paranoá ORIGINAL ARTICLE
TECHNOLOGY, ENVIRONMENT AND SUSTAINABILITY

2026
v.19, e60256

Simplification of a high-fidelity model for generating thermoenergetic metamodels in a university educational building

Simplificação de modelo de alta fidelidade para geração de metamodelos termoenergéticos em uma edificação educacional universitária

Simplificación de un modelo de alta fidelidad para la generación de metamodelos termoenergéticos en un edificio educativo universitario

Kahena Marx Silva
Federal University of Viçosa; Center for Exact Sciences; Department of Architecture and Urbanism.
Viçosa (MG), Brazil.

Rafael de Paula Garcia*
Federal University of Viçosa; Center for Exact Sciences; Department of Architecture and Urbanism.
Viçosa (MG), Brazil.
rafael.pgarcia@ufv.br

Ligiana Pricila G. Fonseca
Federal University of Viçosa; Center for Exact Sciences; Graduate Program in Architecture and Urbanism.
Viçosa (MG), Brazil.

Joyce Correna Carlo
Federal University of Viçosa; Center for Exact Sciences; Department of Architecture and Urbanism.
Viçosa (MG), Brazil.

* Corresponding author.

CRediT

Author contribution: Analysis; Methodology; Writing – review and editing: SILVA, K.M.; GARCIA, R.P.; FONSECA, L.P.G.; CARLO, J.C.; Conception: GARCIA, R.P.; CARLO, J.C.; Data curation; Data collection; Software; Validation; Visualization; Writing - original draft: SILVA, K.M.; Supervision: GARCIA, R.P.; Software: FONSECA, L.P.G.

Conflicts of interest: The authors certify that there is no conflict of interest.

Funding: National Council for Scientific and Technological Development (CNPq) - Project Process 406426/2022-8; Ibero-American Program of Science and Technology for Development – CYTED through the TRAPECIO Network; Capes Brazil (funding code 001) through the Graduate Program in Architecture and Urbanism PPG.au/UFV and international mobility; CNPq under the Institutional Scientific Initiation Program PIBIC/CNPq 2023-2024 and 2024-2025.

Ethics approval: The authors certify that there was no need for approval by an Ethics Committee.

Use of AI: DeepL translating tool was used in the process of translating the accepted work in Portuguese (original) to English. The Editor-in-Chief and the authors reviewed and edited the content as needed and assume full responsibility for the content of the published article.

Editors: Daniel Sant'Ana (Editor-in-Chief); Leandro S. Cruz (Associate Editor); Pedro G. Cardoso (Editorial Assistant).

Submitted on: Nov. 5, 2025 | Accepted on: Dec. 3, 2025.

Abstract

With climate change and the increased demand for cooling, thermoenergetic computational modeling has become essential for studying and adapting buildings. However, constructing models that accurately represent reality, especially for complex buildings, is time-consuming and requires significant computational processing power. This study investigates the use of metamodels to reduce building simulation time without compromising the accuracy of the results. Using the Lecture Hall II (PVB) of the Federal University of Viçosa as a case study, three metamodels were developed, via EnergyPlus, based on simplifications of the original calibrated model. Validation was performed based on ASHRAE Guideline 14, considering the Normalized Mean Bias Error (NMBE) and Coefficient of Variation of the Root Mean Square Error (CV(RMSE)) indices, and complementary statistics. Among the three metamodels compared, the simplest one obtained the best performance and shortest simulation time. The results indicate that reducing the number of thermal zones allows for a decrease in computational time, while keeping errors within acceptable limits.

Keywords: Metamodels; Thermoenergetic simulation; Thermal zones; Calibration.

Resumo

Com as mudanças climáticas e o aumento da demanda por refrigeração, a modelagem computacional termoenergética torna-se essencial para estudar e adaptar edificações. No entanto, a construção de modelos que representem fielmente a realidade, especialmente para edificações complexas, tem alto custo de tempo e processamento computacional. Este estudo investiga o uso de metamodelos para reduzir o tempo de simulação de edificações sem comprometer a precisão dos resultados. Utilizando o Pavilhão de Aulas II (PVB) da Universidade Federal de Viçosa como estudo de caso, três metamodelos foram desenvolvidos, via EnergyPlus, a partir de simplificações do modelo original calibrado. A validação foi realizada com base na ASHRAE Guideline 14, considerando os índices Normalized Mean Bias Error (NMBE) e Coefficient of Variation of the Root Mean Square Error (CV(RMSE)), e estatísticas complementares. Dentre os três metamodelos comparados, o mais simples obteve o melhor desempenho e menor tempo de simulação. Os resultados indicam que a redução do número de zonas térmicas permite diminuir o tempo computacional, mantendo os erros dentro dos limites aceitáveis.

Palavras-chave: Metamodelos; Simulação termoenergética; Zonas térmicas; Calibração.

Resumen

Con el cambio climático y el aumento de la demanda de refrigeración, la modelización termoenergética computacional se vuelve esencial para el estudio y la adaptación de los edificios. Sin embargo, la construcción de modelos que representen fielmente la realidad, especialmente en edificaciones complejas, implica un alto costo en tiempo y procesamiento computacional. Este estudio investiga el uso de metamodelos para reducir el tiempo de simulación de edificaciones sin comprometer la precisión de los resultados. Utilizando el Pabellón de Aulas II (PVB) de la Universidad Federal de Viçosa como estudio de caso, se desarrollaron tres metamodelos en EnergyPlus, a partir de simplificaciones del modelo original calibrado. La validación se realizó conforme a la Guía 14 de ASHRAE, considerando los índices Normalized Mean Bias Error (NMBE) y Coefficient of Variation of the Root Mean Square Error (CV(RMSE)), y estadísticas complementarias. Entre los tres metamodelos comparados, el más simple obtuvo el mejor desempeño y el menor tiempo de simulación. Los resultados indican que la reducción del número de zonas térmicas permite disminuir significativamente el tiempo computacional, manteniendo los errores dentro de los límites aceptables.

Palabras-clave: Metamodelos; Simulación termoenergética; Zonas térmicas; Calibración.

1 Introduction

Climate change has resulted in an increase in the average temperature of the Earth's surface, as well as in the intensification and greater frequency of extreme events, such as heat waves, torrential rains, and severe droughts (IPCC, 2022). Given this scenario and the consequent increase in demand for cooling systems (WMO, 2022), one of the most advanced methods for investigating strategies for adapting buildings to climate change, estimating their thermal performance, is computer simulation of scenarios (Olinger et al., 2023). However, although modeling and computer simulation of buildings offer numerous benefits, their high cost, both in terms of time and processing, is still an obstacle to their full application in architectural production (Cui et al., 2016; Touloupaki and Theodosiou, 2017; Olinger et al., 2019; Veiga et al., 2021). This obstacle is exacerbated when simulation is applied to large-scale complex buildings such as universities, and when the objective of the study is to investigate a large number of design alternatives. Despite this complexity, thermoenergetic modeling has proven to be a fundamental investment for studying and improving the quality of educational environments in several countries, including Brazil, as it enables the simulation of retrofit alternatives or the creation of new buildings with proven energy efficiency (Maciel et al., 2021; Lopes et al., 2023).

