [10], TRNSYS [11], DeST [12,13], and eQuest [14].

The black-box model usually refers to the data-driven modeling method. Establishing an accurate black-box model requires abundant and high-quality training data and an appropriate algorithm [15]. Black-box modeling does not require researchers to have specialized and in-depth engineering knowledge background because it is a mapping of relationship between input and output data. Moreover, the black-box model has strong adaptability. The models can adjust with changes in training data. The drawbacks of the black-box model are also evident: high data quality requirements, low interpretability, high model training costs, and inability to be generalized to other buildings. For heating, ventilation, and air conditioning (HVAC) systems, the black-box model is mainly used for energy consumption estimation or load prediction, which can then be used for control, energy management, etc. Typical algorithms include linear regression [16], change-point linear methods [17], neural networks [18], tree-based algorithms [19], and support vector machines [20,21]. Recent developments in data-driven building energy modeling have expanded to include deep learning architectures, digital twin frameworks, and hybrid physics-data approaches. Transformer-based models and graph neural networks have shown promise in capturing complex temporal and spatial dependencies in multizone buildings [22,23]. Digital twins integrating real-time data with physics-based models are increasingly deployed for building energy management and predictive control [24]. These emerging approaches demonstrate the growing trend toward combining physics-based knowledge with data-driven techniques, which aligns with the hybrid modeling philosophy adopted in this study. The gray-box model, on the other hand, combines the features of both white-box and black-box models and is also known as a reduced-order model or a simplified model [25].

Compared to the black-box model, the gray-box model has physical significance and is easier to interpret. Compared to the white-box model, its model is simpler with lower computational costs. Among them, the resistance-capacitance (RC) model is the most widely studied and applied gray-box model. The development of RC models for building thermal simulation has a rich history dating back several decades. The RC model is established through the circuit analogy, wherein parameters R analogous to resistance represent building thermal resistance, and parameters C analogous to capacitance represent building thermal capacitance. While connecting nodes represent temperature, and the temperature difference between nodes is analogous to the voltage difference in circuits. The order of the RC model depends on the number of Cs, because for each C, its control method includes a differential equation [26]. Currently, common RC model establishment methods include forward and inverse methods.

2.1. RC-based modeling for building energy simulation

Most of the current building load or energy consumption modeling based on the RC model requires the use of historical load or energy consumption data, with the main work being the parameter identification process of the RC model. This method is commonly referred to as the inverse method, which focuses on using data-driven methods to obtain resistance (R) and capacitance (C) values. In the domain of inverse modeling for buildings, ASHRAE's Inverse Model Toolkit (IMT) represents an important contribution to the field. Developed as part of ASHRAE Research Project 1050 [27], the IMT provides standardized implementations of linear, change-point linear, variable-base degree-day, and multi-linear regression data-driven models specifically designed for analyzing building energy consumption data.

Currently, data-driven methods for inversely determining system parameters mainly include regression algorithms [28,29], Gaussian regression models [30], sequential quadratic programming [31], and genetic algorithms [32,33]. Tian et al. invented a building dynamic room temperature prediction method based on the equivalent RC model, using a genetic algorithm (GA) for step-by-step model parameter identification to obtain a parameter-determined equivalent RC model for predicting dynamic average room temperature [34]. Massa Gray and Schmidt proposed a hybrid model that combines the Gaussian regression model as part of machine learning with the 4R4C model to simulate building energy consumption of a single-zone office building in Stuttgart, Germany, which is heated and cooled by radiators and chilled ceilings. The study found that, in terms of energy consumption prediction, the hybrid method has better predictive performance compared to the Gaussian-process-only or RC-only model [30]. Shen et al. adopted genetic algorithm to calibrate RC models for building thermal load calculation (sensible load) and energy use simulation under a changing future climate for retrofit optimization purposes at both building scale [35] and regional scale using archetype-based bottom up modeling [36,37]. Li et al. established a radiating floor-water coupled 2R1C model for a radiant heating system, using genetic algorithm to identify optimal lumped thermal parameters including thermal resistances and capacitances from experimental data. The study focused on sensible heating loads only and included a simplified system model for the radiant floor heating system. The parameter identification process involved optimizing multiple RC parameters simultaneously from measured temperature and energy consumption data, addressing the inherent challenge of isolating individual thermal parameters from coupled system responses [38]. The above studies used various data-driven methods for RC model parameter identification (determining the R and C values). Moreover, the size of the training dataset has a significant impact on system identification. The time span of the training set ranges from as short as 7 days [28] to as long as 60 days [39]. Massa Gray and Schmidt compared Gaussian processes with a physics-based gray-box RC model, finding that while Gaussian processes achieved better accuracy during

occupied periods, the RC model performed more consistently across different operational conditions [40]. However, their referenced building energy use data that was used to calibrate the RC model is simulated by TRNSYS instead of metered data of a real building. The premise of using the inverse method is the appropriate calibration algorithm and a large amount of experimental data to obtain reasonable system parameters and establish an energy consumption prediction model. Recent advances in inverse modeling have increasingly incorporated machine learning techniques and uncertainty quantification methods. Bayesian calibration approaches have gained prominence for their ability to quantify parameter uncertainty while optimizing model predictions [41, 42]. Physics-informed neural networks have emerged as promising tools for building energy modeling, combining data-driven learning with physical constraints [43]. Ensemble-based methods have also been developed to address prediction uncertainty in calibrated building models [44]. While these advanced techniques offer sophisticated uncertainty quantification, they often require substantial computational resources and expertise. In contrast, the differential evolution-based approach proposed in this work provides a balance between calibration effectiveness, computational efficiency, and ease of implementation, making it particularly suitable for practical building energy applications.

Another way to establish an RC model is the forward method, which establishes a thermal balance model for the building by defining building thermal parameters such as the thickness of the enclosure structure, heat transfer coefficient, and thermal storage heat capacity. In our previous work [33], we demonstrated that considering thermal transfer between multiple thermal zones can significantly impact the accuracy of RC model predictions for building load and indoor temperature, which motivated the development of the interzonal thermal coupling models presented in this study. R and C values are typically obtained based on fundamental heat transfer principles including conductive, convective, and radiative heat transfer through building envelope components (walls, windows, roofs), thermal mass properties (density, specific heat capacity, thickness), and surface heat transfer coefficients, or alternatively through field measurements of temperature and energy consumption data. In urban block-scale microclimate research, Bueno and others proposed a city canopy and building energy consumption prediction model based on the forward RC model to study the impact of urban heat island effects on building energy consumption [45]. Regarding building-scale energy consumption simulation, Vivian et al. compared a simplified 7R2C RC model against TRNSYS simulations for a single-zone residential building under four European climates (Continental, Mediterranean, Oceanic, and Subarctic) [46]. The hypothetical case study building was a single-floor structure with typical residential envelope characteristics and windows. The RC model parameters were calibrated using building physics principles rather than inverse methods, and the results showed good agreement with TRNSYS energy consumption predictions across different climate zones. The premise of using the forward method is to obtain building physical parameters (envelope thermal properties, geometry, thermal mass characteristics) to determine the RC network parameters (thermal resistances and capacitances), building operation conditions (occupancy schedules, internal loads, weather conditions), and establish the building thermal model that characterizes the building's thermal behavior. The literature search reveals that while RC thermal network models have a rich history dating back to early pioneers like Paschkis (1942) [47] and Robertson and Gross (1958) [48] who developed foundational RC approaches that later evolved into TETD/TA and CLTD/CLF/SCL methods, current research using forward RC model methods specifically for multi-thermal zone whole-building load simulation at the building scale still remains limited.

2.2. Whole-building energy simulation based on RC model

Table 1 summarizes the current building energy simulation models and platform development based on the RC model theory. Giretti et al. developed an energy consumption prediction model based on a three-order RC model in Modelica, which includes building modules, air conditioning modules, personnel information modules, and weather modules, and established a model parameter calibration process [49]. Coninck et al. established building models, thermal zone models, and air-conditioning models based on the Fastbuildings in the Modelica standard library. Using the Gray-box library in Python, they separately established energy consumption prediction models for single thermal zones and double thermal zones [50]. Remmen et al. established a reduced-order urban-scale building energy consumption prediction tool (TEASER) in Modelica, addressing issues of long computation time and large errors in urban-scale energy consumption predictions [51]. RC_BuildingSimulator is a building energy consumption simulation tool based on the Python programming language. This model is established based on the ISO 13790 standard recommended model and is a single thermal zone load simulation calculation tool [52]. However, the tool is not easy to use due to limited access to building modeling inputs. Additionally, there are some energy consumption simulation tools based on the RC model developed in C++ and R language [53].

However, current research on multi-thermal-zone RC models rarely considers the coupled heat transfer issue between thermal zones. For example, Bacher and Madsen developed a multiple thermal zone RC model using inverse parameter identification based on data from a single-story experimental building, but their approach treated each zone independently without modeling interzonal heat transfer, limiting its applicability to buildings with significant thermal coupling between zones [54]. Similarly, Giretti et al. established a four-order RC model for energy consumption prediction that includes multiple building modules but does not account for thermal interactions between adjacent zones, which can lead to significant prediction errors in buildings with open-plan layouts or shared thermal masses [49]. A preliminary attempt by the author indicates that considering the thermal transfer between multiple thermal zones can significantly impact the accuracy of the RC model's predictions of the load and indoor temperature within a building's thermal zone [33]. These limitations highlight the need for RC models that explicitly represent interzonal thermal coupling, which is a key contribution of our proposed 6R1C, 7R1C, and 7R2C models.

As the number of thermal zones increases, the complexity of the model also increases, which will put pressure on the computational cost. Recent studies have demonstrated that computational time for multi-zone building simulations can increase exponentially with zone count,

Table 1
Current progress on RC based building energy simulation tools.

Tool name	Subject	Order of the model	Model type	Thermal zone	Load calculation	Energy calculation	Thermal comfort calculation	Automated model calibration	Platform	Ref
FastBuildings	Building	Multi-order	Inverse	Multiple	✓	×	×	✓	Modelica	[50]
RC_BuildingSimulator	Building	First	Forward	Single	✓	✓	×	×	Python	[52]
ISOmodel	Building	First	Forward	Single	✓	✓	×	×	C++	[53]
Giretti model	Building	Third	Inverse	Single	✓	✓	×	✓	Modelica	[49]
TEASER	Urban	Third	Forward	Multiple	✓	×	×	×	Python/ Modelica	[51]

particularly for detailed energy models [55,56]. Therefore, reasonable model order reduction and solution algorithms are needed to solve multi-thermal-zone problems [57,58]. Overall, most of the existing related research is limited to the development of single thermal zone models or multizone models with limited zones and lack universal tool that can be widely applied to RC modeling for buildings [25]. Their building simulation functions are usually incomplete (such as lack of energy consumption calculations, thermal comfort models and calculations, and automatic model calibration functions), and the computational accuracy does not look promising. Moreover, by treating multizone building as single thermal zone building can lead to biased simulation results [59]. The fundamental problem lies in the current lack of a comprehensive, accurate, and universally applicable forward RC model-based multi-thermal-zone building energy consumption modeling method and scheme.

2.3. Motivation and contribution of this research

Since current research based on RC models mainly focuses on the inverse method, and research related to the forward method is currently only for single thermal zone buildings (or simplifying multizone buildings into a single model), its practicality and applicability cannot address the huge computational cost incurred by "massive simulation scenarios" in the process of building energy-saving design optimization and iteration. Recent reviews of gray-box modeling approaches [25,60] have identified the need for comprehensive, validated RC modeling frameworks that can handle multizone buildings with interzonal thermal coupling while maintaining computational efficiency. In some of the latest research, advanced building simulation methods increasingly emphasize the integration of physics-based and data-driven approaches [61], which motivates our hybrid modeling strategy. The fundamental gap here for RC modeling lies in the lack of a unified framework that can leverage both theoretical building physics and empirical calibration data depending on data availability. In view of this, this study aims to develop a fast simulation method for building multi-thermal-zone building energy consumption based on hybrid RC model that combines theoretical building physics principles (forward modeling) with real-world data calibration (inverse modeling), which has robust model interpretability, fast solution speed, and affordable calculation bias. The research attempts to explore the computational cost and accuracy of the model algorithm at different model orders, solution methods, and building types.

The development of the proposed dedicated simulation engine (RCBldEng) in this research helps address a critical gap: existing RC modeling implementations lack the integration of forward and inverse capabilities within a unified, accessible tool. While academic studies have demonstrated various RC modeling techniques, practitioners face significant barriers including: (1) absence of standardized multizone RC simulation tools with comprehensive building physics modules, (2) fragmented implementations requiring users to develop custom models for each application, and (3) lack of automated calibration workflows for existing buildings. RCBldEng overcomes these limitations by providing a complete, validated simulation engine with structured text-based inputs, automated parameter optimization, and progressive model complexity options. The developed engine can provide architects and engineers with a lightweight alternative for low-cost modeling, fast simulation, and comparison of building energy consumption, and a non-user-intervention module for automatic calibration of existing building model parameters. Executable program on Windows platform has been developed based on the proposed modeling method and published on GitHub (can be found in the Code Availability section). The research deliverables of this work can be generalized and widely applied to rapid iteration and comparative evaluation of low-energy consumption building design schemes for new and existing buildings. RCBldEng has prospective practical application value in building design, energy-saving and emission reduction optimization related problems.

Hence, the primary objectives of this research are threefold: (1) to develop and validate four progressive RC model configurations (4R1C through 7R2C) that systematically incorporate interzonal thermal coupling and dual capacitance mechanisms for improved multizone building prediction; (2) to create a hybrid modeling framework integrating forward physics-based calculations with inverse data-driven calibration, enabling universal application to both new building design and existing building retrofit; and (3) to implement these methods in RCBldEng, a computationally efficient simulation engine with comprehensive validation across building types and climate zones to establish practical model selection guidelines. Fig. 1 provides a visual summary of the research rationale, showing the challenges in current building simulation approaches, the proposed hybrid RC modeling solution, and the expected benefits of this methodology.

3. Methodology

3.1. The proposed hybrid RC models

3.1.1. RC model without interzonal thermal coupling

The proposed 4R1C model has two temperature nodes (see Fig. 2), which are T_a and T_m. T_a and T_m represent internal air node and thermal mass node, respectively.

$$\frac{T_o - T_a}{R_v} + \frac{T_m - T_a}{R_{im}} + Q_{hvac} + Q_{air} = 0 \tag{1}$$

$$C_m \frac{dT_m}{dt} = \frac{T_o - T_m}{R_{win}} + \frac{T_o - T_m}{R_{ex}} + \frac{T_o - T_m + (Q_{hvac} + Q_{air})R_v}{R_{im} + R_v} + Q_{sol} + Q_{int} \tag{2}$$

where.

T_o: Outside air temperature (K);
T_a: Indoor air temperature of the zone (K)
T_m: Temperature of the building's envelope thermal mass (K)
R_v: Thermal resistance for ventilation and infiltration heat transfer between outdoor and indoor air (m^2K/W);
R_{win}: Thermal resistance of windows (m^2K/W);
R_{im}: Thermal resistance between internal opaque structures and thermal mass (m^2K/W);
R_{ex}: Thermal resistance of external opaque envelope elements (m^2K/W);
Q_{hvac}: Heat flux delivered by the HVAC system to the zone (W);
Q_{air}: Convective heat flux from internal sources directly to indoor air (people, equipment, lighting) (W);
Q_{int}: Radiative heat flux from internal sources to surfaces (people, equipment, lighting) (W);
Q_{sol}: Solar radiation heat flux absorbed by the building's thermal mass(W);
C_m: thermal capacity of internal thermal mass per building area (J/m^2K)

The 4R1C model presented is a simplified yet effective representation of a building's thermal behavior, particularly when considering scenarios where interzonal thermal interactions are negligible. With only two temperature nodes—T_a and T_m—the model can capture the primary thermodynamic interactions in most typical building environments. The thermal mass node (T_m) represents the combined thermal mass of the building opaque structures, including both internal structures and the mass embedded within the building envelope. In this model, the temperature node T_a denotes the air temperature inside the building. It is influenced by external weather conditions, characterized by T_o, and the building's HVAC system Q_{hvac}, and internal air exchanges Q_{air}. Thermal resistances, namely R_v and R_{im}, dictate the heat flow that is exchanged between the external environment and the internal air, and

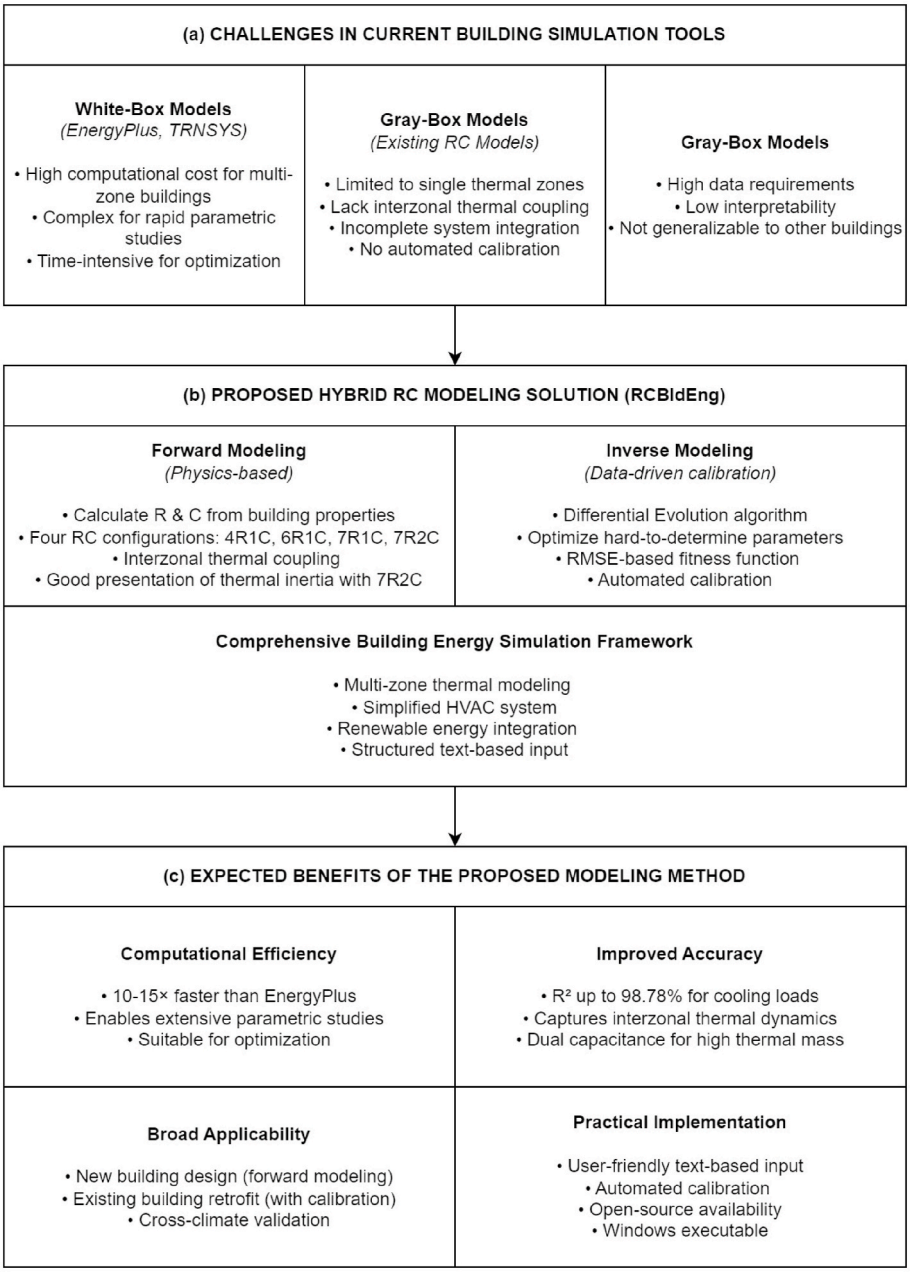


Fig. 1. Research rationale flowchart.

between the internal air and the building's thermal mass, respectively.

The second node, T_m, represents the temperature of the building's thermal mass, capturing the aggregate behavior of building structures and contents. It balances the influences from external conditions T_0, internal loads Q_{int}, solar gains Q_{sol}, and interactions with the internal air node. Its dynamic behavior is defined by C_m, which is the thermal capacity of the building's internal mass.

3.1.2. RC models with interzonal thermal coupling and different parameters

Interzonal thermal coupling can play a critical role in capturing the intricate thermal interactions between different zones within a building. Hence, three distinct RC models—6R1C, 7R1C, and 7R2C—have been proposed to address these complexities. The structures of the three models, including the 4R1C model without interzonal thermal coupling, are plotted together in Fig. 3.

- The 6R1C model

$$\frac{T_o - T_a}{R_v} + \frac{T_m - T_a}{R_{im}} + Q_{hvac} + Q_{air} = 0 \tag{3}$$

$$C_m \frac{dT_m}{dt} = \frac{T_o - T_m}{R_{win}} + \frac{T_o - T_m}{R_{ex}} + \sum_i \frac{T_{az,i} - T_m}{R_{if,i}} + \sum_i \frac{T_{az,i}}{R_{iw,i}} + \frac{T_o - T_m + (Q_{hvac} + Q_{air})R_v}{R_{im} + R_v} + Q_{sol} + Q_{int} \tag{4}$$

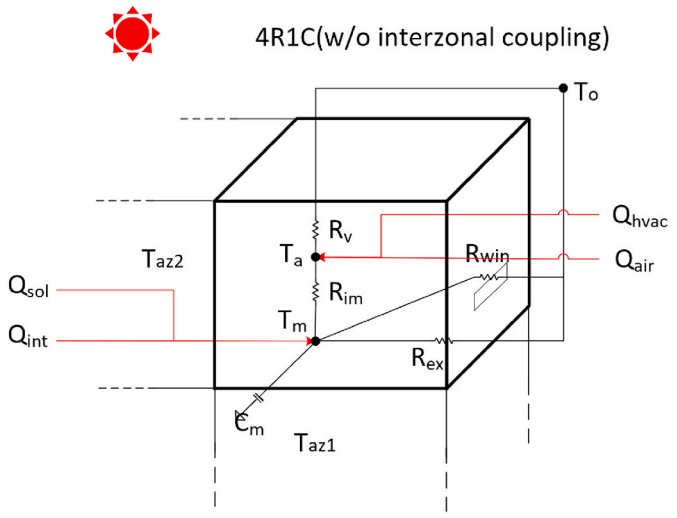


Fig. 2. The proposed 4R1C model without interzonal thermal coupling.

where.

$T_{az,i}$: Indoor air temperature of the ith adjacent zone(K);
$R_{if,i}$: Thermal resistance of the interior floor/ceiling between the current zone and the ith adjacent zone (m^2K/W);
$R_{iw,i}$: Thermal resistance of the interior wall between the current zone and the ith adjacent zone(m^2K/W);

The 6R1C model has two temperature nodes T_a and T_m as well, but it takes into account the heat fluxes between the current zone and all the adjacent zones. Its structure recognizes the importance of the building's internal zones in shaping its thermal behavior, which is essential for larger structures or those with prominent internal heat sources or sinks. By considering both the external environment and neighboring zones, this model offers a more comprehensive thermal view of single zone behaviors, and their adjacent interactions compared to the 4R1C model.

- The 7R1C model

The 7R1C model has three temperature nodes, namely T_a, T_s and T_m, in which T_s represents the central mass node.

$$Q_{hvac} + Q_{air} = \frac{T_a - T_o}{R_v} + \frac{T_a - T_s}{R_{ia}} \tag{5}$$

$$\frac{T_m - T_s}{R_{ia}} + \frac{T_o - T_s}{R_{win}} + \sum_i \frac{T_{az,i} - T_s}{R_{iw,i}} + \frac{T_o - T_s + (Q_{hvac} + Q_{air})R_v}{R_{ia} + R_v} + Q_{int} = 0 \tag{6}$$

$$C_m \frac{dT_m}{dt} + \frac{T_m}{\frac{1}{\frac{1}{R_{ia}+R_v}+\frac{1}{R_{win}}+\frac{1}{R_{iw}}} + R_{im}} = \frac{T_o - T_m}{R_{ex}} + \frac{T_s - T_m}{R_{im}} + \sum_i \frac{T_{az,i} - T_m}{R_{if,i}} + \sum_i \frac{T_{az,i}}{R_{iw,i}} + Q_{sol} + Q_{int} \tag{7}$$

where.

T_s: Temperature of the zone's interior surfaces (m^2K/W);
R_{ia}: Thermal resistance between indoor air and interior surfaces (represents the inverse of the convective heat transfer coefficient) (m^2K/W)

Advancing from the 6R1C configuration, the 7R1C model introduces an additional node, T_s, to represent the temperature of interior surfaces within the thermal zone (such as internal walls, floors, and ceilings). This distinction separates the interior surfaces and indoor furnishings (T_s) from the thermal mass embedded within building opaque elements (T_m). The resistances related to internal walls and floors of neighboring zones ($R_{iw,i}$ and $R_{if,i}$) enable the modeling of heat transfer between adjacent thermal zones. In this way, the 7R1C provides a layered approach to capturing both the different thermal time constants within a zone and the heat exchange between zones.

- The 7R2C model

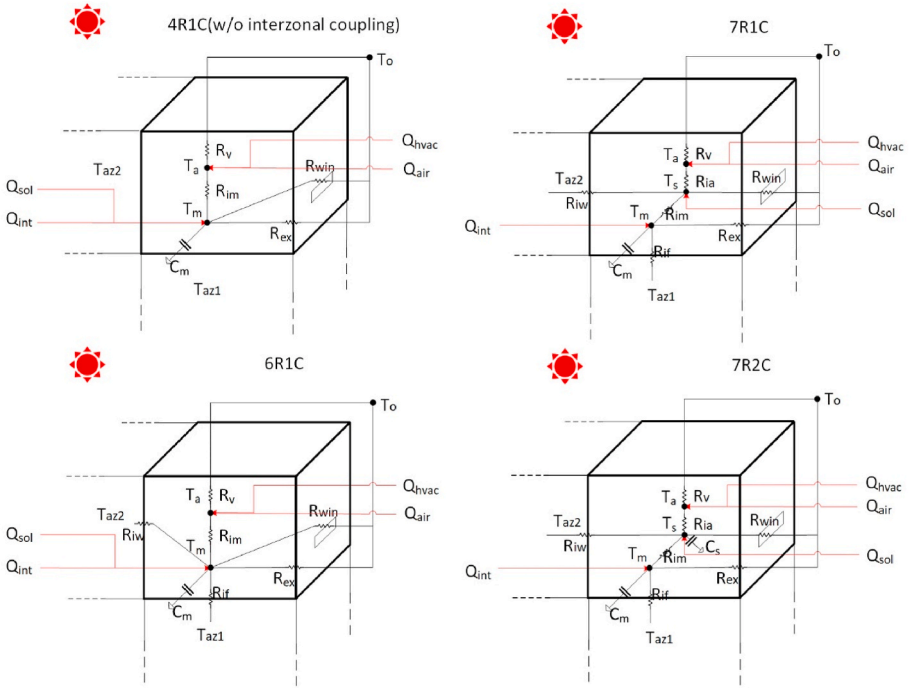


Fig. 3. The proposed RC models with different modeling parameters: a) 4R1C w/o coupling; b) 6R1C; c) 7R1C; d) 7R2C.

$$Q_{hvac} + Q_{air} = \frac{T_a - T_o}{R_v} + \frac{T_a - T_s}{R_{ia}} \tag{8}$$

$$C_s\frac{dT_s}{dt} = \frac{T_m - T_s}{R_{im}} + \frac{T_o - T_s}{R_{win}} + \sum_i \frac{T_{az,i} - T_s}{R_{iw,i}} + \frac{T_o - T_s + (Q_{hvac} + Q_{air})R_v}{R_{ia} + R_v} + Q_{int} \tag{9}$$

$$C_m\frac{dT_m}{dt} + \frac{T_m}{\frac{1}{R_{ia}+R_v} + \frac{1}{R_{win}} + \frac{1}{R_{iw}}} + R_{im} = \frac{T_o - T_m}{R_{ex}} + \frac{T_s - T_m}{R_{im}} + \sum_i \frac{T_{az,i} - T_m}{R_{if,i}} + \sum_i \frac{T_{az,i}}{R_{iw,i}} + Q_{sol} + Q_{int} \tag{10}$$

where.

C_s: Thermal capacitance of interior surfaces and furnishings per unit floor area(J/m^{2}K)

Building upon the foundation of the 7R1C, the 7R2C model further introduces the C_s term, capturing the thermal storage capability of the central thermal mass mode T_s. This distinction in capacitance—C_m for peripheral mass and C_s for central mass—gives the 7R2C model a refined representation of energy storage and discharge dynamics. This dual-capacitance mechanism can ensure a more detailed exhibition of how heat flows within different building components compared with the 7R1C model, making it especially useful for buildings undergoing pronounced day-night thermal variations.

The 7R2C model introduces a second thermal capacitance C_s for interior surfaces and furnishings in addition to existing envelope thermal capacitance C_m which represents progress in our RC modeling framework. Building upon a dual capacitance model allows the recognition of different thermal mass sources in building elements, the 7R2C model separates building interior from opaque structure elements to create different thermal time constants in a single zone since the thermal mass in indoor furnishings and building opaque elements can behave differently due to their response to indoor and outdoor heat flows. The 7R2C model can deliver benefits for simulations of buildings with substantial mass distributions in either their envelope or interiors as well as for fast-changing environmental conditions. The two-capacitance model can also improve calibration procedure because thermal properties from interior and envelope masses can be calibrated respectively.

For multizone buildings, all thermal zones are modeled simultaneously within a coupled system of differential equations. At each time step, the thermal states (temperatures) of all zones are solved together, with interzonal heat transfer terms $\left(\sum\left(T_{az,i} - T_m\right)/R_{if,i} and \sum\left(T_{az,i} - T_m\right)/R_{iw,i}\right)$ explicitly linking adjacent zones through shared internal surfaces. This simultaneous solution approach ensures that thermal interactions between zones are properly represented throughout the simulation, distinguishing our method from simplified approaches that treat zones independently or use sequential solution procedures.

3.2. Modeling structure of forward parameters

The forward modeling and the model inputs collection of a building's energy consumption encompasses several key modules, each intricately contributing to the building's overall thermal performance and energy use.

- Meteorological Input:

The standard weather files, such as the widely used EPW (EnergyPlus Weather) files, serve as input.

- Building Geometry Model:

This model proposes the use of structured text as a means to provide parameter entries for modeling. Users can parametrically modify model inputs in the text domain. All related model parameters are organized and compiled using an object-oriented modular mindset, setting the stage for the integration of the simulation algorithms and parametric design platforms proposed in this project.

- Occupancy Behavior Model:

To reflect the heterogeneous user behavior habits in multi-thermal zone models, a parameterized textual input system is constructed. This system covers space occupancy habits, HVAC system usage patterns, appliance usage, and lighting habits, thereby structuring user behavioral parameters for the simulation algorithms of this project.

- Infiltration or Natural Ventilation Model:

The heat exchange from infiltration is typically calculated using lumped parameters. The model offers two infiltration calculation methods as detailed in Appendix A: a 'constant' method using predefined air change rates (ACH), and a 'real' method that dynamically computes infiltration based on temperature differences and wind conditions according to ISO 13789 and ISO 15242 standards. The 'real' method accounts for both stack effect and wind-driven infiltration, combining them through a quadrature sum approach. For natural ventilation, the model implements behavior-dependent algorithms with opening factors that vary with environmental conditions.

- Solar Radiation Model:

The model uses obtained climatic parameters to generate hourly calculations for direct beam radiation and diffuse and reflected radiation and total radiation across the building's different oriented sides. The calculations conform to the ISO 52010 standard method using adaptations specified in Appendix B. The model includes solar geometry elements (declination, hour angle, altitude, and azimuth) together with surface orientation and anisotropic sky conditions. The transmission of direct beam radiation depends on the outcome of incidence angle calculations. The diffuse component calculates radiation through an anisotropic sky model that adjusts its correction factors from 0.45 to 1.0 according to the incident angle. Moreover, the model calculates ground-reflected radiation based on ground reflectivity factors. The solar radiation heat gain for the building is calculated through the method in Appendix B while using outdoor temperature together with relative humidity and solar azimuth and altitude measurements.

- Convective Heat Transfer Coefficient Modeling and Calculation:

For convective heat transfer, the convective heat transfer coefficient significantly affects the model's predictive accuracy. For exterior surfaces, the model implements convective heat transfer coefficient calculations in accordance with McAdams's method as described in EnergyPlus Engineering Reference [62]. The wind-dependent heat transfer coefficients are implemented using the correlation $h_{c,ext} = 5.7 + 3.8v$, where v is wind speed in m/s.

- Calculation of Shading Coefficient:

The heat from solar radiation inside a building is closely related to the shading coefficient of the window. This coefficient is influenced by shading devices and solar incidence angle. Most RC modeling related studies treat windows as pure resistance without adequately considering the shading coefficient provided by shading devices. Therefore, the developed engine separately considers the indoor solar radiation heat gain, integrating the shading coefficient analogous to the heat transfer coefficient into the energy consumption simulation algorithm. Shading coefficient calculations follow geometric modeling principles as detailed

in Appendix C, providing solar reduction factors (SRF) for common shading devices including overhangs, fins, and horizon obstructions. For overhangs, projection angles typically range from 15° to 60°, with the calculation accounting for both geometric shading of direct beam radiation and partial reduction of diffuse radiation (typically 10–50% depending on configuration). Similar geometric principles apply to vertical fins, with solar azimuth replacing altitude in the calculations. The model includes dynamic shading control with temperature thresholds (26 °C for residential and 23 °C for commercial buildings) to activate shading during overheating periods.

- Thermal Parameters of Building Envelope Model:

The thermal parameters of the building envelope are modularized based on their types. For instance, the window module, which is treated as one of the objects in the envelope settings, includes parameters like heat transfer coefficient, solar radiation absorption rate, emissivity, and transmittance. This model integrates all the thermal parameters of envelopes, including windows, walls, roofs, and floors, according to thermal zones.

- Simplified HVAC System Performance Curving Method:

Central to this study is a simplified performance curve method for the HVAC system. Concerning the primary energy consumption for heating and cooling, a performance curve approach is endorsed. Users are required to input the energy efficiency of the heating and cooling source at intervals of partial load conditions. This method simulates the energy performance of the heating and cooling system under varying load conditions. Upon receiving the energy efficiency inputs for each partial load stage, linear interpolation is employed to emulate the system's performance curve. This approach is designed to streamline model inputs. Moreover, if no HVAC sizing details are provided, the system is assumed to cater to the heating and cooling demands of all zones. If the heating or cooling load surpasses the system's capacity, the indoor temperature will be calculated by governing energy conservation equation as stated in Equations ((1), (3), (5) and (8). Moreover, detailed calculation methods for building energy use and onsite renewable energy production can be found in Appendix D and Appendix E.

By establishing the above modules with hierarchical modeling architecture (zone to building) as shown in Fig. 4, combined with the coupling of thermal flows in thermal zones, a building thermal load simulation method based on the forward RC model is formed. A multi-thermal zone building energy simulation model architecture is thus formed, allowing the further assessment of its performance in calculating the load and energy consumption of an entire building.

The engine uses tailor-made structured text-based modeling files that are designed with simplicity and clarity so that users can easily define and tweak model parameters without steep learning curve. This is to ease the fight with engineers, preventing them from developing their own custom models. This simplifies the data input process and the process of modifications to allow easy iterations and modifications for the researchers and the building professionals. Secondly, the model information collection structure of the RC based simulation engine was implemented in Python 3.9 environment.

3.3. Solving approaches

The RCBldEng simulation engine implements three distinct numerical methods for solving the thermal network equations: the Euler method, Crank-Nicolson method, and Runge-Kutta method. Each method offers different characteristics in terms of stability, accuracy, and computational efficiency.

3.3.1. Euler method

The Euler method, also known as the forward Euler method, represents the simplest explicit time integration scheme. For a given time step Δt, the method advances the solution by approximating the derivative using forward differences. In the context of building thermal simulation, the temperature at the next timestep is calculated as:

$$T^{n+1} = T^n + \frac{\Delta t}{C}(q_{tot} - BT^n) \tag{11}$$

where T^n represents the temperature at the current time step, C is the thermal capacitance, q_{tot} is the total heat gain, and B represents the sum of heat transfer coefficients. While computationally efficient, this

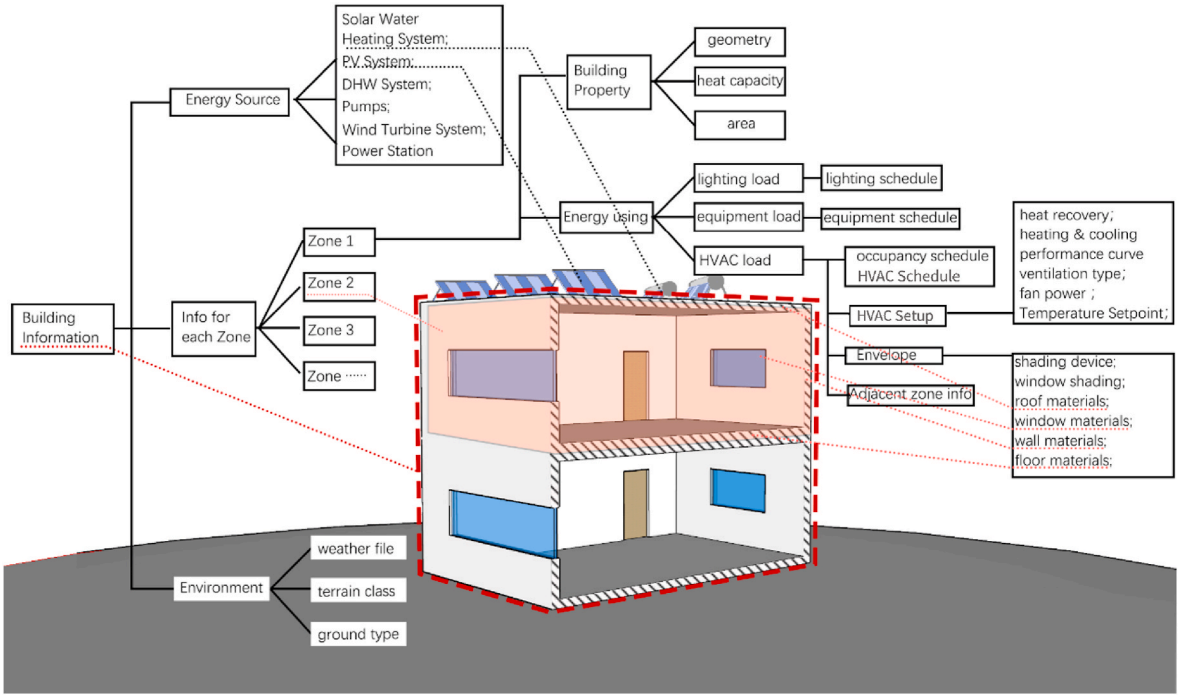


Fig. 4. The hierarchical architecture of model information collection.

method's accuracy is proportional to the time step size, and it may exhibit numerical instability for large time steps or when modeling systems with rapid thermal responses.