The simplification of modeling and computational simulation processes is an indispensable strategy for expanding the feasibility of energy studies in Brazil. In order to promote such simplification, the adoption of metamodels, or surrogate models, has emerged as a viable alternative to overcome the limitations of the high time spent on complex computational modeling, as well as its simulation and post-processing (Olinger et al., 2019; Veiga et al., 2021). These simplified models, also called surrogate models, are approximate representations of complex models, but with a less demanding computational structure. They are capable of accurately preserving and reproducing, but at a lower computational cost, the phenomena that complex models simulate (Hensen and Lamberts, 2011; Westermann and Evins, 2019).

Techniques such as sensitivity analysis have proven to be fundamental in identifying and prioritizing variables of greater relevance, helping to simplify parameters and enabling the construction of these representative models (Yang et al., 2016). Research such as that by Coakley, Raftery, and Keane (2014) highlights the potential of metamodels in contexts of high computational complexity, pointing out that their application can facilitate analysis, reduce operating costs, and make the use of advanced simulations more democratic.

In this scenario, the development of energy metamodels for buildings is not only a technical solution but also a strategic tool for integrating energy efficiency and sustainability into the architectural design process. Its results can have a significant impact on the quality of educational environments and on adapting to climate change in a more dynamic way.

This study seeks to develop and validate a thermoenergetic metamodel that represents the thermal performance of a university educational building. The building selected as a case study was the Classroom Pavilion II (popularly known as PVB), located at the Federal University of Viçosa, Viçosa – MG, Brazil. As it is a public building of high relevance in the academic context of the city and due to its high occupancy density, typical of teaching environments, it has been the subject of research in academic works. As an example, we sought to investigate the effect of natural ventilation and different occupancy densities on the risk of infection by airborne diseases (Souza, 2024). Still in the context of researching strategies for climate adaptation, the development of a metamodel can enable faster computer simulations and, thus, allow for a greater number of evaluations of solutions within a reduced time interval.

This work therefore seeks to answer the following questions: "Can a metamodel accurately represent the thermal performance of an existing and previously validated computational model?"; "What simplifications can be applied to the model without compromising its fidelity and the quality of the prediction of the physical phenomena under evaluation?" The answers are directed at the aforementioned building in order to make its modeling and the computational simulations that support decision-making more feasible. Consequently, facilitating performance simulation contributes to the promotion of more resilient and sustainable buildings, a need that has been exacerbated in recent years due to climate change.

2 A brief discussion of metamodels and simplifications

The application of computational simulation and optimization processes for building performance analysis is not new. Bouchlaghem and Letherman (1990) developed a computational model of numerical optimization applied to the thermal design of buildings, proposing a system capable of adjusting architectural parameters to improve the passive thermal performance of buildings. In more recent reviews, such as those by Westermann and Evins (2019) and Worthmann, Cichocka, and Waibel (2022), it is noted that there is extensive academic literature on the incorporation of computational simulations, optimization processes, and metamodeling as design tools, mainly to solve energy issues in buildings.

There are two main approaches to building computational models of buildings: physics-driven (or white-box) models and statistical (data-driven or black-box) models (Foucquier et al., 2013). White-box models, for example, are models based entirely on thermodynamic physical principles, offering high accuracy and interpretability, but with a high modeling and calibration effort (Arendt et al., 2018). Black-box models, on the other hand, are empirical and rely on statistical methods, including machine learning, to predict patterns from input and output data. They are numerically more complex (Sun; Burton; Huang, 2020), which makes the process faster since they do not explicitly consider thermodynamic mechanisms themselves (Arendt et al., 2018). A middle ground between these two are grey-box models, which combine simplified physical approaches and statistical calibration to balance accuracy and computational cost (Foucquier et al., 2013; Arendt et al., 2018; Yu et al., 2019).

Other methods present in the literature are bottom-up and top-down models. Models built from a bottom-up perspective start from a cumulative logic, where there is a detailed characterization of individual components that, when aggregated, can estimate global results, such as energy consumption, temperature, and others (Kavgic et al., 2010). Xia, Wu, and Zou (2025) developed a bottom-up model to assess the energy consumption of urban residential buildings in China, using prototype units simulated in EnergyPlus associated with machine learning techniques. In this case, instead of analyzing the energy consumption of the entire city based on already aggregated data (top-down), the authors simulated the energy behavior of typical residential units, called prototypes, and then progressively scaled up these results to represent the set of urban buildings in Guangzhou.

In top-down models, the logic is reversed: it is one of decomposition. It starts from a less robust global system and performs a decomposition based on aggregated statistical data. Typically, these models work for macroeconomic indicators, fuel prices, urban energy demand estimates, etc. (Wong et al., 2021; Ali et al., 2021).

In addition to the construction methods presented, the design of energy metamodels for simulation is possible with the aid of different programs, analysis methodologies, degrees of complexity, and numerical approaches (algorithms), depending on the context and its objective (Ostergard, 2017). These approaches can employ techniques such as multiple linear regression models, polynomial regression, Gaussian process regression, and machine learning, as systematized by Westermann and Evins (2019) and Sánchez-Zabala and Gómez-Acebo (2024).

Olinger et al. (2019) presented a meta-model based on artificial neural networks (ANN) for naturally ventilated office buildings. The objective was to estimate the Exceedance Hour Fraction (EHF), an indicator of thermal comfort, from 12 input parameters representing construction and operational characteristics. The metamodel achieved an average absolute error of only 0.04, demonstrating high accuracy and feasibility for use in early design stages. The research consolidated the application of machine learning as a tool to reduce the complexity of thermo-energy simulations and provided rapid support to architects and engineers. Later, the same authors, Olinger et al. (2023), developed a metamodel to estimate the thermal performance of naturally ventilated offices, based on simulations performed in EnergyPlus and a database of real buildings in São Paulo. The model predicted the fraction of hours of thermal discomfort (EHF) according to the ASHRAE 55 adaptive method, considering two approaches, single-zone and multi-zone. Sensitivity analyses showed that factors such as window opening, wall transmittance, and facade exposure strongly influence performance. The study demonstrated that the metamodel can help designers predict thermal comfort in the early stages of architectural design.

Veiga et al. (2021), in turn, proposed a metamodel using machine learning to predict the percentage of hours occupied within the operative temperature range (PHFT) in naturally ventilated apartments, according to Brazilian standard NBR 15575. The model was developed from parametric simulations in EnergyPlus, with geometric simplifications and parameters based on Brazilian housing typologies. The result was a robust model, with low computational cost and applicable to different climates in Brazil, with the potential for direct integration into the thermal performance assessments of the standard.

Observing more manual metamodeling, via trial and error, Gil (2017) and Silva and Ghisi (2014) applied the reduction in the number of thermal zones using the EnergyPlus program based on what Coakley, Raftery, and Keane (2014) define as the development of evidence-based models. Thermal zones are modeling units within EnergyPlus defined by thermal criteria and calculated by heat balance equations. They are used to represent the average thermal behavior of a space, considering that the air and surfaces within that zone exchange heat uniformly, since they are treated as a single control volume (U.S. Department of Energy, 2025). These zones do not necessarily correspond to a physical environment delimited in the modeling, but rather a region of similar thermal behavior.