3.3.2. Crank-Nicolson method

The Crank-Nicolson method improves upon the Euler method by employing a semi-implicit approach that averages the solution between the current and next time steps. This results in a second-order accurate scheme in time, expressed as:

$$T^{n+1} = \frac{T^n(C - 0.5B\Delta t) + \Delta t q_{tot}}{C + 0.5B\Delta t} \tag{12}$$

This method provides better stability characteristics compared to the Euler method and maintains reasonable accuracy even with larger time steps. The Crank-Nicolson scheme is particularly effective for building energy simulations where thermal processes typically evolve gradually, making it the default solver in the engine.

3.3.3. Runge-Kutta method

For applications requiring higher accuracy, the fourth-order Runge-Kutta method (RK4) is implemented. This method evaluates the temperature derivative at multiple points within each time step, providing a more accurate approximation of the solution trajectory. The RK4 implementation follows:

$$k_1 = \frac{\Delta t}{C}(q_{tot} - BT^n) \tag{13}$$

$$k_2 = \frac{\Delta t}{C}\left(q_{tot} - B\left(T^n + \frac{k_1}{2}\right)\right) \tag{14}$$

$$k_3 = \frac{\Delta t}{C}\left(q_{tot} - B\left(T^n + \frac{k_2}{2}\right)\right) \tag{15}$$

$$k_4 = \frac{\Delta t}{C}(q_{tot} - B(T^n + k_3)) \tag{16}$$

$$T^{n+1} = T^n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \tag{17}$$

While computationally more intensive than the other methods, RK4 provides good accuracy and stability, particularly valuable when modeling buildings with complex thermal interactions or rapid temperature changes.

3.3.4. Adaptive error control and solver selection

The solver implementation includes adaptive error control mechanisms to ensure solution accuracy. For each time step, the temperature change is monitored, and if it exceeds specified thresholds, the solution process can automatically adjust by either reducing the time step or switching to a more stable numerical method. This adaptive approach helps maintain simulation stability while optimizing computational resources.

The simulation engine allows users to select the preferred numerical method based on their specific requirements. The Crank-Nicolson method is set as the default solver due to its balanced performance characteristics. For cases where computational speed is paramount, the Euler method can be selected, while the RK4 method is available for applications demanding higher accuracy. All methods are implemented with an hourly time step (3600s) as the standard temporal resolution, though the framework supports variable time-stepping if required for specific applications.

The solving approaches are integrated with the engine's thermal capacity limitation handling (by user input), ensuring that when the thermal load is over the top of HVAC system capacity, the numerical solution remains stable and physically meaningful while respecting the specified heating and cooling capacity limits.

3.4. Calibration of hard-to-determine modeling parameters (inverse modeling)

For existing buildings, it is inevitable to determine uncertain parameters that have a bearing on the model's outcome. Some of the parameters like the building geometry or the material properties are easy to obtain or quantify while others like the thermal capacitors and system efficiencies are not easy to obtain. These difficult to quantify parameters are important to model the real thermal performance of the building and the efficiency of the HVAC systems. However, their identification is usually more complex especially when the model is for existing buildings where design information might be incomplete or inaccessible.

3.4.1. Differential evolution algorithm implementation

To address this challenge, this study adopts the differential evolution (DE) algorithm [63]—a robust and efficient global optimization technique known for its capability to tackle complex, non-linear, and multidimensional problems. In essence, DE is a population-based optimization algorithm. It begins with a randomly generated population of potential solutions (vectors). Over successive iterations, these vectors evolved by combining the processes of mutation, crossover, and selection. The primary advantage of DE over other optimization techniques lies in its ability to maintain population diversity and avoid premature convergence to local minima, making it particularly suitable for calibrating complex building energy models. It has good convergence when it is dealing with model parameter identification if the number of targeted parameters is not too large [64].

3.4.2. Fitness function definition

In the context of this study, each vector in the DE algorithm represents a potential set of values for the hard-to-determine parameters. The goodness of fit of each vector (i.e., how well the associated parameters enable the simulation model to reproduce actual building behavior) is evaluated using a fitness function based on the root-mean-square error (RMSE) between the model's predicted heating and cooling energy use and the observed data from the building (with equal weights). A lower RMSE indicates a better fit between the model's predictions and the observed data, guiding the DE algorithm in its search for the optimal parameter set. To mathematically represent, if P denotes the predicted energy use and O represents the observed data, the RMSE is defined as:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n}} \tag{18}$$

where n is the total number of data points. The aim of the DE algorithm, in this scenario, is to minimize this RMSE value, thereby deriving the most suitable values for the hard-to-determine parameters. The fitness function of the DE algorithm then can be described as:

$$F = \frac{1}{2}\sqrt{\frac{\sum_{i=1}^{n}(PCE_i - OCE_i)^2}{n}} + \frac{1}{2}\sqrt{\frac{\sum_{i=1}^{n}(PHL_i - OHL_i)^2}{n}} \tag{19}$$

where PCE, OCE, PHE, OHE represent predicted cooling energy consumption, observed cooling energy consumption, predicted heating energy consumption, observed heating energy consumption. Three different modes of time granularity are provided in this developed calibration procedure: hourly, daily, and monthly. Users can define at what time granularity they need the calibration scheme to be run. It is worth noting that while the DE-based calibration approach is employed here for existing building modeling, it is not used for the building predesign stage due to the lack of real-world observational data in this case.

3.4.3. Hyperparameter configuration and user flexibility

To further enhance the adaptability and effectiveness of the DE-based calibration approach, the engine offers users the flexibility to tweak the hyperparameters of the DE algorithm. Such customization

ensures that the calibration process can be finely tuned to meet the specific needs of individual building scenarios or to align with the preferences of the modeler. The key hyperparameters available for tuning include.

1. Mutation Rate - This parameter determines the extent to which the population's vectors are perturbed. It essentially controls the degree of exploration in the search space. Its default values are provided as a range from 0.7 to 1, allowing for variability in exploration intensity. For all validation cases in this study, F = 0.8 was used.
2. Recombination Rate - Set at a default of 0.8, this hyperparameter governs the fraction of the donor vector that's mixed with the target vector, influencing the information exchange rate between different potential solutions. This default value (CR = 0.8) was maintained across all case studies.
3. Population Multiplier - The population size is a critical factor in the DE algorithm's success. By default, it is set at a multiplier of 20, ensuring a sufficiently diverse set of initial solutions for the calibration problem. For a calibration problem with n parameters, the population size is 20n. For example, for a case study with 3 calibrated parameters (Cm, Cs, and internal wall U-value), the population size was 60.
4. Maximum Iteration - This parameter dictates the number of iterations the algorithm runs before terminating. Its default value is set at 100, balancing the need for a thorough search with computational efficiency. All calibration runs used this maximum iteration limit.
5. Convergence Tolerance - Defined at a default of 0.01, this hyperparameter establishes the threshold for convergence. If the change in fitness values between consecutive iterations is less than this threshold, the algorithm deems the solution as having converged. Convergence was achieved when the improvement in fitness function between consecutive iterations fell below 0.01, typically occurring within 40–60 iterations for the case studies presented.

Table 2
Information collection for model parameter calibration.

Parameter	Range		Parameter	Range	
Building Heat Capacity (m)	40000	400000	Internal Floor Material U-value (W/m^{2}K)	1	10
Building Heat Capacity (s)	40000	400000	Air Infiltration Rate (h^{-1})	−1	
Effective Mass Area (m^2)	0.1	5	Air Infiltration Style	−1	
External Wall Material U-value (W/m^{2}K)	−1		Lighting Load (W/m^2)	−1	
External Wall Material Absorptivity	−1		Plug Load (W/m^2)	−1	
Internal Wall Material U-value (W/m^{2}K)	1	6	Heating COP	−1	
Window Material U-value (W/m^{2}K)	−1		Cooling COP	−1	
Window Material Emissivity	−1		Heating Supply Air Temperature (°C)	−1	
Window Material SHGC	−1		Cooling Supply Air Temperature (°C)	−1	
Roof Material U-value (W/m^{2}K)	−1		HVAC Distribution Loss Coefficient	−1	
Roof Material Absorptivity	−1		Heating Temperature Setpoint (°C)	−1	
External Floor Material U-value (W/m^{2}K)	−1		Cooling Temperature Setpoint (°C)	−1	
Hyperparameters of DE					
Mutation Rate	0.7	1			
Recombination Rate	0.8				
Population Multiplier	20				
Maximum Iteration	100				
Convergence Tolerance	0.01				

The DE algorithm follows the rand/1/bin strategy, where mutation uses the formula: Vi,g+1 = Xr1,g + F(Xr2,g - Xr3,g), where Vi,g+1 is the mutant vector, Xr1,g, Xr2,g, Xr3,g are three randomly selected distinct population members, and F is the mutation rate. Binomial crossover is applied with probability CR. Parameter bounds for the education building calibration are specified in Table 2, with the search space constrained by physical feasibility ranges derived from building audit data and material property databases."

The process becomes more versatile because users are allowed to modify these hyperparameters. The defaults provided here are derived from extensive testing and have reliably performed across a wide range of building types, but the ability to adjust them provides some robustness of the engine even in unusual or challenging calibration environments. To make the calibration procedure more versatile and applicable to diverse building types and scenarios, users are given the flexibility to choose which modeling parameters they want to calibrate. Table 2 shows an example of the modeling parameters available for calibration in this version. The calibration information shown in Table 2 is used for the calibration of the case study education building that will be described in Section 3.6.

The program takes the value of "-1″ as a unique identifier that signals certain parameters won't be included in the calibration procedure. That allows users to remove parameters you know or perceive as unimportant for calibration. If users want to calibrate a parameter, they can give a range of reasonable value within which the DE algorithm searches the optimal value.

Furthermore, to maintain clarity, organization, and ease of access, all this calibration information—both for modeling parameters and DE hyperparameters—is stored in a structured text-based file, specifically named “calibration_parameter.txt”. The format shown here is an intuitive way to review and modify the calibration parameters, enabling the user to quickly prepare the simulation for calibration and ensuring reproducibility and transparency throughout the calibration process. This design choice ensures that the calibration settings integrate seamlessly with other computational tools, and that users have a straightforward path to audit, review, or modify the calibration settings.

3.4.4. *Calibration workflow*

The calibration process follows a systematic workflow designed to optimize hard-to-determine parameters while preserving physics-based calculations for well-characterized building properties. The procedure consists of six main steps as illustrated in Fig. 5.

Step 1: Data Preparation and Parameter Selection - Users specify which modeling parameters require calibration by defining search ranges in the 'calibration_parameter.txt' file, with '-1′ values indicating parameters to be held constant at their physics-based values. Measured energy consumption data (heating and cooling) is prepared at the selected time granularity (hourly, daily, or monthly). For the education building case study, three parameters were selected for calibration (Cm, Cs, internal wall U-value) while all other parameters were calculated from building audit data.

Step 2: Initial Population Generation - The DE algorithm generates an initial population of candidate parameter sets using Latin Hypercube Sampling within the specified bounds. Population size is determined by the population multiplier (default 20) times the number of calibrated parameters, ensuring adequate search space coverage.

Step 3: Forward Simulation - For each candidate parameter set, RCBldEng performs a complete annual simulation using the forward modeling framework described in Section 3.2, with calibration parameters overriding their default physics-based values while all other parameters remain unchanged.

Step 4: Fitness Evaluation - The fitness function (Equation (19)) calculates the weighted RMSE between simulated and measured energy consumption for both heating and cooling. Lower fitness

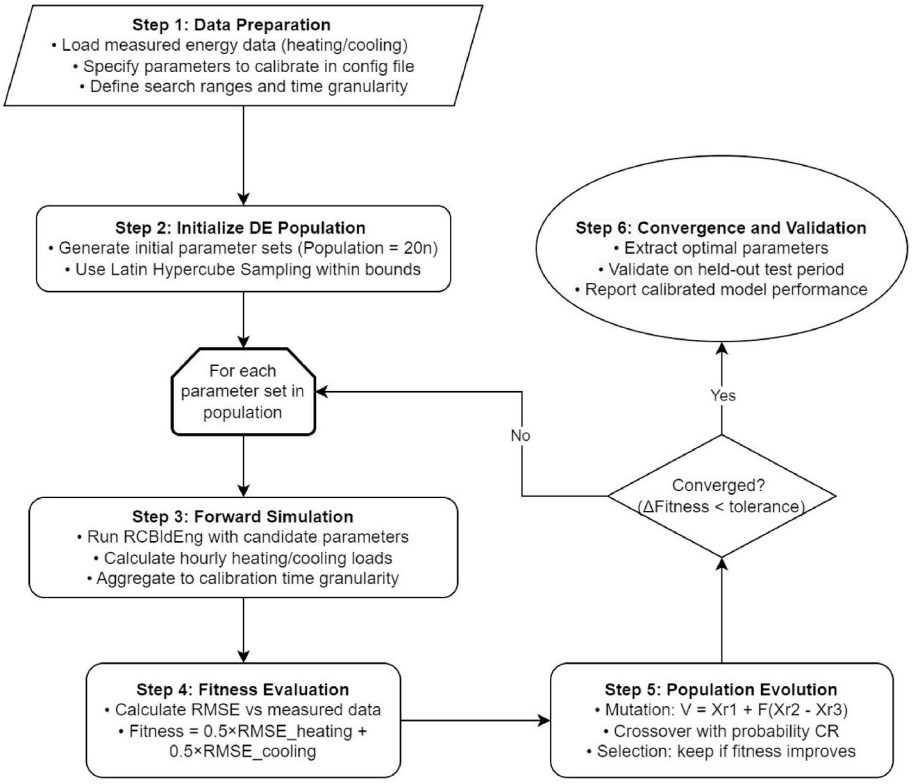


Fig. 5. Flowchart of the automated calibration workflow using Differential Evolution algorithm for parameter identification of existing building cases in RCBldEng.

values indicate better agreement between model predictions and observations.

Step 5: Population Evolution - The DE algorithm applies mutation ($V_{i,g+1} = X_{r1,g} + F(X_{r2,g} - X_{r3,g})$), crossover with probability CR, and selection operations to evolve the population toward optimal parameter values. The algorithm tracks convergence by monitoring the improvement in best fitness value between consecutive generations.

Step 6: Convergence and Validation - When the fitness improvement falls below the convergence tolerance (default 0.01) or the maximum iteration limit is reached, the algorithm terminates and returns the optimal parameter set. The calibrated model is then validated against a held-out test period to assess predictive performance on unseen data.

This automated workflow requires no user intervention beyond initial parameter range specification, enabling consistent and reproducible calibration across different building types and datasets.

3.5. Simulation of prototype buildings

3.5.1. Prototype building models

The validation and performance assessment of the developed RC-based simulation engine are grounded on two prototype buildings: a detached house and a medium-sized office building as shown in Fig. 6. The EnergyPlus models of two prototype buildings from the U.S. Department of Energy (DOE), specifically from the commercial and residential prototype model series, are remodeled in RCBldEng, ensuring the representation of climate-dominant and internal-load dominant building types [65]. Both prototype buildings adhere to the ASHRAE 90.1–2013 standard, underlining their commitment to energy efficiency and sustainable design. ASHRAE 90.1–2013 is a well-established standard in the building industry, widely adopted for its benchmarks and guidelines on energy-efficient design and practices [66]. Adherence to this standard ensures that the case study buildings reflect contemporary building design practices, especially in the context of energy efficiency. The occupancy and building use schedule have been visualized in Fig. 7, and the indoor cooling and heating setpoint schedule can be found in Fig. 8.

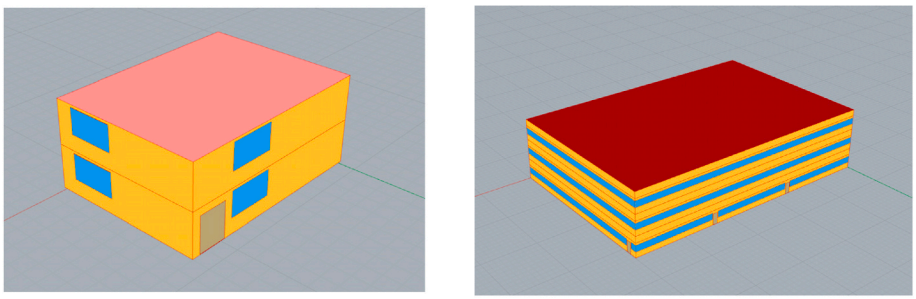


Fig. 6. The two DOE EnergyPlus prototype building models.

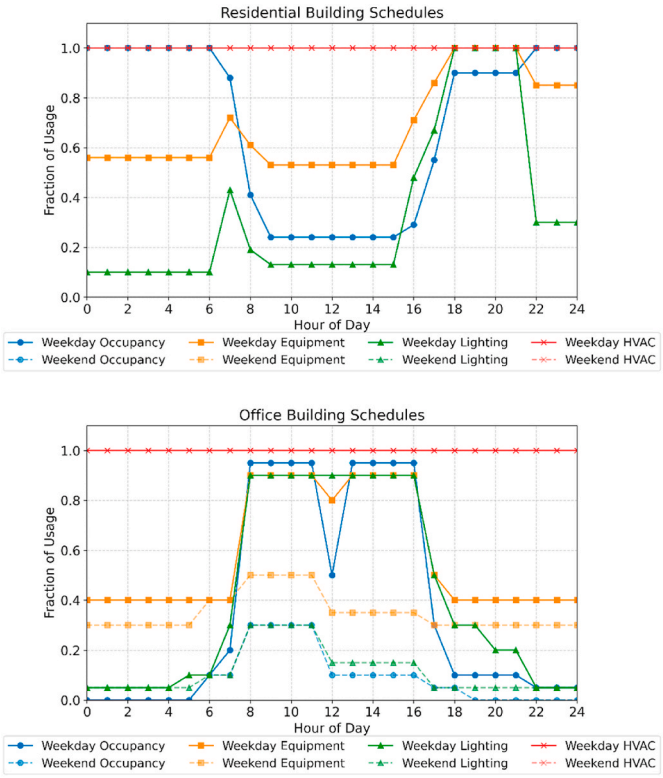


Fig. 7. Building occupancy and use schedule for the two prototype buildings.

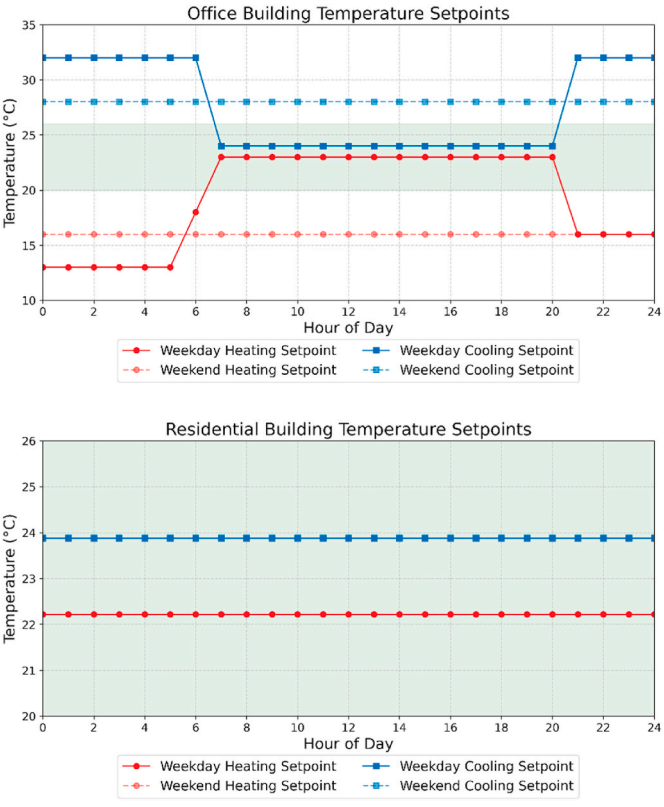


Fig. 8. Indoor cooling and heating setpoint schedule of the two prototype buildings.

The building simulation uses IWEC (International Weather for Energy Calculations) weather data in EPW format for Shanghai, China. The IWEC dataset, developed by ASHRAE, represents typical weather conditions derived from long-term hourly observations and is comparable to the TMY (Typical Meteorological Year) datasets used in the United States. Shanghai is in the ASHRAE 3A climate zone, which has a warm and moist climate, hot summers and mild winters. The IWEC data here provides the annual baseline for the developed RC model simulation engine, which encapsulates several thermal load conditions across different seasons, and consequently presents a complete assessment. As a reference simulation engine for building energy simulation, EnergyPlus [10], the popular and well-known building energy simulation software is used, to compare and to establish a performance baseline. The version of EnergyPlus that was used for this study is 9.5. It is a good yardstick against which the RC based simulation engine outcomes can be gauged. This comparative approach provides an understanding of the fidelity and reliability of the developed RC based engine. The proposed four RC model configurations are considered, 4R1C, 6R1C, 7R1C and 7R2C, for a more complete comparison. The structure of this assessment can facilitate a thorough analysis of the developed engine's accuracy, robustness, computational cost, and versatility with respect to various modeling complexities.

3.5.2. *Cross-climate zone validation*

We further conducted the testing for the proposed RC modeling methods by integrating three cities with different climates into the assessment of two prototype buildings. For all climate zones in this analysis, we used IWEC (International Weather for Energy Calculations) weather data in EPW format. The selected climate zones for this study encompass a wide variety of climatic environments to assess the methodology.

1. Shanghai (ASHRAE Climate Zone 3A): This climate comprises of hot summers alongside mild winters due to its warm and moist environment. Significant seasonal changes can occur in this climate region which requires both heating systems and cooling systems throughout the entire year.
2. Guangzhou (ASHRAE Climate Zone 1A): Hot and humid climate with minimal seasonal variation. The climate zone has cooling-dominated conditions that create higher cooling requirements than heating requirements during all periods of the year.
3. Shenyang (ASHRAE Climate Zone 7): This climate zone faces severe cold conditions which extend across long heating seasons while remaining short in summer periods. At this climate zone the heating necessities outweigh cooling requirements due to its lengthy winter season.

We evaluated the performance of four RC model configurations from 4R1C model all they way to 7R2C model across the three climate zones using TMY weather data for both residential houses and medium office buildings. Each climate zone received analyzes of model accuracy based on R^2 values plus NRMSE (normalized root mean squared error) for hourly load predictions along with NRMSE for daily peak load predictions.

3.5.3. *EnergyPlus validation configuration*

To ensure fair comparison between RCBldEng and EnergyPlus simulations, consistent modeling assumptions and calculation settings were implemented. The EnergyPlus simulations used the following key configurations:

Timestep and Convergence Settings.

- Simulation timestep: 4 timesteps per hour (15-min intervals) for EnergyPlus; 1-h timestep for RCBldEng
- Heat balance algorithm: Conduction Transfer Function (CTF) method
- Zone air heat balance algorithm: Third Order Backward Difference

•Convergence limits: 0.04 °C for zone temperature, 0.4 W for zone load

Solar and Shading.

- Solar distribution: FullInteriorAndExterior
- Shading calculations: AverageOverDaysInFrequency with 20-day frequency
- Surface convection algorithm: TARP (Thermal Analysis Research Program) for interior and exterior surfaces

HVAC System Modeling.

- Ideal loads air system for load calculations (no detailed HVAC system modeling to enable direct thermal load comparison)
- Heating and cooling availability: Based on schedules shown in Fig. 7
- Temperature setpoints: As specified in Fig. 8

Weather Data Processing.

- IWEC weather files for Shanghai, Guangzhou, and Shenyang in EPW format
- Identical weather data used for both EnergyPlus and RCBldEng simulations
- Ground temperatures calculated using Kusuda-Achenbach model in EnergyPlus; monthly average ground temperatures extracted and used in RCBldEng

The RCBldEng simulations replicated these thermal boundary conditions using the forward modeling parameters calculated from the same building geometry, envelope properties, internal loads, and schedules. This approach isolates the impact of the RC model simplifications from differences in HVAC system modeling or calculation algorithms.

3.6. Test on existing building

The developed simulation engine RCBldEng was also tested on an existing building to evaluate the efficacy and accuracy of its application to real world scenarios. An education building in Philadelphia, U.S., was selected as a testbed for this purpose. Local weather station on site has been installed for building simulation with onsite hourly weather data. This building, which houses laboratories, classrooms, lecture halls, and offices, is an appropriate candidate for a comprehensive evaluation since the thermal loads and the energy usage pattern of each function are different.

The first step was the collection of known (forward) modeling parameters. Essential parameters were extracted leveraging an energy audit of the building. In this case, the data gathered from energy audit becomes the modelling parameters to determine the modelling inputs and the test was based on real world situation. Nevertheless, although energy audits are by their very nature very detailed, some modeling parameters are inherently difficult to determine with high precision. It is these hard-to-determine parameters which serve an essential role in capturing the intricate building thermal behavior, and overall energy efficiency. To address this challenge, the study used available designed calibration procedures as described in Section 3.4. The DE algorithm can be used to identify and optimize these elusive parameters underpinned by this procedure. The objective is for the RC model to be not only informed but also fine-tuned with the building specific characteristics derived from the calibration. The results from this test can provide some tangible evidence of the engine's capability and its applicability to non-trivial simulation of building energy situations.

3.6.1. Known modeling parameters (forward modeling)

Table 3 provides an exhaustive summary of the building envelope, detailing both the opaque and transparent segments. The data showcases the surface areas of each envelope component, segregated by orientation – South (S), East (E), North (N), West (W), and Roof. Notably, it also offers insights into below-grade opaque sections, an indispensable component when considering subterranean heat gains or losses.

Table 3
Summary of modeling information of the education building envelope in each orientation.

Orientation	Opaque envelope (m^2)	Window (m^2)	Below Grade Opaque (m^2)
S	343.6	275.7	121.9
E	632.7	422.5	159.6
N	400.4	221.3	118.1
W	604.5	457.1	134.0
Roof	2111.6	0.0	0.0

The building's functional use and its consequent energy consumption patterns are deeply influenced by its occupancy and usage schedule. depicts the building's typical weekday and weekend schedules for various aspects: occupancy, equipment usage, lighting, and HVAC operations. As described in Section 3.2, the modeling information is fed into the engine via a tailor-designed structured text file, ensuring that the model is both accurate and reflective of the building's actual operational conditions. The occupancy and building use schedule for the existing education building is depicted in Fig. 9.

3.6.2. Calibration of hard-to-determine modeling parameters (inverse modeling)

Hard-to-determine modeling parameters are those difficult to determine due to their complex interactions, latent effects, or the absence of direct measurement methods. The inverse modeling procedure, the solution of adjusting the unknown parameters using observational data and simulation results, can be applied as a solution. Section 3.4 presents a calibration scheme for the identification of these parameters, which are effective. The calibration process adjusts these hard to determine parameters such that the simulation outcomes for the model match observations of how the building performs in terms of energy. Most parameters were derived directly from the energy audit for the case of the education building, particularly the thermophysical properties of the structure and envelopes. This left a handful of crucial parameters that needed identification via the calibration scheme. The calibration information of the following parameters for the education building case study are reflected in Table 2:

Thermal Capacity of the Building: The ability of the building's thermal energy storage can be described by two distinct parameters: Cm for the 1st order RC model, Cs for the 2nd order RC model. Thermal capacities,

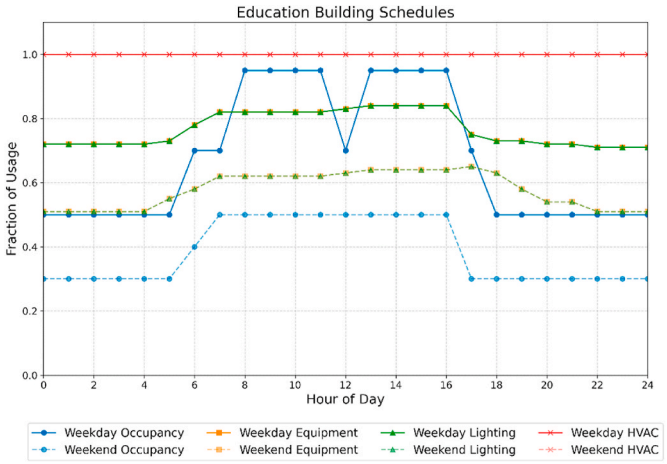


Fig. 9. Occupancy and building use schedule of the education building.

as difficult to measure directly, play a key role in the time a building takes to heat up or cool down.

U-value of Internal Wall: The U value is a description of how well a certain material is an insulator. Internal wall U values, although important to energy efficiency, are not always easily available, unlike external components.

U-value of Internal Floor: presents how much heat transfers through the floor structure is determined by its U-value, just like the internal wall. This becomes crucial in multistoried buildings.

It is noteworthy that although system efficiency of heating and cooling sources is typically a key calibration parameter, it is not required for this building. The education building receives district steam and chilled water with directly metered consumption, eliminating the need to calibrate source efficiency. Only distribution losses are considered in the model, as the metered data already accounts for the source energy.

3.6.3. Data processing and quality control

In this research, measured energy data for the education building was collected from building-level meters for a consecutive year. Hourly steam consumption for heating was converted to thermal energy. Chilled water energy was directly measured by BTU meters and validated against supply/return temperature differential and flow rate. On-site weather station measurements included dry-bulb temperature, relative humidity, wind speed, and global horizontal irradiance at hourly resolution with 98.5% data completeness.

A double MAD (median absolute deviation) method was used to recognize and remove outliers in the building energy use data. One of the common methods to identify and remove outliers in one-dimensional data is to mark as a potential outlier any point that is more than two standard deviations from the mean. However, the presence of outliers is also likely to have a strong effect on the mean and the standard deviation, making this technique unreliable. So, it is recommended to use a measure of distance that is robust against outliers. MAD is good in dealing with this kind of problem because it uses the mean absolute deviation from the median. However, MAD outlier recognition requires that the data distribution not be skewed or asymmetric. It works well with, for example, a symmetric statistical distribution like normal distribution, or uniform distribution. For asymmetric distributions, double MAD should be used. This is a synergy of two MAD methods: (1) the mean absolute deviation from the median of all points less than or equal to the median, and (2) the mean absolute deviation from the median of all points greater than or equal to the median. The former is used to calculate the distance from the median of all points less than or equal to the median; the latter is necessary to calculate the distance for points that are greater than the median. By using this double MAD-based outlier removal method, it is possible to recognize and remove the outliers that exist in the building energy use data. The percentage-based outlier removal method, which screens outliers by the top percentage of biased points, was compared with the double MAD-based method in Fig. 10. Percentage-based removal, most of the time, removes too many incorrectly identified outliers that are actually valid data points. The red points shown in the figure above are the outliers detected by the two different methods. The comparison between the two on campus building's energy use data before and after outlier removing are also plotted below. It can be seen that the double MAD based method is successful at handling the outliers in the building energy use data.

4. Results and analysis

4.1. Simulation of the prototype buildings

For an effective and comprehensive understanding of the accuracy and applicability of the RC-based models introduced, their performance was juxtaposed with the renowned existing building energy simulation tool, EnergyPlus. For the two prototype buildings, the performance of the proposed models with and without interzonal thermal coupling are compared with EnergyPlus simulation results and discussed in the following sections.

4.1.1. The simulation results of the model without interzonal thermal coupling

The hourly heating and cooling load simulation results were first analyzed to explore the applicability of the 4R1C model without interzonal thermal coupling, as shown in Fig. 11. The cooling load was calculated by EnergyPlus and shown by the blue line and by the 4R1C model predictions shown by the orange line. The comparative results between the EnergyPlus predictions and the 4R1C model for a detached house and an office building are shown in this figure. The trend delineated by the EnergyPlus simulations was very well reproduced by the 4R1C model results for the detached house. The reason for this reasonably close adherence is the simpler architectural design of the detached house, which has only three zones. The spatial segmentation of such a simple model is devoid of complexity so that the interzonal effects are minimal, and the 4R1C model produces quite accurate hourly heating and cooling load predictions without accounting for interzonal thermal transfer.

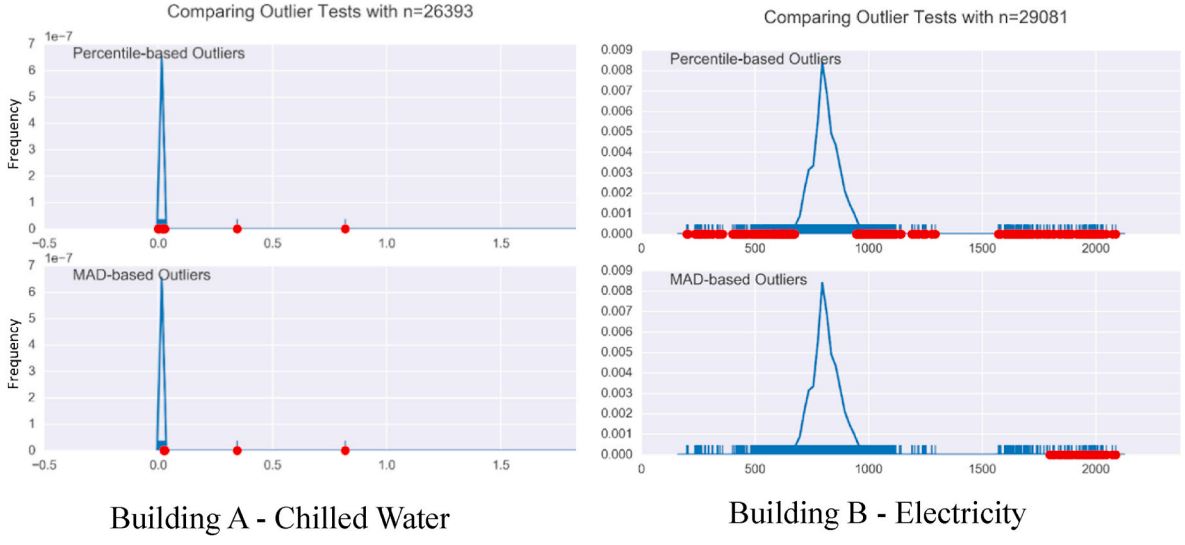


Fig. 10. Comparison of percentile-based and MAD-based outlier removal methods.

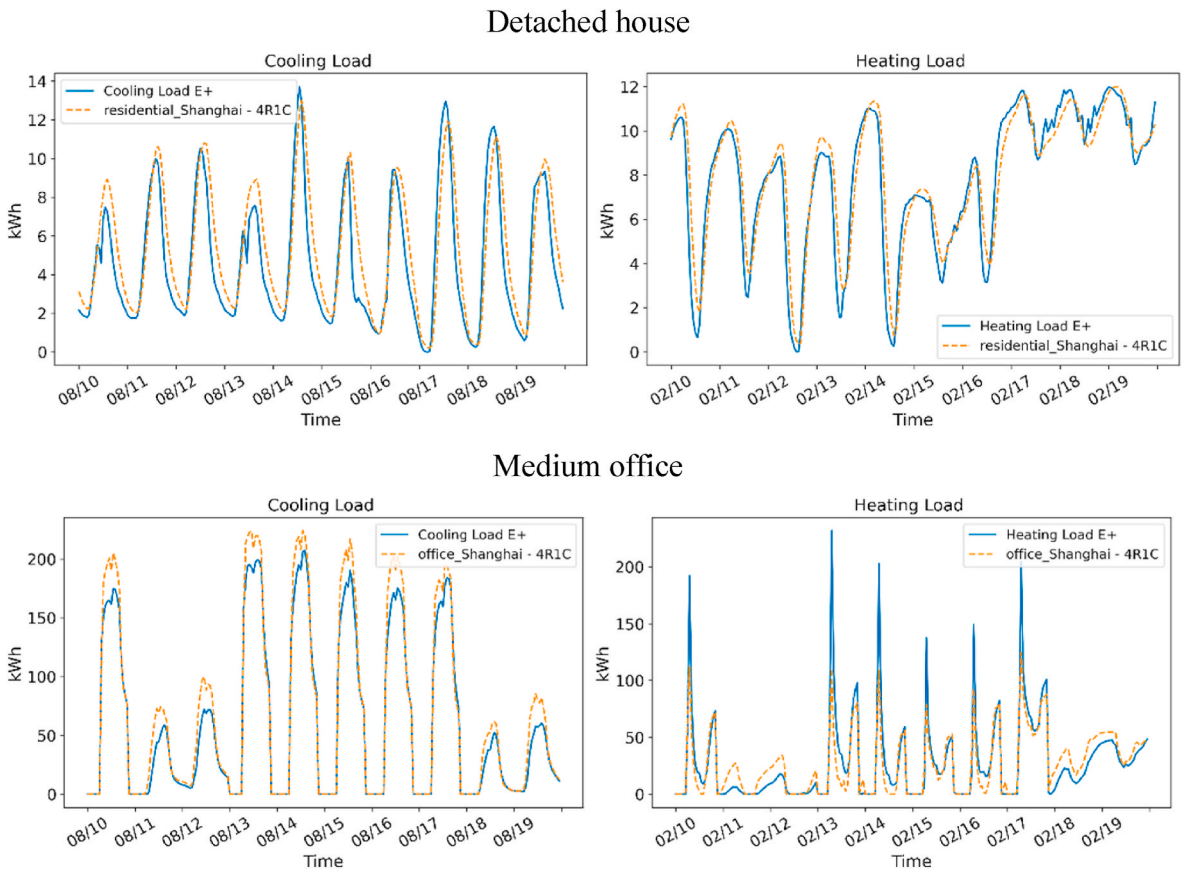


Fig. 11. Hourly simulation results of EnergyPlus and the 4R1C model for the two prototype buildings.