Gil (2017) evaluated the manual simplification of thermal zones in EnergyPlus for single-story, detached social housing, simulated in the program as multizone (MuZ) and single-zone (MoZ) models. To validate the simplification, the author used indicators based on indoor air temperature, operative temperature, and discomfort degree-hours (°Ch), comparing the outputs of the simplified models with those of the more complex model. The analysis showed that the annual average of the differences between the models was low enough to justify the use of a simplified model in the early stages of design. Silva and Ghisi (2014) applied a similar methodology to evaluate 15 simplified geometric variations of a multifamily educational building with complex geometry, with geometric simplification being the central focus of the study. The authors modeled a real educational building at the Federal University of Santa Catarina (UFSC), creating 15 model variations in EnergyPlus, divided according to three levels of simplification. The simplified models were created manually and validated by comparing annual energy consumption and simulation time. The acceptance criteria considered the reduction in simulation time and a maximum difference of 10% in energy consumption compared to the base model. Both Silva and Ghisi (2014) and Gil (2017) concluded that simplification in building modeling for energy simulations, even with a manual approach, reduces computational and time costs without compromising the accuracy of the results, being able to replace complete calibrated models with satisfactory performance.

These approaches aimed at saving time and optimization have been widely explored in studies and professional practices that use a variety of simulation and optimization tools aimed at simplifying and integrating computational models. Among the most widely used energy simulation programs are EnergyPlus, TRNSYS, DOE-2, ESP-r, and IDA ICE, the former being the most used in the literature, accounting for about 40% of the studies reviewed by Shi et al. (2016). Its wide acceptance is due to its accuracy, reliability, and integration with external platforms. At the same time, tools such as GenOpt, MATLAB, modeFRONTIER, and ModelCenter have been applied to connect simulation programs to optimization algorithms, favoring the automated search for energy-efficient design solutions (Shi et al., 2016). In the contexts of parametric modeling and metamodeling, graphical interfaces such as Rhino/Grasshopper, associated with extensions such as Ladybug Tools and Dragonfly, have expanded the interaction between architectural design and energy simulation, allowing for more agile and interactive representations. This diversity of programs and methodologies represents a set of complementary possibilities that have been applied to facilitate and accelerate the process of energy simulation in face of growing demand for more efficient and sustainable buildings. The pursuit of LEED and BREEAM certifications in the United Kingdom and the Green Building Label in China are evidence of this (Shi et al., 2016).

3 Materials and methods

The Classroom Pavilion II (PVB) is a naturally ventilated metal-framed building (Figure 1) designed to meet the educational needs of various courses at the Federal University of Viçosa (UFV). With a built area of 6,705 m^2 distributed over three floors, the building houses 30 classrooms, as well as complementary spaces such as monitoring rooms, two auditoriums, a computer lab, a cafeteria, technical support rooms, and restrooms on all floors. The spatial layout of the PVB follows a "square eight" format (Figure 1), structured around two open internal courtyards. These courtyards play an essential role in the ventilation and thermal comfort of the environment, since the building does not have an active air conditioning system, basing its thermal performance predominantly on passive natural ventilation strategies. To enhance this ventilation, each of the 30 classrooms has thermal chimneys, maxim-air windows, doors connected to external balconies, and fans.

Figure 1: Photo PVB, 2023. Viçosa, MG.

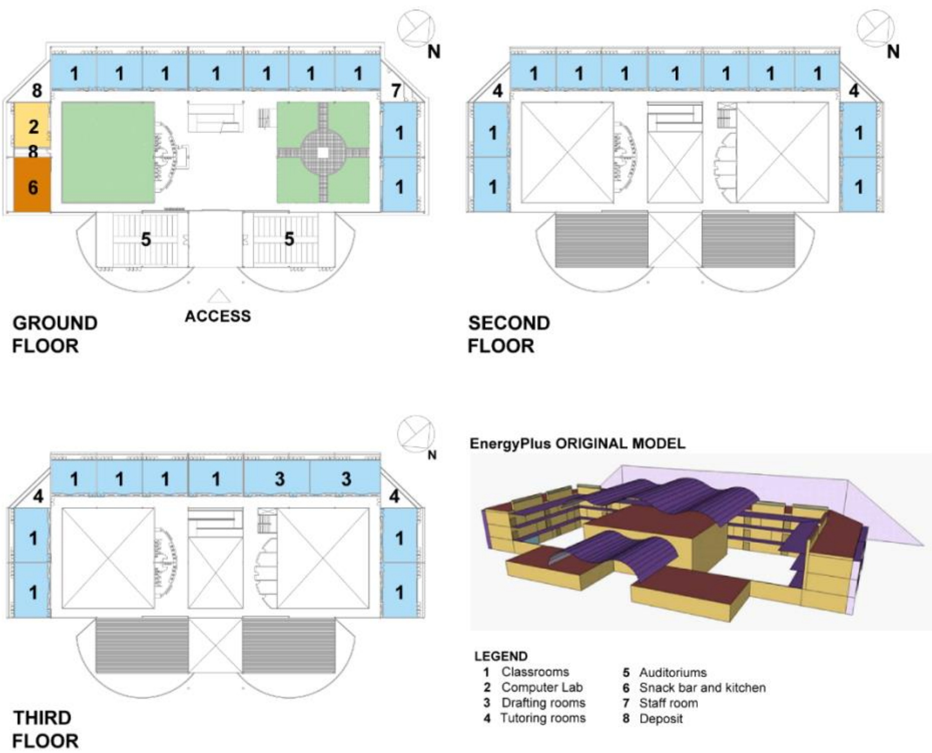
Source: left, the authors; right, Google Earth.

The building was designed so that the facades with classrooms face northwest, southeast, and southwest, resulting in different variations in sun exposure throughout the day (Figure 1).

3.1 Original Computational Model

In order to understand the thermo-energy performance of PVB, Souza (2024) developed a detailed computational model of the building (Figure 2) using *EnergyPlus* version 22.1. This is an open-source program distributed by *the U.S. Department of Energy* (DOE) and developed to perform energy performance simulations in buildings and their systems. The program is capable of simulating and computationally calculating the behavior of thermal loads and building systems such as heating, cooling, ventilation, indoor and outdoor lighting, shading, air quality, energy generation, environmental emissions, and facade performance, among others. It is widely used in academic research and architectural and engineering projects for its ability to reproduce energy behaviors in buildings (ASHRAE *Standard* 140, 2023).

Figure 2: PVB, floor plans, and computational model developed by Souza (2024).



The model has 55 thermal zones that thermally describe classrooms, thermal chimneys, and adjacent spaces, ensuring an accurate representation of their thermal compartmentalization. Thermal chimneys are elements that play a fundamental role in thermo-energy behavior, both in the actual building and in the computational model. Each classroom has an upper opening connected to a thermal chimney, which is responsible for promoting cross ventilation between the windows and the upper air outlet of the chimney to which it is connected (Figure 3). They vary in height according to each floor, which directly influences the air renewal rates of each classroom (Souza, 2024).

Figure 3: Modeling and location of thermal chimneys in the calibrated model.

The spaces belonging to the central atrium of the building were modeled as a single thermal zone. This spatial simplification was carried out because of the simulation objective for which it was designed: the evaluation of classroom performance. Each classroom was modeled with five openings: an access door, a row of windows facing the external balcony, two doors to the same balcony, and an entrance to the thermal chimney inside the room. The openings facing the balconies were simplified only in terms of number, while maintaining the equivalence of area in relation to the actual building. Thus, instead of the 12 maxim-air windows positioned at the top of the wall facing the balcony, the simulation considered a single opening equivalent in area and located in the same position as in the actual building. The external balconies, the circulation corridors that provide access to the classrooms, and the roof over the corridors and balconies were modeled as shading elements. Natural ventilation was simulated by the *AirflowNetwork* module of *EnergyPlus*, which allows a dynamic analysis of internal and external air flows based on the interaction between the openings present in the model and the climatic conditions of Viçosa, MG. Table 1 shows some parameters defined in the original model.