However, when analyzing the office building, a building that is inherently more complex structure with more than 10 zones, the scenario changes. The 4R1C model simulation results show significant discrepancies, which are consistent with this heightened complexity. In particular, the 4R1C model is seen to have interzonal thermal coupling that is absent. Interzonal thermal transfer biases sum together to lead to significant overprediction of the cooling load by the 4R1C model in summer months. In particular, this overestimation is particularly pronounced during peak cooling periods. On the other hand, during winter months, the model exhibits a tendency to underestimate heating loads, and the variance turns out to be most significant during the peak heating periods. The bifurcation in the results between the detached house and office building highlights the important role that interzonal thermal transfers play, particularly in buildings containing multiple zones. This provides a critical indication of the limitations of the 4R1C model when applied to complex structures and highlights the importance of models that are capable of representing interzonal thermal dynamics in such situations.

4.1.2. *The simulation results of the models with interzonal thermal coupling*

The hourly simulation results with internzonal thermal coupling over a spread of ten randomly selected continuous days during winter, summer, and transitional seasons are further studied in Fig. 12. When comparing the 7R2C model to the 6R1C and 7R1C models, it becomes clear that the 7R2C model outperforms the other models in accuracy. However, this superior performance is subtly shown and the differences between models are almost imperceptible using time series plots. Thus, analytical efforts were continued in the scatter plots shown in Fig. 13 that compare the hourly predictions of EnergyPlus against each of the RC models throughout the year. These plots also support a noticeable increase in prediction accuracy as we move from the 4R1C model to the more complicated 7R2C model, especially for the residential building prototype.

The office building exhibits markedly different heating and cooling load profiles compared to the residential house, primarily due to differences in surface-area-to-volume ratio, occupancy schedules, envelope properties, and internal load intensity. Fig. 14 illustrates the hourly simulation results for the office building using EnergyPlus and the RC models (6R1C, 7R1C, and 7R2C). The 7R2C model, with its intricate structure, emerges as the best performing model in terms of minimizing biases relative to EnergyPlus. Its capability in predicting heating and cooling loads is especially conspicuous during transitional seasons. Complementing this observation, Fig. 15, the scatter plot reflecting the simulation accuracy of the models for the office buildings, demonstrates hierarchical performance. While both the 4R1C and 6R1C models exhibit less persuasive performance relative to the 7R models for hourly heating and cooling load predictions, the 7R2C model comfortably takes the lead over the 7R1C.

4.1.3. *Cross-climate zone performance analysis*

To assess the robustness and generalizability of the proposed RC models, we conducted validation across three distinct climate zones: Shanghai (3A, warm-moist), Guangzhou (1A, hot-humid), and Shenyang (7, severe cold). Table 4 presents a comprehensive validation results for both prototypical residential house and medium office building across these climate zones.

Our findings demonstrate universal applicability because model accuracy shows steady growth when complexity increases across each climate region. Across all climates the residential building cooling load predictions show R^2 values increase from 0.88 to 0.90 for the 4R1C model to 0.98 for the 7R2C model. The R^2 values of the 4R1C model for the office building rise from 0.90 to 0.93 before reaching 0.99 when using the 7R2C model. The steady improvement in prediction results demonstrates that complex RC models deliver benefits irrespective of

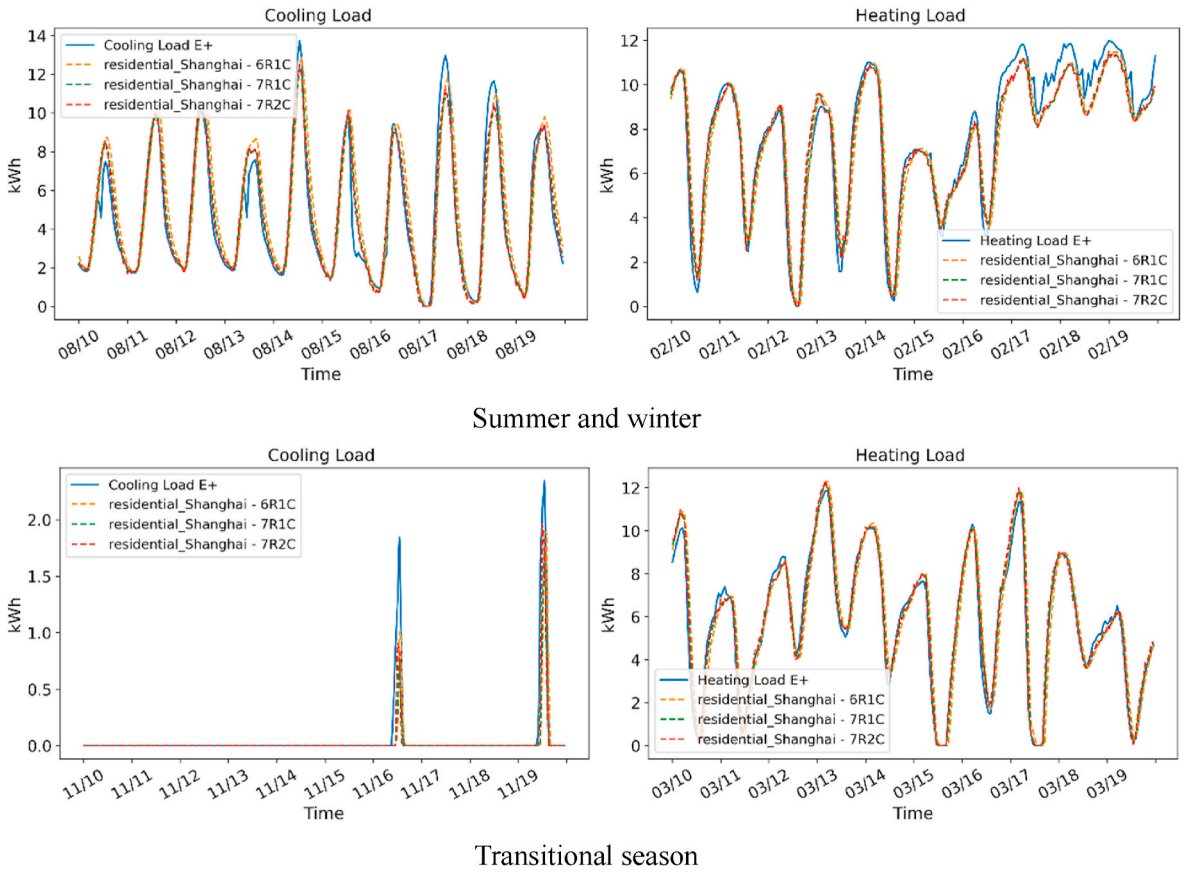


Fig. 12. Hourly simulation results of EnergyPlus and the 6R1C, 7R1C, and 7R2C model for the residential house.

the selected climate zone.

Certain trends in terms of model performance that are related to climate variation become visible during this validation process. The hot and humid conditions of Guangzhou produce superior cooling load prediction accuracy among all climate zones and particularly affect the performance of office buildings. The cooling-dominated building conditions and stable daily temperature patterns explain this observation. The heating load predictions for the office building in Guangzhou display lower correlation values compared to other testing areas, mainly because of the building's limited heating usage period.

The 7R2C model delivers best heating load prediction outcomes in Shenyang's severe winters across all building types as it achieves the highest R^2 values across every climate zone. The dual capacitance mechanism proves essential for capturing thermal mass effects which occur during extended heating periods with major differences between indoor and outdoor temperatures. The increased thermal capacitance in the 7R2C model yields its greatest prediction benefits during heating load forecasting in the Shenyang climate zone.

The cross-climate validation showed how different climatic conditions affect which RC model configurations prove optimal. Buildings located in severe cold regions with noticeable seasonal changes will gain significant accuracy improvements when using the 7R2C model to forecast heating loads. The 7R1C model also shows good accuracy and operational efficiency for hot and humid climates when focused on cooling-dominated building operations. The 7R1C model has achieved good performance measures in Shanghai's mild climate zone where heating and cooling durations exist while the 7R2C model provides minimal additional benefits. The identified findings enrich our model selection process by adding climate-condition adaptability to the assessment framework which includes building design characteristics and simulation end goals we previously introduced.

4.2. Simulation of the existing case study building

For the task of simulating energy use in the real-world educational structure, the 7R2C model was the most suitable according to the results. This decision was influenced by the building's distinct attributes, especially its thermal capacity, internal load intensity, and the distribution of thermal zones. Given the multifaceted nature of the building edifice, a simplified approach was imperative for efficiency in the RC-based modeling. Consequently, the building's layout was delineated into eighteen primary thermal zones. This consolidation was achieved by merging zones that exhibited analogous functionalities and occupancy dynamics. These thermal zones, coupled with the detailing of the building's thermophysical characteristics, form the backbone of the simulation process. Table 5 lists these specifics, ranging from the net building floor area to the calibrated modeling parameters, Cm and Cs, which have been derived through inverse modeling. The subsequent sections will provide a closer look at these simulations, elucidating the 7R2C model's performance against the educational building's real-world energy dynamics.

Fig. 16 provides an illustrative comparison between the energy use predicted by the 7R2C model and the actual metered data for the educational building. This juxtaposition is particularly significant for understanding the model accuracy and precision in replicating real-world energy use patterns, especially concerning heating and cooling. As elucidated in the preceding sections, the building's heating and cooling provisions are sourced from the local district energy provider, specifically in the forms of steam and chilled water. Looking into the winter season's data, it is evident that the 7R2C model's predictions for steam usage align well with the actual metered consumption. However, the situation is slightly different when examining the chilled water usage for cooling. While the general trend of the predicted values aligns with the metered data, there are certain deviations with metered hourly

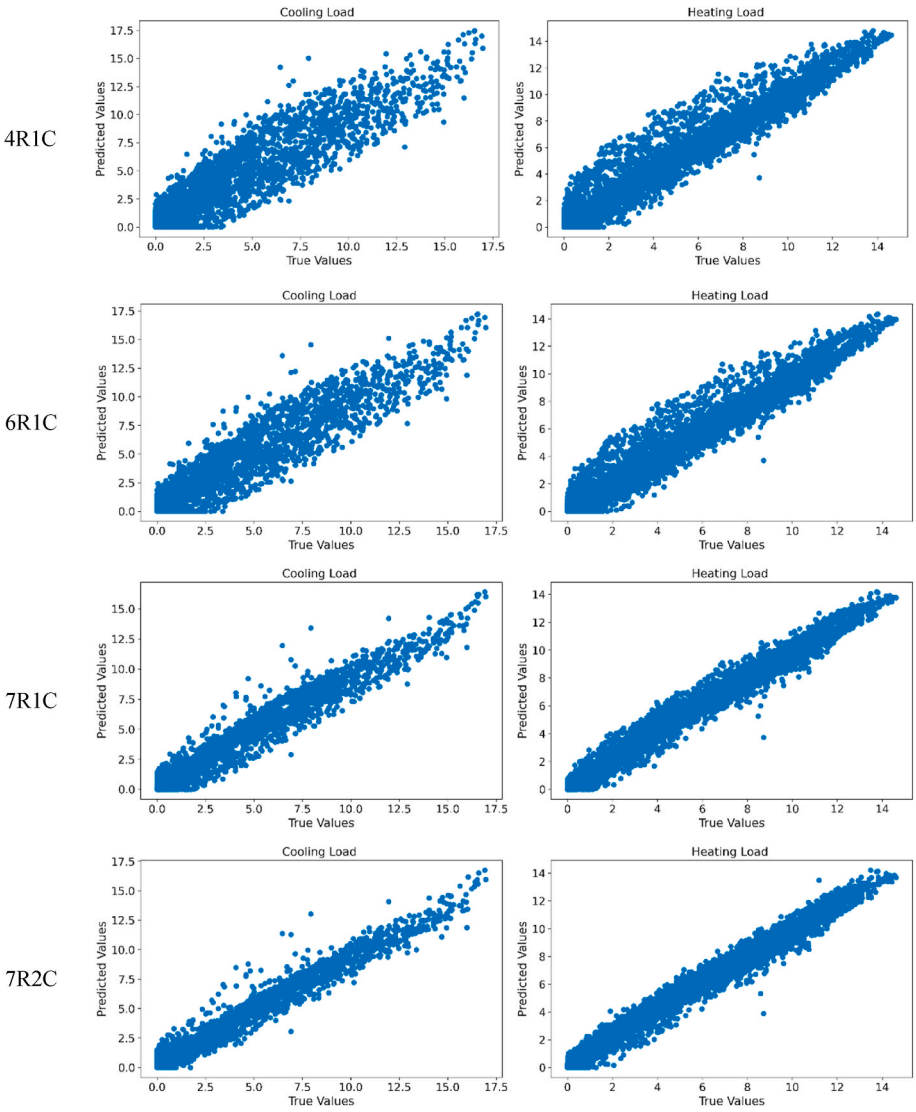


Fig. 13. Scatter plots of simulation accuracy of the 4R1C, 6R1C, 7R1C, and 7R2C model for the residential buildings (True value (kWh): EnergyPlus, Predicted value (kWh): RC model).

energy use. These discrepancies are more pronounced in predicting cooling energy use.

The challenge of simulating real-world energy consumption in buildings can be filled with intricacies primarily due to the uncertainties associated with operational conditions and building occupancy. The pronounced seasonal discrepancies stem from varying operational uncertainties between heating and cooling seasons. During summer (July 15–30), the education building experiences irregular occupancy due to reduced course schedules and variable laboratory usage, introducing substantial uncertainty in internal heat gains (±40% from nominal values). Occupant-controlled window opening during mild conditions creates unmodeled natural ventilation, while dynamic blind adjustments alter solar heat gains beyond the simplified shading model's capability, contributing to the cooling underprediction. Conversely, winter heating (January 15–30) exhibits better accuracy due to predictable semester schedules, minimal window operation, and reduced solar variability. The heating overprediction suggests the model does not fully capture thermostat setback practices during unoccupied periods and district heating system response delays. These seasonal variations highlight the fundamental trade-off in RC modeling between capturing primary thermal dynamics and accepting bounded uncertainty in occupant-driven variables, making the approach suitable for monthly energy trends and retrofit analysis rather than hour-by-hour operational control.

As depicted in Fig. 17, a comparative analysis of the aggregated monthly energy usage for both chilled water and steam presents insightful conclusions. The side-by-side representation of actual consumption and the 7R2C model's predictions serve to underline the model's efficacy in mirroring monthly energy use trends. The accuracy with which the proposed RC model simulates monthly energy consumption is good. Despite the intricacies that might influence a building's energy profile, the model manages to capture the overarching trends and patterns. This level of precision, especially when applied at a monthly aggregate level, provides a promising indication of the model's robustness.

The ability of the 7R2C model to simulate monthly energy consumption with such precision underscores its potential utility in broader applications. For instance, when attempting to optimize a building's energy consumption, designers and engineers are often faced with innumerable combinations of physical parameters. Each of these combinations can influence the building's energy profile in unique ways. The RC model's proficiency in accurately predicting monthly energy trends showcases its potential as a powerful tool in such optimization processes.

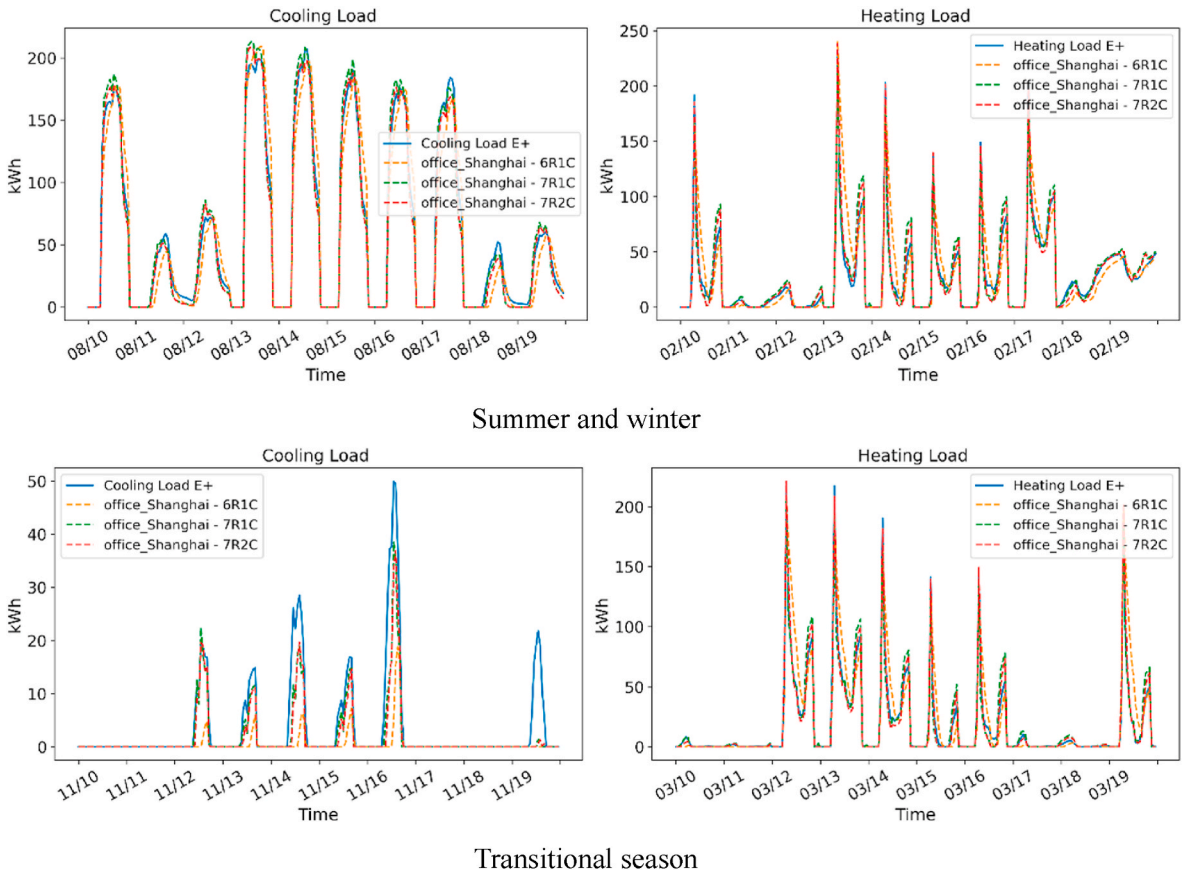


Fig. 14. Hourly simulation results of EnergyPlus and the 6R1C, 7R1C, and 7R2C model for the office buildings.

To provide a comprehensive quantitative assessment of the model performance, Table 6 summarizes the statistical metrics for the education building case study at hourly and monthly time granularities.

Table 6 demonstrates that the calibrated 7R2C model achieves good predictive accuracy with monthly NRMSE of 17.85% for heating and 10.99% for cooling, and minimal annual bias (MBE <4%). At hourly resolution, NRMSE increases to 18.04–21.46%, reflecting challenges in capturing short-term operational variations. The hourly MBE reveals systematic directional biases: 11.37% heating overprediction and 11.38% cooling underprediction, consistent with the operational uncertainty discussion in this section regarding irregular occupancy patterns and thermostat practices. Despite these hourly-level biases, the low monthly MBE values demonstrate that errors largely offset over coarser granularity. The higher cooling RMSE indicates greater prediction variability due to complex interactions between solar gains, internal loads, and occupant-controlled shading. Overall, these metrics validate that the model provides reliable predictions for retrofit analysis and energy management, particularly for monthly and seasonal assessments.

5. Discussions

5.1. Comparative analysis of different RC models and their selection

In this study, a comparative analysis of RC model configuration is performed to gain insights into the relation between model complexity and simulation accuracy. The introduction of interzonal thermal coupling brings marked improvement, which is most apparent in the transition from the 4R1C model to more sophisticated ones. The 7R2C model shows superior accuracy than the 4R1C model for residential buildings, with R^2 values of 97.81% and 98.74% for cooling and heating loads respectively, compared to the 89.51% and 96.12% R^2 value of the 4R1C model. The improvement aligns with Vivian et al.'s study [46]. Office buildings exhibit inherently higher complexity of thermal interactions and thus the performance differentiation is more pronounced. With R^2 values for cooling load prediction up to 98.78%, 7R2C model demonstrates the importance of simulating interzonal thermal dynamics in complex building structure.

For peak cooling load predictions in residential buildings, the 7R2C model demonstrates superior stability (NRMSE of 4.47%) due to its dual capacitance representation of thermal storage effects. While the 7R1C model achieves comparable accuracy in many situations, thermal mass representation remains a crucial factor in model selection. In particular, the increase in the number of capacitors in the 7R2C model was found to be especially advantageous when dealing with buildings with significant thermal mass, as in the case of the office building. This is in line with previous findings by that thermal mass modeling is needed to make accurate building energy predictions.

The influence of interzonal thermal coupling varies substantially by building typology. The simpler 6R1C model is able to provide viable accuracy for residential prototypes with three thermal zones. Yet in office buildings with many zones, more sophisticated interzonal thermal coupling in the models 7R1C and 7R2C is needed for reliable predictions. The relationship between building complexity and model sophistication required can add practical guidance for model selection in different applications as per this study. This work provides a feasible framework for model selection based on the building characteristics and simulation requirements. The 6R1C model tends to produce sufficient accuracy for simple residential structures with few thermal zones. The 7R2C model can capture thermal dynamics with enhanced capability in predicting the peak load of complex commercial buildings as well as applications where the peak load prediction needs to be predicted with high precision. This selection guidance will help the practical application of RC modeling with building energy simulation by providing a balance between model sophistication and practicality.

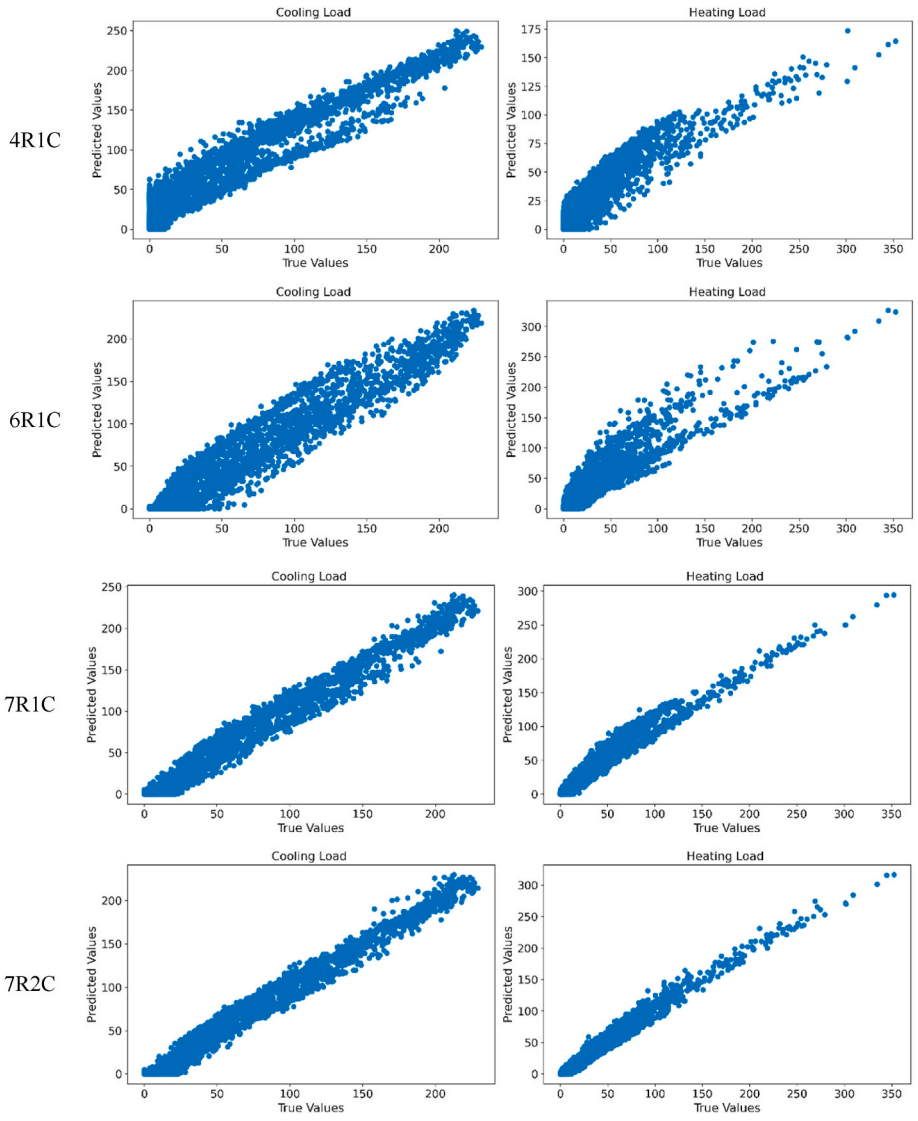


Fig. 15. Scatter plots of simulation accuracy of the 4R1C, 6R1C, 7R1C, and 7R2C model for the office buildings (True value (in kWh): EnergyPlus, Predicted value (kWh): RC model).

To provide more specific guidance for model selection, Table 7 presents a decision matrix based on building characteristics and simulation requirements. This matrix synthesizes the quantitative performance metrics presented in Table 4 with computational efficiency considerations.

The selection criteria are derived from the following findings.

1. *Building zoning complexity:* Buildings with more than 4 thermal zones show at least 3–4% improvement in R^2 values when using 7R1C or 7R2C models compared to simpler models.
2. *Thermal mass consideration:* Buildings with high thermal mass (concrete structures, massive floors/walls) benefit from the dual capacitance representation in the 7R2C model, improving NRMSE for peak load predictions by approximately 0.5–1.5% compared to the 7R1C model.
3. *Accuracy requirements:* For applications requiring heating/cooling load predictions with NRMSE <3%, at least the 7R1C model should be used for complex buildings. For NRMSE <5%, the 6R1C model is typically sufficient for residential buildings.
4. *Seasonal considerations:* During transition seasons, the 7R2C model reduces NRMSE for cooling load predictions by 0.8–1.2% compared to the 7R1C model for complex commercial buildings.
5. *Computational constraints:* For large-scale parametric studies requiring thousands of simulations, the computational efficiency of the 6R1C model (12x faster than EnergyPlus) may be prioritized over the marginal accuracy improvements of more complex models.

5.2. Computational efficiency versus accuracy loss

The computational performance analysis of different RC model configurations reveals significant efficiency advantages over traditional simulation methods while maintaining acceptable accuracy levels. The evaluations are conducted on a Surface laptop equipped with an 11th Gen Intel(R) Core (TM) i7-1185G7 processor clocked at 3.00 GHz and complemented by 16 GB RAM. As shown in Table 8, the computational time requirements for RC models are substantially lower than EnergyPlus simulations across both building types. For the detached residential house, the simulation time reduces from 22.3 s with EnergyPlus to a range of 1.5–2.1 s across different RC models. Similarly, for the medium office building, the computation time decreases from 54.7 s to 3.2–5.3 s.

As RC model complexity increases, the 4R1C model remains fastest (1.5s for residential, 3.2s for office). Adding interzonal thermal coupling in the 6R1C model increases computation time by 27–44%, while the

Table 4
Performance of RC models across different climate zones for residential and office buildings.

Climate Zone	Building Type	Model	R^2 of CL	R^2 of HL	NRMSE of CL	NRMSE of HL	NRMSE of daily peak CL	NRMSE of daily peak HL
Shanghai	Residential	4R1C	0.895	0.961	5.92%	5.42%	4.14%	3.21%
		6R1C	0.92	0.969	5.04%	4.66%	3.93%	3.10%
		7R1C	0.967	0.986	3.25%	3.19%	5.00%	3.14%
		7R2C	0.978	0.987	2.69%	2.99%	4.47%	3.50%
	Office	4R1C	0.925	0.869	8.41%	3.75%	13.40%	14.20%
		6R1C	0.955	0.893	4.73%	3.21%	4.42%	3.69%
		7R1C	0.985	0.969	2.84%	1.56%	3.92%	3.36%
		7R2C	0.988	0.983	2.48%	1.17%	3.63%	2.09%
Guangzhou	Residential	4R1C	0.903	0.954	5.83%	5.56%	4.08%	3.34%
		6R1C	0.927	0.961	4.97%	4.83%	3.86%	3.25%
		7R1C	0.972	0.966	3.18%	3.42%	4.95%	3.31%
		7R2C	0.981	0.978	2.62%	3.21%	4.41%	3.62%
	Office	4R1C	0.931	0.852	8.21%	3.82%	12.95%	14.62%
		6R1C	0.962	0.882	4.63%	3.25%	4.36%	3.73%
		7R1C	0.987	0.951	2.75%	1.67%	3.87%	3.45%
		7R2C	0.989	0.964	2.42%	1.39%	3.61%	2.24%
Shenyang	Residential	4R1C	0.881	0.945	6.01%	5.29%	4.21%	3.15%
		6R1C	0.916	0.963	5.12%	4.53%	3.98%	3.04%
		7R1C	0.964	0.987	3.31%	3.07%	5.06%	3.08%
		7R2C	0.975	0.988	2.76%	2.88%	4.53%	3.42%
	Office	4R1C	0.907	0.883	8.78%	3.84%	14.51%	13.93%
		6R1C	0.943	0.904	4.95%	3.19%	4.62%	3.57%
		7R1C	0.978	0.974	3.02%	1.45%	4.12%	3.21%
		7R2C	0.981	0.983	2.57%	1.06%	3.75%	1.94%

Table 5
Forwardly modeled parameters and the calibration of hard-to-determine parameters of the 7R2C model.

Model parameter	Value	Unit
Net building floor area	7781.4	m2
Number of occupants	650	/
Indoor temperature setpoint (summer)	25	°C
Indoor temperature setpoint (winter)	19	°C
External wall U-value	1.9	W/m^2K
Window U-value	4.3	W/m^2K
Window solar heat gain coefficient	0.56	/
Roof U-value	1.4	W/m^2K
Below-grade U-value	2.85	W/m^2K
Lighting intensity	13.02	W/m^2
Equipment intensity	63.51	W/m^2
Calibrated internal wall U-value	2.54	W/m^2
Calibrated Cm	164427	J/m^2K
Calibrated Cs	38792	J/m^2K

transition from 7R1C to 7R2C introduces minimal overhead, indicating that the second capacitance node has limited computational impact. For preliminary building design and optimization processes, the computational advantages are particularly important. Architects and engineers need to quickly assess many design alternatives in early design stages or for existing building retrofit [67]. The parametric study using EnergyPlus that runs from hours to days to understand the effects of building orientation, window to wall ratios, and envelope materials can now be computed in minutes to hours using the developed RC modeling engine, leading to more detailed design space exploration. The implications can be substantial for optimization related research. For example, to perform genetic algorithm optimization with 100 generations, 50 individuals per generation would require 5000 simulation runs. With the 7R2C model, the computation time of a medium office building will reduce from 76 h to 7 h when compared to EnergyPlus, which helps to make complex optimization studies feasible in affordable project timeframes.

RC configurations also have promising scalability prospects with a near-linear increase in computation time from residential to office building simulations. This scaling behavior indicates that the computational cost of RCBldEng is still viable as building complexity grows. These results suggest that the 4R1C or 6R1C models can provide substantial computational savings without compromising reasonable accuracy where speed of analysis is critical, such as at early design stages or large-scale optimization studies. Yet, the modest extra computational cost of using the 7R2C model is justified for applications where accuracy is important due to its superior performance in predicting peak loads and in dealing with complex thermal interactions.

5.3. *Real-world application and practical implications*

The application of the RC model to the education building case study reveals the practical implementation challenges and capabilities. Comparisons with metered data show good agreement for monthly energy consumption predictions, especially in the spring semester when building operation patterns are more regular. The simulation results provide important considerations for practical applications, with

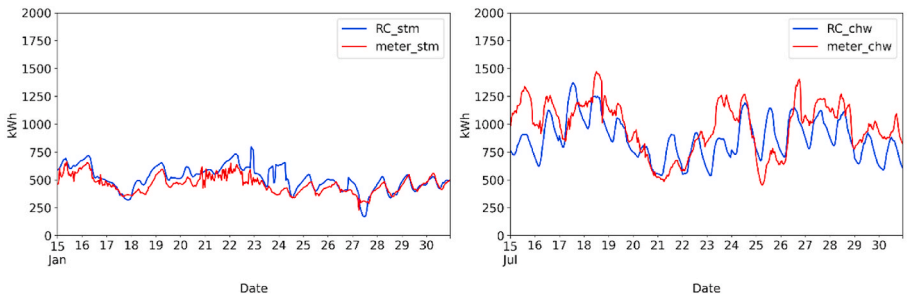


Fig. 16. Hourly energy use predictions versus the metered hourly energy use.

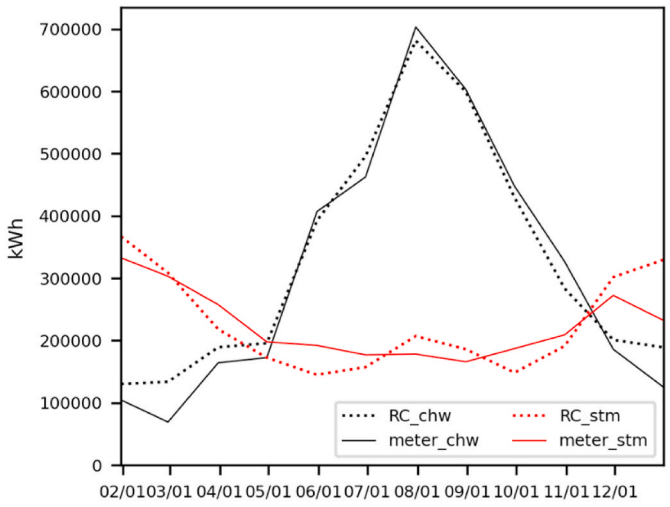


Fig. 17. Monthly energy use predictions versus the metered monthly energy use.

Table 6
Quantitative performance metrics of the 7R2C model for the education building case study.

Time	Granularity	Energy Source	RMSE (kWh)	NRMSE (%)	MBE (%)
Cooling Season (Jul)	Hourly	Chilled Water	134.77	21.46	−11.38
Heating Season (Jan)	Hourly	Steam	83.22	18.04	11.37
Annual	Monthly	Chilled Water	34534.59	10.99	3.97
Annual	Monthly	Steam	40215.15	17.85	1.01

Note: RMSE = Root Mean Square Error; NRMSE = Normalized Root Mean Square Error; MBE = Mean Bias Error.

Table 7
RC model selection decision matrix.

Building Characteristics	Simulation Purpose	Recommended RC Model	Computational Efficiency
Simple residential (1–3 zones)	Early design exploration	4R1C	15x faster than EnergyPlus
Simple residential (1–3 zones)	Load analysis	6R1C	12x faster than EnergyPlus
Simple residential (1–3 zones)	Peak load analysis	7R1C	11x faster than EnergyPlus
Complex commercial (5+ zones)	Early design exploration	6R1C	12x faster than EnergyPlus
Complex commercial (5+ zones)	Load analysis	7R1C	11x faster than EnergyPlus
Complex commercial (5+ zones)	Peak load analysis	7R2C	10x faster than EnergyPlus
High thermal mass buildings	Retrofit analysis	7R2C	10x faster than EnergyPlus
Low thermal mass buildings	Retrofit analysis	7R1C	11x faster than EnergyPlus
Any building during transition seasons	Cooling load prediction	7R2C	10x faster than EnergyPlus

prediction accuracy showing distinct seasonal behavior and larger discrepancies during summer months due to irregular building usage patterns. This observation aligns with Yan et al.'s findings regarding operational uncertainties in building energy modeling [68].

The calibration process indicates that accurate parameter identification is of critical importance and shows that calibrated values for

Table 8
Computational performance of the proposed RC models.

Model	Detached house	Medium office
EnergyPlus	22.3s	54.7s
4R1C	1.5s	3.2s
6R1C	1.9s	4.6s
7R1C	2.0s	5.1s
7R2C	2.1s	5.3s

internal wall U-value and thermal capacitance can be obtained through the differential evolution algorithm under physical constraints. In practice, implementation is highly influenced by data availability and quality. Basic building geometry and envelope properties are generally available, but the detailed operational data that is required for model calibration can sometimes be missing. Hourly metered data was available to support the education building case study in this research, yet such extensive information is not necessarily ubiquitous, jeopardizing the potential in achieving accuracy and reliability of calibration and prediction. Moreover, due to the structured text-based input system of the RCBldEng model, it is practical in terms of model modification and scenario testing. However, simplifications of complex phenomena such as thermal bridges and dynamic occupant behavior may involve systematic uncertainties. For the case study, the current implementation used a simplified approach to represent district heating and cooling, which will be sufficient for the education building case, however, bias might be introduced in modeling more complex HVAC configurations. Beyond energy prediction, the practical implications extend to potential building operation optimization. Implementation of the model in real time control applications can be feasible with careful consideration of prediction uncertainty. Rapidly evaluating different operational scenarios could help make more informed decisions in building design or management for facilities linked to district energy.

5.4. Limitations and future research

In this study, the proposed RC based simulation engine has shown promising prospects in terms of the prototype and real building applications. However, a number of limitations and future development opportunities should be discussed. Although this implementation is useful for many applications, there are still constraints to the current implementation, which could be addressed by further research in the future. Future validation work should incorporate standardized test cases from ASHRAE Standard 140, which provides comprehensive building envelope and system test cases developed by SERI/NREL. These cases would offer additional benchmarks against which to evaluate the performance of different RC model configurations across controlled scenarios.

Regarding model generalization capability, this study included a comprehensive cross-climate validation across three distinct ASHRAE climate zones (Shanghai 3A, Guangzhou 1A, Shenyang 7) for both residential and commercial buildings, totaling 24 building-climate combinations with R^2 values of 0.88–0.99. The validation demonstrates that model accuracy improvements from increased complexity remain consistent across all climates, with the 7R2C model showing particular advantages in severe cold climates for heating loads and the 7R1C model achieving excellent accuracy in cooling-dominated climates. However, the validation focused on buildings with typical envelope constructions meeting ASHRAE 90.1–2013 standards and standard HVAC operation modes; performance for buildings with latest envelope techniques (i.e. highly glazed façades, phase change materials, dynamic insulation) or advanced HVAC strategies (i.e. demand-controlled ventilation, radiant systems, predictive control) requires further investigation.