The complete model was calibrated manually according to the methodology described by Souza (2024), comparing the internal temperatures of classrooms, measured in the field in 2022 and 2023, with the temperatures obtained by simulation. The measurements were taken with *Onset* HOBO U-12 data loggers, configured to record every 5 minutes, positioned in the center of the rooms, in the internal corridors, and on the balconies, protected by ventilated enclosures to avoid radiative interference. The data was collected over four days in winter (August 2022) and ten days in summer (February and March 2023), in empty rooms with no lighting or equipment turned on, ensuring thermal conditions equivalent to those simulated. The windows remained open between 8 a.m. and 4 p.m., and the surface temperatures of walls, floors, and ceilings were periodically checked with a FLIR TG-165X thermal imaging camera, ensuring the consistency of the model's construction parameters.

Table 1: Parameters of the original model.

EnergyPlus version	22.1.0
Orientation	39.5
Latitude	-20.753889
Longitude	-42.881944
Elevation	648m
Terrain	Suburb
Solar distribution model	Complete Interior and Exterior
Time step	6
Simulation Period	January 1 to December 31
Soil Temperature Calculation Model	Xing
Ventilation Module	Airflow Network
Ventilation Operation	- 50% of the windows are open 24 hours a day during the school terms and closed during the holidays; - Doors are closed - whether balcony doors or entrance doors along the internal circulation corridor; - Thermal chimneys available 100% of the time.
Gap coefficient	Maxim-air windows = 0.5
Location	Viçosa - MG
Occupancy	Not modeled
Lighting	Not modeled
Equipment	Not modeled
Climate Archive	Vicosa_1985-2014_TMY_ISO15927-4
Requested outputs	Site Outdoor Air Drybulb Temperature [°C] (Hourly) Zone Mean Air Temperature [°C] (Hourly)

The simulations took into account the climatic characteristics of the city of Viçosa, in Minas Gerais. According to Lucarelli, Oliveira, and Carlo (2022), the municipality has a Cwa climate according to the Köppen classification, characterized by hot and humid summers and mild to dry winters, a condition typical of the Zona da Mata region of Minas Gerais. The average annual temperatures are close to 21°C, with highs reaching 30°C and lows around 13 and 14°C in winter. The average annual precipitation is around 1300 mm, concentrated between November and March, and the average relative humidity is close to 78%. For the simulations, the Viçosa_1985-2014_TMY_ISO15927-4 climate file was used, which represents a typical meteorological year (TMY) in Viçosa, allowing the thermal performance to be evaluated under average conditions representative of the local climate.

The indicators for validation and acceptance of the computational model were the *Normalized Mean* Bias *Error* (NMBE) and the Co*efficient of Vari*ation *of the Root Mean Square Error* (CV(RMSE)), from ASHRAE *Guideline* 14 (ASHRAE, 2018). The limits for acceptance of the calibrated model were set at a maximum of ±10% for NMBE and 30% for CV(RMSE), as defined in the standard for hourly data.

The simulation results were anchored in two *EnergyPlus outputs*:

1. *Outdoor Air Drybulb Temperature [°C] (Hourly)* - outdoor drybulb air temperature, which provides the surrounding climate conditions;
2. *Zone Mean Air Temperature [°C] (Hourly)* - average indoor air temperature in each thermal zone, adopted as a variable for analysis and calibration.

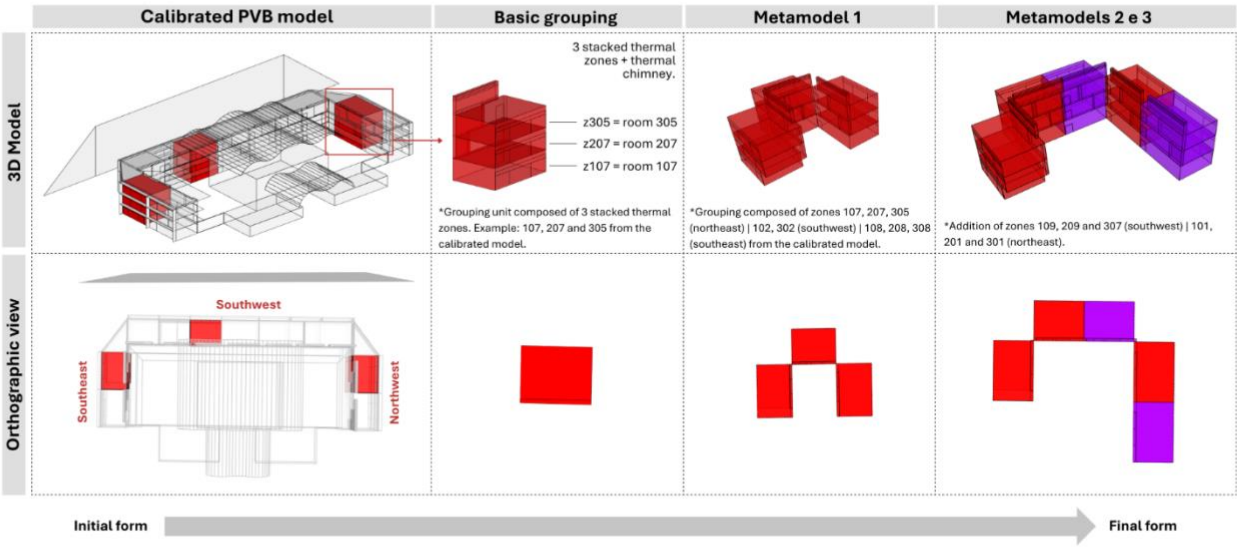
These two parameters were chosen because they represent, respectively, the thermal behavior of the external environment and the internal response of the building to external thermal variations.

3.2 Metamodels

In this study, three metamodels were developed with different levels of geometry simplification and number of zones, named Metamodel 1, Metamodel 2, and Metamodel 3 (Figure 4). All were simulated directly in *EnergyPlus* v.22.1.0, using the calibrated model and its settings as a modeling reference to ensure comparability. In all cases, the central zone existing in the calibrated version was eliminated, since the objective of the metamodel is to represent the thermal behavior of classrooms, as in the calibrated model. In addition, as all metamodels are geometrically smaller than the calibrated model, there is no space to insert volumes in the central area of the metamodels, so it was decided to eliminate it. External shading and thermal chimneys were maintained, as these are elements that could significantly impact the thermal behavior of the metamodels. In addition, the dimensions of the southwest facade were considerably reduced.

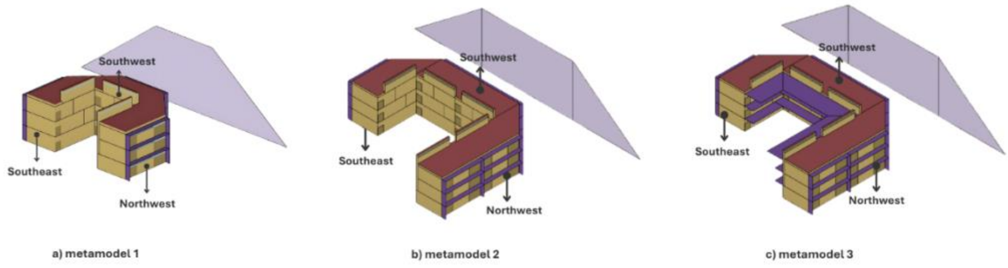
The simulations were performed on a computer with a 10th generation *Intel Core* i5 processor, 8 GB of RAM, and a 512 GB SSD storage unit. It is important to note that the simulation time values presented in the results are directly related to the processing capacity of the machine used, and may vary significantly on equipment with higher or lower configurations.

Figure 4: Design process by integrating thermal zones into metamodels.



Metamodel 1 was structured with three thermal zones stacked on each arm of the building (southeast, southwest, and northwest), each zone representing a room on each floor of the model. The thermal chimneys corresponding to each region were maintained, and the internal shading provided by the corridors leading to the rooms was eliminated in relation to the original model. Only external shading was modeled. The model had 13 zones. Based on the results obtained in metamodel 1, decisions were made to create metamodel 2, with the aim of improving the initial performance obtained. In this version, a separate thermal zone was added for each floor of the northwest and southwest orientations, as shown in Figure 5. In this version, the number of thermal zones increased to 21. Finally, metamodel 3, developed from the results of metamodel 2 also to improve its performance, incorporated external shading elements to mitigate direct solar gains. Figure 5 visually complements the configuration information for each metamodel.

Figure 5: Metamodels evaluated in comparison with the calibrated PVB model.



3.3 Metamodel evaluation indicators

The validation of the metamodels developed in this study was performed in comparison with the calibrated model using two approaches. The first was based on the same indicator s recommended by ASHRAE *Guideline* 14 (ASHRAE, 2018), previously used by O'donovan, O'Sullivan, and Murphy (2019) and Sakiyama (2021), and also employed by Souza (2024) in the calibration process of the original computational model. The second follows a statistical method of polynomial curve fitting, used to assess the statistical and visual similarity between the thermal patterns obtained.

Although ASHRAE indicators are traditionally applied to the calibration of energy models based on consumption comparisons, Souza (2024) demonstrated their applicability also in temperature-based models. Following this same logic, the present study validated the metamodels by comparing the temperatures simulated by the calibrated PVB model with those generated by the metamodels, verifying that the deviations between the two remained within the limits established by ASHRAE, 10% for NBME and 30% for CV(RMSE).

It should be noted that this is, therefore, a comparison between simulated results: the reference model, calibrated by Souza (2024), was previously validated with actual measured data, while the metamodels were evaluated based on their ability to reproduce the thermal behavior obtained from this calibrated model.

Figure 6: Distribution of thermal zones aggregated by solar orientation and floor in the PVB.

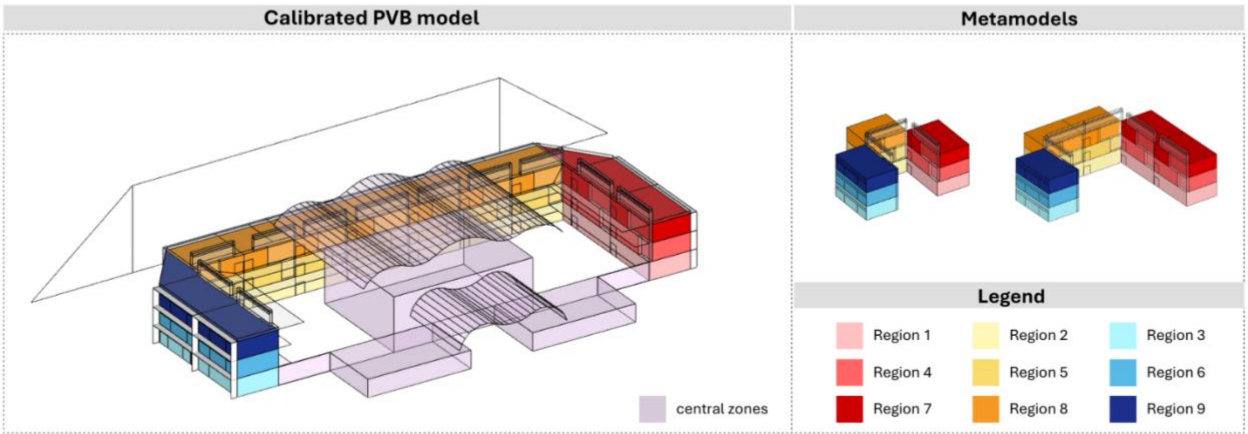


Table 2: List of analysis regions in the models.

Pavement	**Solar orientation**		
	Northwest	**Southwest**	**Southeast**
Ground floor	Region 1	Region 2	Region 3
First Floor	Region 4	Region 5	Region 6
Second Floor	Region 7	Region 8	Region 9

Since the number of thermal zones representing classrooms differs between the original calibrated model and the metamodel, it was necessary to aggregate the data between the model zones. To select the zones to be aggregated, the spatial distribution of the building was taken into account, dividing it into the three solar orientations of the classroom facades (northwest, southeast, and southwest) and its three floors (ground floor, first floor, and second floor) (Figure 6). In total, nine distinct regions were defined according to their conditions of contact or exposure to the outside environment (Table 2).

The comparison between the models was performed using the average hourly temperatures between corresponding regions. Thus, the average hourly temperatures of all rooms in that region are calculated through a data grouping process, regardless of the number of rooms (thermal zones) present in each region analyzed by the calibrated model and metamodels. This makes it possible to pair corresponding regions and assess whether the simplifications applied in each metamodel are consistent with the actual behavior of the building.

The results were also segmented into two seasonal periods: summer (September 24 to April 3) and winter (April 4 to September 23). This division considers the extreme conditions of the climate archive used for the calibration of the metamodels. Calibrating the models by the extremes allows us to cover the entire spectrum of building performance in response to different external environmental variations and reduce uncertainties.

3.3.1 ASHRAE *Guideline* 14 indicators: CV(RMSE) and NMBE

ASHRAE *Guideline* 14 (2018) defines criteria for calibrating energy models modeled in building simulation programs—such as *EnergyPlus*—establishing an NMBE less than or equal to the 10% module for hourly data, and a CV(RMSE) less than or equal to 30% also for hourly data. The indicators are calculated using equations 1 and 2.

$$\boldsymbol{NMBE} = \frac{\mathbf{1}}{\overline{\boldsymbol{m}}} \cdot \frac{\sum_{i=1}^{n}(\boldsymbol{mi} - \boldsymbol{Si})}{\boldsymbol{n} - \boldsymbol{p}} \cdot \mathbf{100}(\%) \tag{1}$$

$$\boldsymbol{CV}(\boldsymbol{RMSE}) = \frac{\mathbf{1}}{\overline{\boldsymbol{m}}} \cdot \sqrt{\frac{\sum_{i=1}^{n}(\boldsymbol{mi} - \boldsymbol{Si})^2}{\boldsymbol{n} - \boldsymbol{p}}} \cdot \mathbf{100}(\%) \tag{2}$$

Where:
$NMBE$ = *Normalized mean bias error*
$CV(RMSE)$ = Coefficient of Variation *of the Root Mean Square Error*;
$\overline{m}$ = Average of measured values
mi = Measured temperature value
Si = Value simulated by the model
n = Number of samples
p = Number of adjustable model parameters

Unlike CV(RMSE), which only assumes positive values, NMBE can assume both positive and negative values. This is because it indicates the direction of the average error between simulated data and reference data. In this study, the metamodel data were treated as simulated values to be calibrated, while the calibrated PVB model data acted as reference data. Positive values indicate that the model tends to underestimate the actual data (i.e., the simulated values are smaller than the measured values), while negative values indicate overestimation (the simulated values are larger than the measured values) (ASHRAE, 2018). Thus, it is possible not only to quantify the average error between the data sets, but also to identify trends in the model's behavior, which is relevant for directing adjustments to the input parameters based on the direction of the deviation observed by the indicators.