As for calibration data requirements, when metered energy consumption is unavailable for existing buildings, alternative approaches include short-term measurement campaigns (1–2 weeks), utility bill-based monthly calibration, or benchmarking against similar building

archetypes with conservative assumptions validated through sensitivity analysis. In addition, the current simplified HVAC representation may not adequately capture complex control strategies, variable refrigerant flow systems, or thermal storage, limiting accuracy for buildings with advanced HVAC configurations. The lack of explicit psychrometric modeling—including moisture balance equations and humidity dynamics—restricts applicability to buildings with strict humidity control requirements (laboratories, museums) and may underestimate latent loads in humid climates, though sensible loads remain accurately predicted. These limitations, along with simplified occupancy modeling and the assumption of well-mixed zone air, define the model's appropriate scope as a tool for monthly energy trends, early-stage design comparison, and retrofit measure evaluation rather than detailed operational control or specialized environmental conditioning applications. Additionally, while the current validation focuses on whole-building aggregated loads, future work should include zone-level validation to diagnose potential systematic biases across different orientations (south/north), locations (core/perimeter), and floor levels (upper/lower), which would provide deeper insights into the model's spatial accuracy and help identify specific scenarios where model refinements may be needed.

At this point, urban scale application both offers challenges and presents opportunities for future development and extension of this study [55]. This suggests the potential in scaling to district or urban scale simulations, but additional complexities like building shadowing effects, urban heat island impacts, and district energy system interactions should be addressed [69,70]. Further exploration of the potential for real time applications is warranted in the context of developing robust control algorithms that can leverage the model's predictions while dealing with operational uncertainties.

6. Conclusions

In this research, a novel hybrid modeling approach that integrates the strengths of both forward and inverse modeling methods through the development of the RCBldEng simulation engine is introduced. The approach bridges the theoretical and practice by combining theoretical building physics principles with real-world energy use data calibration. Four RC model configurations (4R1C, 6R1C, 7R1C, and 7R2C) are proposed and developed to cater to varying levels of building complexity, with each model configured to provide an advantage in specific applications.

A dual-capacitance model which coupled the thermal behavior of the interzonal zones and modeled the effect of interzonal thermal coupling and heat transfer was developed (7R2C) and finds superior performance in modeling complex building thermal behavior, especially in the case of buildings with high indoor thermal mass. The 4R1C model is computationally efficient but lacks the ability to describe the dynamics of interzonal interactions. The 7R2C model aligns the closest with EnergyPlus results, particularly for buildings with significant thermal mass, while still maintaining substantially faster computation times. For situations where real time simulations are needed or when limited computational resources are available, this computational efficiency becomes vital. Moreover, the application of the model to an educational building case study shows that the model can handle real-world scenarios. The differential evolution algorithm is shown to effectively identify hard-to-determine parameters, but seasonal prediction accuracy variations suggest that consideration of operational uncertainties is essential. The structured text-based modeling approach can be useful for early-stage design analysis and retrofit studies since it allows rapid scenario testing and model modification.

It is shown that the RC-based approach has significant advantages in computational cost for design optimization and parametric studies. The comprehensive exploration of design alternatives in a matter of minutes rather than hours, allows for better informed decision making in building design and operation. For applications where rapid analysis is desired, i.e., early-stage design evaluation or large-scale optimization studies, these benefits are particularly pronounced. This research offers a methodology for choosing and implementing RC models depending on the building characteristics and simulation requirements. The results provide practical guidelines to strike a balance between model sophistication and computational efficiency and suggest opportunities for future development of RC based simulation engines. Future work includes enhanced representation of complex HVAC systems, more user-friendly interfaces, and urban-scale applications. In short, this research shows that RC-based building energy simulation engines can provide viable alternative to traditional dynamic simulation methods, especially for cases where fast analysis or extensive parametric studies are needed, which may facilitate more efficient and sustainable building energy solutions.

Code availability

The RCBldEng simulation engine (v1.2) developed in this research is available as a Windows executable, which can be downloaded on GitHub at https://github.com/andersonspy/RCBIdEng/releases/tag/v1.2.0. The program runs on Windows operating systems and requires minimal installation. The executable and example files can be downloaded in compressed file, and users can follow the instruction in the user manual to perform building energy simulation using the RC modeling approach introduced in this paper.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

I would sincerely thank Professor Bill Braham and Alex Waegel from the Center for Environmental Building + Design at the University of Pennsylvania for their remarkable assistance and support at the early stages of the development of RCBldEng when I was pursuing my Ph.D. a decade ago. I am also particularly indebted to the late Professor Godfried Augenbroe from Georgia Institute of Technology, who offered profound insights on RC building simulation as well as encouragement that inspired this work. My ten-year winding journey in developing the RCBldEng would not have been possible without their mentorship and encouragement.

Appendix A: Infiltration and Natural Ventilation Model

This appendix describes the mathematical formulation of the infiltration and natural ventilation model used in the simulation engine.

A.1 Infiltration Model

The infiltration model is based on the principle of air leakage through building envelopes due to pressure differences caused by temperature differentials (stack effect) and wind.

A.1.1 'Real' Infiltration Method

The 'real' infiltration method combines models for stack effect and wind-driven infiltration.

A.1.1.1 Stack Effect. The stack effect model is based on the buoyancy-driven airflow due to temperature differences between indoor and outdoor air:

$$q_{v,stack} = \max\left(0.0146 \cdot F_r \cdot f_{inf} \cdot (0.7 \cdot h \cdot |T_e - T_{set}|)^{0.667}, 0.001\right)$$

where $q_{v,stack}$ is the volumetric flow rate due to stack effect [m^3/s], F_r is the reference flow rate [m^3/s], f_{inf} is the infiltration fraction [−], h is the zone height [m], T_e is the outdoor temperature [K], and T_{set} is the indoor set temperature [K]. The constant 0.7 accounts for the average height of neutral pressure level in buildings.

A.1.1.2 Wind Effect. The wind effect model is based on the pressure differences created by wind around the building. It also follows the orifice equation:

$$q_{v,wind} = 0.0769 \cdot F_r \cdot f_{inf} \cdot \left(\Delta C_p \cdot v_{site} \cdot w_s^2\right)^{0.667}$$

where $q_{v,wind}$ is the volumetric flow rate due to wind effect [m^3/s], ΔC_p is the pressure coefficient difference [−], v_{site} is a site-specific factor [−], and w_s is the wind speed [m/s]. The pressure coefficient difference accounts for the distribution of wind pressure on the building façade.

A.1.1.3 Combined Stack and Wind Effect. The combined effect is modeled using a quadrature sum approach, which accounts for the fact that stack and wind effects are not simply additive:

$$q_{v,sw} = \max\left(q_{v,stack}, q_{v,wind}\right) + \frac{0.14 \cdot q_{v,stack} \cdot q_{v,wind}}{F_r \cdot f_{inf}}$$

A.1.1.4 Final Infiltration Rate. The final infiltration rate considers any additional pressure differences and ensures non-negative values:

$$q_{v,inf} = \max\left(0, -q_{v,diff}\right) + q_{v,sw}$$

where $q_{v,diff}$ is a term accounting for the pressure-induced ventilation flow rate difference between supply and exhaust air streams.

A.1.2 'Constant' Infiltration Method

The constant infiltration method uses a simplified approach based on a reference flow rate and infiltration fraction:

$$q_{v,inf} = F_r \cdot f_{inf}$$

A.2 Natural Ventilation Model

When ventilation type is not mechanical, the natural ventilation model is activated. The natural ventilation model is based on the principles of airflow through naturally opened windows or vents, when the proper opening pattern is taken for buoyancy-driven or wind-driven airflow.

A.2.1 Ventilation Speed

The ventilation speed model combines wind speed and temperature difference effects:

$$V = 0.01 + 0.001 \cdot w_s^2 + 0.0035 \cdot h \cdot |T_{set} - T_e|$$

A.2.2 Environmental Factors

The model includes factors to account for the impact of wind and temperature on occupant behavior regarding window opening:

Wind factor:

$$Y_{wind} = \min(\max(1 - 0.1 \cdot w_s, 0), 1)$$

Temperature factor:

$$Y_{temp} = \min(\max(T_e / 25.0 + 0.2, 0), 1)$$

Opening factor:

$$R_{opw} = Y_{wind} \cdot Y_{temp}$$

A.2.3 Natural Ventilation Rate

The natural ventilation rate is modeled based on the orifice equation, considering the opening area and ventilation speed:

$$q_{v,airing} = R_{opw} \cdot \frac{3.6 \cdot 500 \cdot A_{ow} \cdot \sqrt{V}}{A_f}$$

where A_{ow} is the opening area [m^2] and A_f is the floor area [m^2].

A.2.4 Natural Ventilation Control

The model includes a simple control strategy based on cooling demand and outdoor temperature:

$$V_{hrs} = \begin{cases} 1 \text{ if } q_C > 0 \text{ and } T_e < T_{set} \text{ and } T_e > 16^\circ C \\ 0 \text{ otherwise} \end{cases}$$

where q_C is the cooling demand and NV_{hrs} is the calculated natural ventilation hours.

A.3 Total Ventilation and Heat Transfer

A.3.1 Total Ventilation Rate

The total ventilation rate combines mechanical supply, infiltration, and natural ventilation:

$$q_{v,tot} = q_{v,supp} + q_{v,inf} + q_{v,NV}$$

A.3.2 Ventilation Heat Transfer Coefficient

The ventilation heat transfer coefficient is based on the heat capacity of air and the total ventilation rate:

$$H_{ve} = \frac{1.2}{3.6} \cdot q_{v,tot}$$

where 1.2 is the volumetric heat capacity of air [$kJ/(m^3 \cdot K)$].

Appendix B. Solar Radiation Calculation Model

This appendix describes the simplified solar radiation calculation model used in the building energy simulation engine. The model calculates various solar parameters and radiation components for different surface orientations.

B.1 Solar Position Calculations

B.1.1 Day of Year

The day of the year (Y_{date}) is calculated based on the month and day:

$$Y_{date} = \begin{cases} Y_{date,i-1} + day_i - day_{i-1} & \text{if } month_i = month_{i-1} \\ Y_{date,i-1} + 1 & \text{otherwise} \end{cases}$$

where, Y_{date}: the cumulative date-related value; $Y_{date,i-1}$: The previous value of Y_{date}, where the subscript $i-1$ denotes the preceding iteration or time step; day_i: The current day of the month; day_{i-1}: The previous day of the month; $month_i$: The current month; $month_{i-1}$: The previous month;

B.1.2 Solar Time

The model calculates several time-related parameters:

Solar time variable (τ):

$$\tau = \frac{2\pi(Y_{date} - 1)}{365}$$

Equation of Time (ET):

$$ET = 2.2918(0.0075 + 0.1868 \cos\tau - 3.2077 \sin\tau - 1.4615 \cos 2\tau - 4.089 \sin 2\tau)$$

Apparent Solar Time (AST):

$$AST = hour + \frac{ET}{60} + \frac{longitude - LSM}{15}$$

where LSM is the Local Standard Time Meridian.

B.1.3 Solar Angles

Solar declination (δ):

$$\delta = 23.45 \sin\left(\frac{Y_{date} + 284}{365} \cdot 2\pi\right)$$

Hour angle (θ_h):

$$\theta_h = 15(AST - 12)$$

Solar altitude (β):

$$\sin\beta = \cos\phi\cos\delta\cos\theta_h + \sin\phi\sin\delta$$

where φ is the latitude.

Solar azimuth (ψ):

$$\sin\psi = \frac{\sin\theta_h\cos\delta}{\cos\beta}$$

$$\cos\psi = \frac{\cos\theta_h\cos\delta\sin\phi - \sin\delta\cos\phi}{\cos\beta}$$

B.2 Surface Solar Radiation Calculations

For each surface orientation, the model calculates.

B.2.1 Surface Solar Azimuth (γ):

$$\gamma = \left|\psi - \psi_{surface}\right|$$

B.2.2 Angle of Incidence (θ):

$$\cos\theta = \cos\beta\cos\gamma\sin\sigma + \sin\beta\cos\sigma$$

where σ is the surface tilt angle.

B.2.3 Direct Beam Component ($E_{t,b}$):

$$E_{t,b} = \begin{cases} 0 & \text{if } 90^\circ < \gamma < 270^\circ \text{ or } \cos\theta < 0 \\ E_b\cos\theta & \text{otherwise} \end{cases}$$

where $E_{t,b}$ is the direct normal irradiance.

B.2.4 Diffuse Component ($E_{t,d}$):

$$Y = \max\left(0.45, 0.55 + 0.437\cos\theta + 0.313\cos^2\theta\right)$$

$$E_{t,d} = \begin{cases} E_d(Y\sin\sigma + \cos\sigma) & \text{if } \sigma \le 90^\circ \\ E_d Y\sin\sigma & \text{if } \sigma > 90^\circ \end{cases}$$

where E_d is the diffuse horizontal irradiance.

B.2.5 Ground-Reflected Component ($E_{t,r}$):

$$E_{t,r} = (E_b\sin\beta + E_d)r_g\frac{1-\cos\sigma}{2}$$

where r_g is the ground reflectivity.

B.2.6 Global Solar Radiation (GSR):

$$GSR = E_{t,b} + E_{t,d} + E_{t,r}$$

B.3 Solar Irradiation on Tilted Surfaces

For various tilt angles, the model calculates solar irradiation:

$$E_{sol} = \frac{E_{gh}}{\sin\beta}\left(\cos\beta\sin\sigma\cos\left(\psi - \psi_{surface}\right) + \sin\beta\cos\sigma\right)$$

where E_{gh} is the global horizontal irradiance.

Appendix C. Solar Reduction Factors

The model calculates solar reduction factors for various shading devices.

B.3.1 Overhang:

$$SRF_{overhang} = \frac{\max\left(0, 1 - \frac{0.5 \tan i}{\tan(90^\circ - \max(\beta, 0^\circ))}\right) E_{t,b} + \left(1 - \frac{i}{90^\circ}\right) E_{t,d} + E_{t,r}}{GSR}$$

B.3.2 Fin:

$$SRF_{fin} = \frac{\max\left(0, 1 - \frac{0.5 \tan i}{\tan\left(90^\circ - \left|\psi - \psi_{surface}\right|\right)}\right) E_{t,b} + E_{t,d} + E_{t,r}}{GSR}$$

B.3.3 Horizon:

$$SRF_{horizon} = \begin{cases} 1 - \dfrac{E_{t,b}}{GSR} & \text{if } \max(\beta, 0^\circ) < i \\ 1 & \text{otherwise} \end{cases}$$

where i is the shading device angle.

Appendix D. Building Energy Use Calculation Method

This part describes the simplified HVAC energy use calculation model implemented in the building energy simulation engine. The model calculates energy consumption for heating, cooling, ventilation, and domestic hot water (DHW) for each thermal zone.

D.1 Supply Air Flow Rates

The model calculates supply air flow rates for heating and cooling:

$$V_{h,supply} = \frac{q_H}{3600\rho\left(T_{supply,H} - T_{air,hc}\right)}$$

$$V_{c,supply} = \frac{q_C}{3600\rho\left(T_{air,hc} - \left(T_{supply,C} + \Delta T_{reheat}\right)\right)}$$

where $V_{h,supply}$ and $V_{c,supply}$ are the heating and cooling supply air flow rates [$m^3/h/m^2$], q_H and q_C are the heating and cooling demands [W/m^2], ρ is the air density [kg/m^3], $T_{supply,H}$ and $T_{supply,C}$ are the supply air temperatures for heating and cooling [°C], $T_{air,hc}$ is the zone air temperature [°C], and ΔT_{reheat} is the reheat temperature difference [°C].

D.2 Fan Energy Consumption

The fan volume and energy consumption are calculated as:

$$V_{fan} = \max\left(V_{h,supply}, V_{c,supply}\right) + V_{exh}$$

$$E_{fan} = \max\left(V_{fan}\right) P_{fan} f_{control} f_{BAC,e}$$

where V_{exh} is the exhaust air flow rate, P_{fan} is the fan power [$W/(m^3/s)$], $f_{control}$ is the fan flow control factor, and $f_{BAC,e}$ is the building automation and control (BAC) factor for energy.

D.3 Pump Energy Consumption

The model calculates water flow rates and pump volumes for heating and cooling:

$$V_{w,h} = \frac{q_{H,HVAC}}{\rho_{\left(w\ c_{\left(p,w\Delta T_{w,H}\right)}\right)}}$$

$$V_{w,c} = \frac{q_{C,HVAC}}{\rho_w c_{p,w} \Delta T_{w,C}}$$

where $V_{w,h}$ and $V_{w,c}$ are the heating and cooling water flow rates [m^3/h], ρ_w is the water density, $c_{p,w}$ is the specific heat capacity of water, and $\Delta T_{w,H}$ and $\Delta T_{w,C}$ are the water temperature differences for heating and cooling. The pump volumes are then calculated based on the pump control strategy.

D.4 HVAC System Efficiency

The model calculates heating and cooling system efficiencies using performance curves:

$$eta_{heat} = \mathrm{PLV}_h\left(q_{H,HVAC} \,/\, q_{H,max}\right)\cdot\eta_{heat,nom}$$

$$COP_{cool} = \mathrm{PLV}_c\left(q_{C,HVAC} \,/\, q_{C,max}\right)\cdot COP_{cool,nom}$$

where PLV_h and PLV_c are the part load value curves for heating and cooling, and $\eta_{heat,nom}$ and $COP_{cool,nom}$ are the nominal efficiencies. $q_{H,HVAC}$ and $q_{C,HVAC}$ are the hourly heating and cooling demand, and $q_{C,max}$ and $q_{H,max}$ are the maximum heating and cooling output of the HVAC system.

D.5 Distribution Losses and Final Energy Consumption

The model accounts for distribution losses:

$$f_{dem,heat} = \max\left(0.1, \frac{\sum q_{H,HVAC}}{\sum q_{H,HVAC} + \sum q_{C,HVAC}}\right)$$

$$f_{dem,cool} = \max\left(1 - f_{dem,heat}, 0.1\right)$$

$$\eta_{dist,heat} = \frac{1}{1 + a_{heat} + f_{waste} \Big/ f_{dem,heat}}$$

$$\eta_{dist,cool} = \frac{1}{1 + a_{cool} + f_{waste} \Big/ f_{dem,cool}}$$

where:

$f_{dem,cool}$ represents the fraction of cooling demand; $f_{dem,heat}$ is the fraction of heating demand; $\eta_{dist,heat}$ and $\eta_{dist,cool}$ are the efficiency of heating and cooling distribution; a_{heat} and a_{cool} are coefficients related to heating and cooling losses; f_{waste} represents the fraction of waste energy.

Final energy consumption for heating and cooling:

$$E_{heat} = \frac{\left(q_{H,HVAC} + q_{H,loss}\right) f_{BAC,hc}}{\eta_{heat}}$$

$$E_{cool} = \frac{\left(q_{C,HVAC} + q_{C,loss}\right) f_{BAC,hc}}{COP_{cool}}$$

where $q_{H,loss}$ and $q_{C,loss}$ are the distribution losses, and $f_{BAC,hc}$ is the BAC factor for heating and cooling.

D.6 Domestic Hot Water (DHW) Demand

The DHW demand is calculated based on the occupancy profile:

$$Q_{DHW} = 12\cdot DHW\cdot 4.18\cdot 1000\cdot 45\cdot A_f$$

$$q_{DHW} = Q_{DHW}\cdot \frac{f_{occ}}{\sum f_{occ}}$$

where DHW is the daily hot water consumption [m^3/m^2/day], A_f is the floor area [m^2], and f_{occ} is the occupancy fraction.

Appendix E. Onsite Renewable Energy Production Calculation Model

E.1 Photovoltaic (PV) System Energy Generation

The model calculates the solar irradiation incident on the PV panels and the resulting energy generation.