3.3.2 Statistical Similarity by Polynomial Curves

In addition to traditional metrics, a statistical analysis based on polynomial curve fitting was performed to compare the thermal patterns of the metamodels in relation to the calibrated model. This method allows verifying the point accuracy of the simulated temperatures and the overall behavior of the thermal curves over time. The methodology adopted consists of the following steps:

- Fitting of 4th degree polynomial curves (Equation 3) to represent the variation in temperature over time in the metamodels and in the validated complex model;
- Calculation of the mean square error (MSE) between the adjusted values and the simulated values for each model.

The general equation used to adjust the curves is expressed by Equation 3.

$$\boldsymbol{T(x) = a_0 + a_1x^1 + a_2x^2 + a_3x^3 + a_4x^4} \tag{3}$$

Where:
$\boldsymbol{T(x)}$ = Temperature as a function of time x
$a_0, a_1..., a_4$ are the coefficients of the polynomial regression, determined by least squares to minimize the adjustment error.

To fit the curve to the data, the *polyfit* function from the *numpy.polynomial.polynomial* sublibrary of the *Python* library was used. This function performs a polynomial fit using the least squares method, estimating the polynomial coefficients that minimize the sum of the squares of the differences between the observed values and those predicted by the model. The choice of a fourth-degree polynomial was made based on preliminary tests, observing its ability to capture complex variations in data behavior without resorting to overly sophisticated models.

3.3.3 Refinement via *timesteps*

At the end of the process, after the results were obtained and validated, we sought to refine them by changing the *timestep* value in the simulations and verifying its impact.

One point of attention is the initial choice of the value 6 for the simulation *timestep* for both the calibrated model and the metamodels. The *timestep*, or *Number of Timesteps per Hour*, defines how many times per hour the program updates the energy balance calculations in the one-hour time interval (if the simulation is being performed hourly). Its value can vary from 1 to 60, and *EnergyPlus* itself recommends that the chosen value be at least 6 in order to avoid inaccuracies in the results. This means that the program performs six energy balance calculations to obtain each simulated hourly value. Thus, the larger the chosen *timestep*, the greater the number of times the program will perform calculations to generate a result. The result, in turn, will be detailed, but this refinement increases the simulation time of the models, making it more computationally costly. Therefore, the value 6 for the *timestep* was chosen because it is a good balance between accuracy and computational cost, meeting the objective of a metamodel.

For the final refinement process, the metamodel that performed best was subjected to two new simulations with different *timesteps*, 30 (intermediate refinement) and 60 (maximum refinement), in order to observe the impact of this change on the results.

4 Results and Discussions

Table 3 shows the characteristics of the calibrated model and metamodels in terms of the number of thermal zones, surfaces, and simulation time. It can be seen that the simplification of the calibrated model via the reduction of thermal zones had a direct impact on the simulation time: as the number of zones and surfaces decreases, the total processing time is reduced almost proportionally. Metamodel 1, consisting of only 13 thermal zones, recorded an average simulation time 77% lower than the original model, demonstrating the positive effect of controlled variable simplification on the computational performance of an energy model. This behavior is in line with one of the purposes of a metamodel: to make the simulation process more agile and accessible.

Table 3: Simulation times according to the most impactful characteristics of the models.

Metamodels	Original Model	Metamodel 1	Metamodel 2	Metamodel 3
Number of zones	55	13	21	21
Number of surfaces	844	213	300	349
Simulation time (s)	513	116	229	267

Source: the authors.

In addition to the reduction of thermal zones and consequent geometric simplification, it is worth emphasizing that simulation time is also linked to the number of *output* variables. In this case, simulations of both the calibrated model and the metamodels were performed considering only the *Site Outdoor Air Drybulb Temperature [°C] (Hourly)* and the *Zone Mean Air Temperature [°C] (Hourly),* mentioned earlier in this study. This makes it possible to evaluate the performance of all models using ASHRAE *Guideline* 14, as well as the statistical adjustment of polynomial curves below.

4.1 Performance according to ASHRAE *Guideline* 14

The ASHRAE *Guideline* 14 (2018) indicators are presented in Table 4 using the hourly temperature of each model for the nine regions of analysis, resulting from the combination of the three floors and the three main solar orientations of the building: southeast, southwest, and northwest. They are separated for the winter (W) and summer (S) periods according to Table 4.

Table 4: Results of the ASHRAE *Guideline* 14 indicators obtained for each metamodel.

Floor	Ground floor			First Floor			Second Floor		
Orientation	Southeast	Southwest	Northwest	southeast	southwest	northwest	southeast	southwest	northwest
Metamodel 1									
CV(RMSE) - W	2.19	4.86	4.30	3.99	5.35	4.72	3.48	5.14	6.77
NMBE – W	-0.68	-3.25	-1.80	-2.11	-4.04%	-3.05%	-2.07	-4.11%	-4.46%
CV(RMSE) - S	2.10	2.91	3.00	2.93	2.91%	3.19%	2.74	2.60	5.13
NMBE – S	-0.69	-1.22	-1.68	-1.46	-1.36%	-1.86	-1.50%	-1.25%	-3.77
Metamodel 2									
CV(RMSE) - W	2.48	8.80	8.55	4.23	5.55	9.01	3.71	9.05	10.64
NMBE – W	-1.14	-6.31	-5.30	-2.38	-4.15%	-6.65	-2.33%	-7.52%	-7.63%
CV(RMSE) - S	2.27	4.89	5.79	3.09	2.97	6.25	2.88	4.42	8.30%
NMBE – S	-0.94%	-2.02	-3.94	-1.63	-1.34%	-4.83	-1.65%	-2.51%	-6.74%
Metamodel 3									
CV(RMSE) - W	2.97	9.19	8.36	4.66	5.91	8.76	4.12	9.53	10.42
NMBE – W	-1.62%	-6.79%	-5.21%	-2.82	-4.56	-6.43%	-2.71%	-8.05%	-7.43%
CV(RMSE) - S	2.50	4.99	5.50	3.32	3.07	5.86	3.14	4.60	7.95
NMBE – S	-1.10%	-2.20	-3.58	-1.82	-1.50%	-4.38	-1.84%	-2.73%	-6.35

Source: the authors; * W = Winter, S = Summer.

Metamodel 1 performed best among the three evaluated, keeping NMBE and CV(RMSE) values at a maximum of |4.46|% and 6.77%, respectively, both on the second floor to the northwest. These maximums are lower than the limits established by ASHRAE, $\pm$ 10% and 30%. This suggests that, even with reduced and more simplified geometry when compared to the other two metamodels, the model was able to adequately and better represent the average temperature patterns of the calibrated model. Such behavior is not uncommon, as demonstrated by Gil (2017), a single-zone model (a model with radical simplifications of thermal zones in the *EnergyPlus* program) simulates temperature with almost negligible differences when compared to the reference multi-zone model, with a simulation time 70% lower. Westermann and Evins (2019) have already concluded that simple (low-order) models, or metamodels, are often more stable and faithfully reproduce the behavior patterns of more complex models, provided they are properly calibrated. Veiga *et al.* (2021) also observed in their research that their simplified model developed for naturally ventilated homes had 10% lower CV(RMSE) compared to the complex model, concluding that the simplifications managed to improve the thermal stability of the curves. In other words, complexity in computational building models does not necessarily mean refinement of results.

Still in Metamodel 1, the highest overestimations occurred on the northwest and southwest facades, especially on the second floor, regions more exposed to afternoon solar radiation and less protected by lateral shading. It was observed that, in the summer, the areas facing northwest and located on the second floor had the highest values for CV(RMSE) (up to 6.77%), and NMBE indicated a greater overestimation of temperatures in these areas. This behavior is due to greater exposure to radiation from the northwest exterior façade throughout the day, which receives the highest thermal loads from solar incidence. On the ground floor, although there are still overestimated values indicated by the negative sign in NMBE, the overestimation was lower, reaching maximum values of 4.86% and |3.25|% for CV(RMSE) and NMBE, respectively. This reflects greater thermal stability due to the inertia of the soil and possible self-shading caused by the building's own volume.

The southeast orientation showed the smallest deviations, with NMBE and CV(RMSE) more homogeneous across the pavements, ranging from 2.10% to 6.77% for CV(RMSE) and |0.68|% and |4.43|% for RMSE. After all, it tends to receive solar radiation early in the day, when temperatures are lower, which reduces thermal peaks and provides a more stable temperature distribution throughout the hours. In addition, this orientation is more influenced by the external shading element of the building, which explains the milder temperatures. In all three metamodels analyzed, this solar orientation generated the most satisfactory and representative performances.

Metamodel 2 was created to resolve the thermal overestimation of the first model and optimize it by adding thermal zones. As a result, the maximum and minimum values of NMBE and CV(RMSE) increased slightly, with minimums ranging from 2.10% to 2.27% (CV(RMSE)) and |0.68|% to |0.94|% (NMBE), and maximums ranging from 6.77% to 10.64% (CV(RMSE)) and |4.46|% to |7.63|% (NMBE). Although still within ASHRAE limits, it was noted that the added complexity to the metamodel did not provide better results. The increase in deviation compared to Metamodel 1 is due to greater solar exposure of the building's surfaces, which now have more zones and a larger area exposed to radiation throughout the day. In addition, the internal self-shading that was previously projected by the southwest-facing classrooms of the model during the morning was compromised by the geometric increase of the model. This directly impacts the number of hours of solar exposure of the surfaces of Metamodel 2, especially those facing the PVB's internal courtyard, which was where the aforementioned shading occurred.

Metamodel 3 introduced internal shading elements with the aim of reducing thermal gains from direct solar radiation in the morning, caused by the geometric increase of Metamodel 2, and solving the overestimation of the average temperature of the southwest and northwest zones already described. A slight reduction was observed in the maximum CV(RMSE) and NMBE of the second floor to the northwest (the most critical region), going from 10.64% to 10.42% (CV(RMSE)) and |7.63|% to |7.43|% (NMBE). However, the minimum values increased slightly, from 2.27% to 2.50% (CV(RMSE)) and |0.94|% to |1.10|% (NMBE) on the ground floor in the southeast. This suggests that the added shading elements, although they generated positive results in orientations and floors with a greater tendency for thermal discomfort due to heat, were not sufficient to offset the thermal gains caused by the addition of thermal zones in the model. Metamodel 3 was not improved in its entirety, but only partially.

All metamodels had overestimated results, i.e., higher temperature averages when compared to the calibrated model. The southwest and northwest orientations showed more critical thermal performance due to the area of solar exposure, which, in addition to being larger in metamodels 2 and 3, is affected by solar radiation in the afternoon when temperatures are higher. This effect is reflected in higher CV(RMSE) values, as already presented. The southeast-facing classrooms had the lowest errors due to their surface area, which, in addition to being smaller in metamodels 2 and 3 when compared to other orientations, receives solar radiation only during the morning, when temperatures are milder. In the afternoon, the region is self-shaded by the building's own volume, attenuating thermal loads accumulated by the building in this orientation.

An interesting point to evaluate is the tendency to overestimate metamodels, indicated by negative NMBE in all cases. This trend was also observed in the calibrations of the original model calibrated by Souza (2024), with no exact reason to justify this behavior, which may originate from the climate files used in the simulations. Souza (2024) concludes in his studies that the accuracy of the simulations is directly related to the quality of the climate data, as well as the position of the classrooms and their exposure to radiation and wind. In addition, floor conditions (rooms located on the ground floor or third floor) also influence the accuracy of the model in question.

Another trend also consistent with Souza's (2024) studies is that the results for the winter period are less consistent when compared to the results obtained in the summer period, although he does not discuss the possible reasons for this. However, there are some differences in thermal behavior patterns when comparing the results obtained in the metamodel simulations with those presented by Souza (2024) for equivalent rooms and regions. In Souza's (2024) studies, room 307 of the calibrated model, equivalent to region 7 of this study (second floor, northwest orientation), showed the best calibration performance in relation to the data measured in the field. In contrast, in this same region, the metamodels showed the greatest deviations from the calibrated model. The reason for this divergence has not been proven, but it is important to note that the metamodels were calibrated using average data and, therefore, may interfere with the results obtained for each region analyzed. Even so, the deviations in both studies, this one and that of Souza (2024), were of the same order of magnitude and very similar: CV(RMSE) with maximums of 12.68% and NMBE ranging from +4.5% to -10.90%.

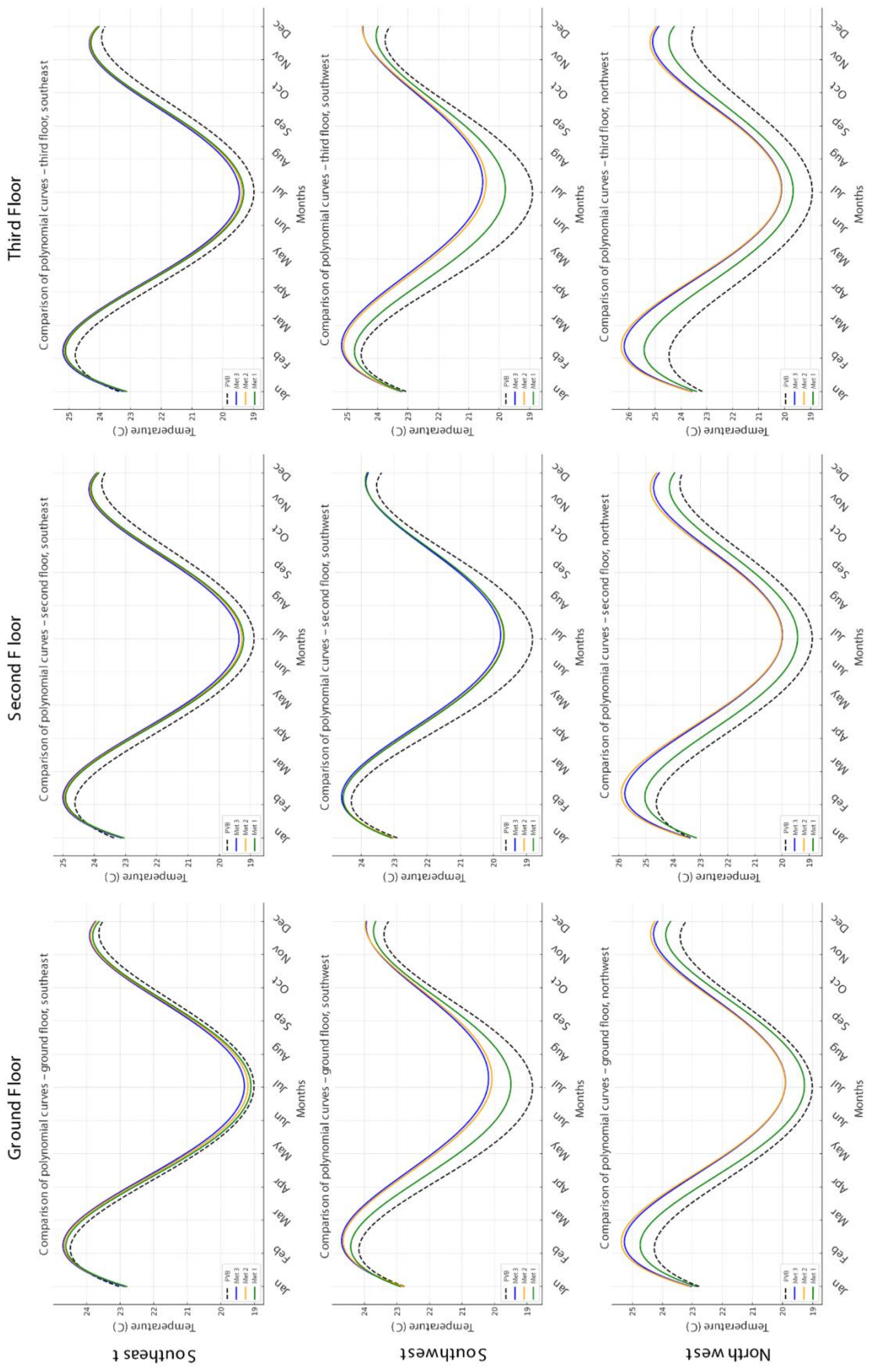
4.2 Performance according to polynomial curves