E.1.1 Solar Irradiation on PV Panels

The solar irradiation on PV panels $E_{sol,PV}$ is determined based on the PV panel area, angle, and orientation:

$$E_{sol,PV} = \begin{cases} E_{gh}/1000 \; \textit{if PV angle} = 0^\circ \\ E_{sol,\theta,\psi}/1000 \; \textit{if PV angle} > 0^\circ \\ 0 \; \textit{if PV area} = 0 \end{cases}$$

where.

E_{gh} is the global horizontal irradiation [W/m^2]
$E_{sol,\theta,\psi}$ is the solar irradiation on a tilted surface with angle θ and orientation ψ [W/m^2]

E.1.2 PV Energy Generation

The energy generated by the PV system is calculated as:

$$E_{gen,PV} = E_{sol,PV} \cdot P_{pk} \cdot f_{perf} \cdot 3600000$$

where.

$E_{gen,PV}$ is the energy generated by the PV system [J]
$E_{sol,PV}$ is the solar irradiation on PV panels [kWh/m^2]
P_{pk} is the peak power of the PV system [kW]
f_{perf} is the performance factor of the PV system [−]
3600000 is the conversion factor from kWh to J

E.2 Solar Water Heating (SWH) System

The model calculates the solar irradiation incident on the SWH collectors.

E.2.1 Solar Irradiation on SWH Collectors

The solar irradiation on SWH collectors ($E_{sol,SWH}$) is determined based on the collector area, angle, and orientation:

$$E_{sol,SWH} = \begin{cases} E_{gh}/1000 \text{ if SWH angle} = 0^\circ \\ E_{sol,\theta,\psi}/1000 \text{ if SWH angle} > 0^\circ \\ 0 \text{ if SWH area} = 0 \end{cases}$$

E.2.2 SWH Energy Gain

The energy gain of the SWH system is calculated as:

$$E_{gain,SWH} = E_{sol,SWH} \cdot 3600000$$

where.

$E_{gain,SWH}$ is the energy gain of the SWH system [J]
$E_{sol,SWH}$ is the solar irradiation on SWH collectors [kWh/m^2]

E.3 Model Implementation Notes

If the PV or SWH area is zero or if there are invalid inputs, the respective solar irradiation is set to zero. The model uses pre-calculated solar irradiation data for tilted surfaces ($E_{sol,\theta,\psi}$) stored in the solar calculation dictionary which is calculated based on Appendix C.

The performance factor (f_{perf}) for the PV system accounts for various system losses and inefficiencies. The model assumes that the SWH energy gain is directly proportional to the incident solar irradiation, without considering system efficiency or heat losses. For a more accurate estimation, additional factors such as collector efficiency, heat exchanger effectiveness, and storage losses should be incorporated.

E.2 Wind Turbine Energy Generation

If wind turbines are present, the energy generation is calculated as:

$$E_{gen,wind} = 0.5 \rho_{air} N A_{swept} \eta_{turbine} v_{wind}^3 \cdot 3600$$

where ρ_{air} is the air density, N is the number of turbines, A_{swept} is the swept area, $\eta_{turbine}$ is the turbine efficiency, and v_{wind} is the wind speed.

Appendix F. Forward Model Parameter Calculations

F.1 Thermal Resistances Calculation

F.1.1 Envelope Thermal Resistances

The thermal resistances of building envelope components are calculated from their U-values and areas:

$$R_{ex} = \frac{1}{\sum_i (U_i \cdot A_i)_{opaque}}$$

$$R_{win} = \frac{1}{\sum_i (U_i \cdot A_i)_{window}}$$

where U_i is the heat transfer coefficient (W/m^{2}K) and A_i is the corresponding surface area (m^2).

F.1.2 Ventilation and Infiltration Resistance

Ventilation resistance is calculated based on the air exchange rate and air properties:

$$R_v = \frac{1}{1.2 \cdot ACH \cdot V_{zone}/3.6}$$

where ACH is the air changes per hour (h^{-1}), V_{zone} is the zone volume (m^3), and 1.2 represents the volumetric heat capacity of air (kJ/m^{3}K).

F.1.3 Internal Resistances

Internal resistances between air and surfaces, and between surfaces and thermal mass:

$$R_{ia} = \frac{1}{h_{in} \cdot A_t}$$

$$R_{im} = \frac{1}{h_{im} \cdot A_m}$$

where h_{in} is the interior convective heat transfer coefficient, h_{im} represents the conduction coefficient between surfaces and mass, A_t is the total surface area, and A_m is the effective mass surface area in m^2.

F.1.4 Interzonal Thermal Resistances

For thermal coupling between zones:

$$R_{iw,i} = \frac{1}{U_{iw} \cdot A_{iw,i}}$$

$$R_{if,i} = \frac{1}{U_{if} \cdot A_{if,i}}$$

where U_{iw} and U_{if} are the heat transfer coefficients of internal walls and floors, and $A_{iw,i}$ and $A_{if,i}$ are the corresponding areas between the current zone and adjacent zone i.

F.2 Thermal Capacitances Estimation

The thermal capacitances are calculated as:

$$C = \sum_i \rho_i \cdot c_{p,i} \cdot d_i \cdot A_i$$

where ρ_i is the density (kg/m^3), $c_{p,i}$ is the specific heat capacity (J/kgK), d_i is the effective thickness (m), and A_i is the area (m^2) of the respective building element.

F.3 Heat Gains Calculation

F.3.1 Internal Heat Gains

Internal heat gains from occupants, equipment, and lighting are calculated as:

$$Q_{int} = q_{occ} \cdot f_{occ} + q_{app} \cdot f_{app} + q_{li} \cdot f_{li}$$

where q_{occ}, q_{app}, and q_{li} are the heat gain densities (W/m^2) from occupants, appliances, and lighting, and f_{occ}, f_{app}, and f_{li} are the corresponding time-dependent fractions.

F.3.2 Solar Heat Gains

Solar heat gains through windows and opaque surfaces:

$$Q_{sol} = \frac{\sum_{i,j}\left(SRF_{i,j}\cdot GSR_j\cdot A_{i,j}\cdot(1-FF_i)\cdot SHGC_i\right)}{A_f} + \frac{\sum_{i,j}\left(GSR_j\cdot A_{i,j}\cdot U_i\cdot\alpha_i\cdot r_{se}\right)}{A_f}$$

where $SRF_{i,j}$ is the solar reduction factor (accounting for shading devices), GSR_j is the global solar radiation on surface j, $A_{i,j}$ is the area, FF_i is the frame factor, $SHGC_i$ is the solar heat gain coefficient, U_i is the U-value, α_i is the absorption coefficient, r_{se} is the external surface resistance, and A_f is the floor area.

Data availability

Data will be made available on request.

References

[1] GhaffarianHoseini A, Dahlan ND, Berardi U, GhaffarianHoseini A, Makaremi N, GhaffarianHoseini M. Sustainable energy performances of green buildings: a review of current theories, implementations and challenges. Renew Sustain Energy Rev 2013;25:1–17.

[2] Shen P, Li Y, Gao X, Chen S, Cui X, Zhang Y, et al. Climate adaptability of building passive strategies to changing future urban climate: a review. Nexus 2025;2(2):1–13.

[3] Ascione F, Bianco N, De Stasio C, Mauro GM, Vanoli GP. Simulation-based model predictive control by the multi-objective optimization of building energy performance and thermal comfort. Energy Build 2016;111:131–44.

[4] USDOE. The encyclopedic reference to EnergyPlus input and output. 8.0. Champaign, IL: Univ. of Ill: EnergyPlus Development Team; 2014.

[5] Picco M, Lollini R, Marengo M. Towards energy performance evaluation in early stage building design: a simplification methodology for commercial building models. Energy Build 2014;76:497–505.

[6] Shen P. Building retrofit optimization considering future climate and decision-making under various mindsets. J Build Eng 2024;96:110422.

[7] Afram A, Janabi-Sharifi F. Theory and applications of HVAC control systems – a review of model predictive control (MPC). Build Environ 2014;72:343–55.

[8] Torcellini P, Pless S, Deru M, Crawley D. Zero energy buildings: a critical look at the definition. In: 2006 ACEEE summer study on energy efficiency in buildings. Golden, CO (United States): Pacific Grove, California: National Renewable Energy Lab.(NREL); 2006.

[9] Li Y, O'Neill Z. An innovative fault impact analysis framework for enhancing building operations. Energy Build 2019;199:311–31.

[10] Crawley DB, Lawrie LK, Winkelmann FC, Buhl WF, Huang YJ, Pedersen CO, et al. EnergyPlus: creating a new-generation building energy simulation program. Energy Build 2001;33(4):319–31.

[11] Klein SA. TRNSYS-A transient system simulation program. University of Wisconsin-Madison, Engineering Experiment Station Report; 1988. p. 38. 12.

[12] Yan D, Xia J, Tang W, Song F, Zhang X, Jiang Y. DeST—An integrated building simulation toolkit Part I: fundamentals. Conference DeST—An integrated building simulation toolkit Part I: fundamentals, vol. vol. 1. Springer, p. 95-110.

[13] Yan D, Zhou X, An J, Kang X, Bu F, Chen Y, et al. DeST 3.0: a new-generation building performance simulation platform. Build Simulat 2022;15(11):1849–68.

[14] Hirsch JJ. eQUEST: the quick energy simulation tool. https://www.doe2.com/equest/2006.

[15] Harish VSKV, Kumar A. A review on modeling and simulation of building energy systems. Renew Sustain Energy Rev 2016;56:1272–92.

[16] Asadi S, Amiri SS, Mottahedi M. On the development of multi-linear regression analysis to assess energy consumption in the early stages of building design. Energy Build 2014;85:246–55.

[17] Sreshthaputra A, Haberl J, Claridge D. Development of a toolkit for calculating linear, change–point linear and multiple–linear inverse building energy analysis models. ASHRAE Research Project 1050-RP, Detailed Test Results; 2001.

[18] Płoszaj-Mazurek M, Ryńska E, Grochulska-Salak M. Methods to optimize carbon footprint of buildings in regenerative architectural design with the use of machine learning, convolutional neural network, and parametric design. Energies 2020;13(20):5289.

[19] Wang H, Chen Y, Kang J, Ding Z, Zhu H. An XGBoost-Based predictive control strategy for HVAC systems in providing day-ahead demand response. Build Environ 2023;238:110350.

[20] Li Q, Meng Q, Cai J, Yoshino H, Mochida A. Applying support vector machine to predict hourly cooling load in the building. Appl Energy 2009;86(10):2249–56.

[21] Deng M, Fu B, Menassa CC, Kamat VR. Learning-based personal models for joint optimization of thermal comfort and energy consumption in flexible workplaces. Energy Build 2023:113438.

[22] Olu-Ajayi R, Alaka H, Sulaimon I, Sunmola F, Ajayi S. Building energy consumption prediction for residential buildings using deep learning and other machine learning techniques. J Build Eng 2022;45:103406.

[23] Wang Z, Hong T, Piette MA. Building thermal load prediction through shallow machine learning and deep learning. Appl Energy 2020;263:114683.

[24] Hosamo HH, Nielsen HK, Kraniotis D, Svennevig PR, Svidt K. Digital twin framework for automated fault source detection and prediction for comfort performance evaluation of existing non-residential Norwegian buildings. Energy Build 2023;281:112732.

[25] Li Y, O'Neill Z, Zhang L, Chen J, Im P, DeGraw J. Grey-box modeling and application for building energy simulations - a critical review. Renew Sustain Energy Rev 2021;146:111174.

[26] Deb C, Schlueter A. Review of data-driven energy modelling techniques for building retrofit. Renew Sustain Energy Rev 2021;144:110990.

[27] John Kelly K, Haberl JS, Claridge DE. Inverse modeling toolkit: numerical algorithms. ASHRAE Transactions 2003;109:425–34.

[28] Braun JE, Chaturvedi N. An inverse gray-box model for transient building load prediction. HVAC R Res 2002;8(1):73–99.

[29] Okuyama H, Onishi Y. System parameter identification theory and uncertainty analysis methods for multi-zone building heat transfer and infiltration. Build Environ 2012;54:39–52.

[30] Massa Gray F, Schmidt M. A hybrid approach to thermal building modelling using a combination of Gaussian processes and grey-box models. Energy Build 2018;165:56–63.

[31] Dinh vB, Delinchant B, Wurtz F. The importance of derivatives for simultaneous optimization of sizing and operation strategies: application to buildings and HVAC systems. Conference the importance of derivatives for simultaneous optimization of sizing and operation strategies: application to buildings and HVAC systems, Newcastle, United Kingdom. https://hal.science/hal-01369149/document https://hal.science/hal-01369149/file/hal-01369149%2C%20version%201%20BSO%202016.pdf. .

[32] Zhu Q, Xu X, Wang J, Xiao F. Development of dynamic simplified thermal models of active pipe-embedded building envelopes using genetic algorithm. Int J Therm Sci 2014;76:258–72.

[33] Shen P, Braham W, Yi Y. Development of a lightweight building simulation tool using simplified zone thermal coupling for fast parametric study. Appl Energy 2018;223:188–214.

[34] Tian Z, Wang J, Lan B. In: Administration CNIP, editor. A building dynamic room temperature prediction method based on equivalent RC model; 2019. China.

[35] Shen P, Braham W, Yi Y, Eaton E. Rapid multi-objective optimization with multi-year future weather condition and decision-making support for building retrofit. Energy 2019;172:892–912.

[36] Shen P, Wang Z, Ji Y. Exploring potential for residential energy saving in New York using developed lightweight prototypical building models based on survey data in the past decades. Sustain Cities Soc 2021;66:102659.

[37] Shen P, Yang B. Projecting Texas energy use for residential sector under future climate and urbanization scenarios: a bottom-up method based on twenty-year regional energy use data. Energy 2020;193:116694.

[38] Li A, Sun Y, Xu X. Development of a simplified resistance and capacitance (RC)-network model for pipe-embedded concrete radiant floors. Energy Build 2017;150:353–75.

[39] Hazyuk I, Ghiaus C, Penhouet D. Optimal temperature control of intermittently heated buildings using Model Predictive Control: part I – building modeling. Build Environ 2012;51:379–87.

[40] Massa Gray F, Schmidt M. Thermal building modelling using Gaussian processes. Energy Build 2016;119:119–28.

[41] Chong A, Menberg K. Guidelines for the Bayesian calibration of building energy models. Energy Build 2018;174:527–47.

[42] Hwang J, Lim H, Lim J. Reducing uncertainty of building shape information in urban building energy modeling using Bayesian calibration. Sustain Cities Soc 2024;116:105895.

[43] Pandiyan SV, Gros S, Rajasekharan J. Physics informed neural network based multi-zone electric water heater modeling for demand response. Appl Energy 2025;380:125037.

[44] Manfren M, Nastasi B, Groppi D, Astiaso Garcia D. Open data and energy analytics - an analysis of essential information for energy system planning, design and operation. Energy 2020;213:118803.

[45] Bueno B, Norford LK, Pigeon G, Britter R. A resistance-capacitance network model for the analysis of the interactions between the energy performance of buildings and the urban climate. Build Environ 2012;54:116–25.

[46] Vivian J, Zarrella A, Emmi G, Carli Md. Analysis of simplified lumped-capacitance models to simulate the thermal behaviour of buildings. Conference Analysis of simplified lumped-capacitance models to simulate the thermal behaviour of buildings..

[47] Paschkis V, Baker H. A method for determining unsteady-state heat transfer by means of an electrical analogy. Trans Am Soc Mech Eng 1942;64(2):105–10.

[48] Robertson AF, Gross D. An electrical-analog method for transient heat-flow analysis. J Res Natl Bur Stand 1958;61(2):105–15.

[49] Giretti A, Vaccarini M, Casals M, Macarulla M, Fuertes A, Jones RV. Reduced-order modeling for energy performance contracting. Energy Build 2018;167:216–30.

[50] De Coninck R, Magnusson F, Åkesson J, Helsen L. Toolbox for development and validation of grey-box building models for forecasting and control. J Build Perform Simul 2016;9(3):288–303.

[51] Remmen P, Lauster M, Mans M, Fuchs M, Osterhage T, Müller D. TEASER: an open tool for urban energy modelling of building stocks. J Build Perform Simul 2018;11(1):84–98.
[52] Jayathissa P, Luzzatto M, Schmidli J, Hofer J, Nagy Z, Schlueter A. Optimising building net energy demand with dynamic BIPV shading. Appl Energy 2017;202:726–35.
[53] Li DHW, Yang L, Lam JC. Impact of climate change on energy use in the built environment in different climate zones – a review. Energy 2012;42(1):103–12.
[54] Bacher P, Madsen H. Identifying suitable models for the heat dynamics of buildings. Energy Build 2011;43(7):1511–22.
[55] Yang J, Wang H, Cheng L, Gao Z, Xu F. A review of resistance–capacitance thermal network model in urban building energy simulations. Energy Build 2024;323:114765.
[56] Westermann P, Evins R. Surrogate modelling for sustainable building design – a review. Energy Build 2019;198:170–86.
[57] Hong T, Jiang Y. A new multizone model for the simulation of building thermal performance. Build Environ 1997;32(2):123–8.
[58] Kramer R, van Schijndel J, Schellen H. Inverse modeling of simplified hygrothermal building models to predict and characterize indoor climates. Build Environ 2013;68:87–99.
[59] Shin M, Haberl JS. Thermal zoning for building HVAC design and energy simulation: a literature review. Energy Build 2019;203:109429.
[60] Foucquier A, Robert S, Suard F, Stéphan L, Jay A. State of the art in building modelling and energy performances prediction: a review. Renew Sustain Energy Rev 2013;23:272–88.
[61] Feng C, Zhang C, Lu J, Zhao Y. Hybrid data-driven and physics-based fast building cooling demand modeling method for large-scale building demand response control. J Build Eng 2025;100:111808.
[62] DOE. The EnergyPlus engineering reference, vol. 8. EnergyPlus Development Team; 2015. 0 ed, http://www.energyplus.gov.
[63] Storn R, Price K. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. J Global Optim 1997;11(4):341–59.
[64] Shen P, Wang M, Liu J, Ji Y. Hourly air temperature projection in future urban area by coupling climate change and urban heat island effect. Energy Build 2023;279:112676.
[65] Shen P, Wang H. Archetype building energy modeling approaches and applications: a review. Renew Sustain Energy Rev 2024;199:114478.
[66] Halverson MA, Rosenberg MI, Hart PR, Richman EE, Athalye RA, Winiarski DW. ANSI/ASHRAE/IES standard 90.1-2013 determination of energy savings: qualitative analysis. Richland, WA (United States): Pacific Northwest National Lab. (PNNL); 2014.
[67] Li S, Wang M, Shen P, Cui X, Bu L, Wei R, et al. Energy saving and thermal comfort performance of passive retrofitting measures for traditional rammed Earth house in Lingnan, China. Buildings 2022;12(10):1716.
[68] Yan D, O'Brien W, Hong T, Feng X, Burak Gunay H, Tahmasebi F, et al. Occupant behavior modeling for building performance simulation: current state and future challenges. Energy Build 2015;107:264–78.
[69] Shen P, Ji Y, Li Y, Wang M, Cui X, Tong H. Combined impact of climate change and heat island on building energy use in three megacities in China. Energy Build 2025:115386.
[70] Shen P, Ji Y, Zhong M. Performance of district energy system under changing climate: a case study of Shenzhen. Appl Energy 2025;379:124986.