The polynomial regression curves (Figure 7) describe the seasonal behavior of the simulated hourly average temperatures for each of the nine regions analyzed. It can be observed that the curves of the metamodels, especially Metamodel 1, adequately reproduce the shape and trend of the calibrated model curves, with discrepancies of less than 2 °C over the annual period.

The polynomial analysis complements the indicators of ASHRAE *Guideline* 14 (2018) by allowing a continuous temporal reading of thermal variations. Unlike point metrics such as NMBE and CV(RMSE), the polynomial fit highlights the consistency of the building's dynamic response, identifying lags, amplitudes, and seasonal patterns that indicate thermal stability and good representation of the physical behavior of the models.

It can be seen that the curves that best fit the calibrated model curve (PVB) are those of the ground floor, especially those referring to the southeast-facing region, a behavior indicated by the ASHRAE metrics with the smallest deviations. It is also possible to notice the marked overestimation of the output data in the evaluated models compared to the calibrated one, which is consistent with the negative NMBE in all metamodels (Table 4).

Figure 7: Polynomial performance curves of the metamodels and the calibrated model.



In general, the curves obtained showed good similarity between the metamodels and the PVB, since the largest discrepancy between the metamodel and the calibrated model did not reach 2°C. This shows that thermal patterns over time were maintained. This statistical curve fitting analysis reinforces the findings obtained through ASHRAE *Guideline* 14, visually validating the metamodels and demonstrating that, even with geometric simplifications, it is possible to maintain the consistency of results in computer simulations.

4.3 Refinement via *timesteps*

In order to refine the results, Metamodel 1 was subjected to two new simulations with *timesteps* adjusted to 30 and 60. The results can be seen in Table 5. The variations in the indicators between the different *timesteps* were marginal, only in the second decimal place, with no statistically relevant gains in accuracy.

Table 5: Results of the comparison of *timesteps* (6, 30, and 60) obtained for Metamodel 1.

Pavement Orientation	Ground floor		First Floor				Second Floor		
	Southeast	Southwest	Northwest	southeast	southwest	northwest	southeast	southwest	northwest
***Timestep* 6**									
CV(RMSE) - W	2.19	4.86	4.30	3.99	5.35	4.72	3.48	5.14	6.77
NMBE – W	-0.68	-3.25	-1.80	-2.11	-4.04%	-3.05%	-2.07	-4.11%	-4.46%
CV(RMSE) - S	2.10	2.91	3.00	2.93%	2.91%	3.19	2.74	2.60	5.13
NMBE – S	-0.69	-1.22	-1.68	-1.46	-1.36%	-1.86	-1.50	-1.25%	-3.77%
***Timestep* 30**									
CV(RMSE) - W	2.15	4.79	4.26	3.93	5.29	4.67	3.41	5.07	6.70
NMBE – W	-0.63	-3.18	-1.72	-2.05	-3.98	-2.95%	-1.96%	-4.04%	-4.34%
CV(RMSE) - S	2.08	2.86	2.97	2.91	2.87	3.16	2.69	2.55	5.07
NMBE – S	-0.64	-1.16	-1.63	-1.40	-1.29%	-1.79	-1.42	-1.18%	-3.69%
***Timestep* 60**									
CV(RMSE) - W	2.14	4.78	4.27	3.92	5.28	4.67	3.40	5.06	6.70
NMBE – W	-0.63	-3.18	-1.73	-2.04	-3.97	-2.95	-1.95	-4.04	-4.34
CV(RMSE) - S	2.08	2.85	2.97	2.90	2.86	3.15	2.68	2.55	5.06
NMBE – S	-0.64	-1.15	-1.62	-1.39	-1.29	-1.79	-1.41	-1.17	-3.68

* W = Winter, S = Summer.

On the other hand, simulation time increased from 116 seconds (6 timesteps/h) to 870 seconds (60 timesteps/h), with *timestep* 30 being the intermediate value between the two, with a simulation time equal to 446 seconds. Given the negligible gain in accuracy compared to the increase in computational cost caused by the 650% increase in simulation time, 6 timesteps/h was adopted as the standard. It is therefore conclusive that a *timestep* value of 6 is sufficient for consistent results in thermo-energy performance analyses with *EnergyPlus,* especially when it comes to metamodeling.

Table 6: Impact of *timestep* on the refinement of Metamodel 1 results.

No. of timesteps/h	Simulation time (s)	Percentage increase (%)
6	116	0
30	446	284
60	870	650

*reference value.

5 Conclusion

The objective of this study was to develop and validate a thermo-energetic metamodel capable of representing the thermal behavior of an existing and previously validated computational model, with lower computational cost and without compromising the accuracy of the results. In addition to evaluating which simplifications can be implemented in the model during the process.

The results obtained demonstrate that the three metamodels evaluated were able to adequately represent the thermal behavior of the original calibrated model, respecting the error limits established by ASHRAE, a maximum of $\pm$ 10% for NMBE and a maximum of 30% for CV(RMSE). However, Metamodel 1, the most simplified model of the three, obtained the best results with a maximum NMBE of |4.46|% and a maximum CV(RMSE) of 6.77%, both below the limits of ASHRAE *Guideline* 14 (2018), and this model was chosen for validation. Even though it was the most simplified model with the shortest simulation time (about 77% shorter than the calibrated model), it maintained the thermal standards within the technical criteria and with due visual adherence to the curves of the calibrated model.

This shows that the strategic reduction in the number of thermal zones was an efficient solution to enable faster simulations without compromising the quality of the temperature results for the Case of Classroom Pavilion II, where there is a high degree of geometric and physical similarity between the modeled environments. These results confirm that metamodels with volumetric representations, such as in this study, can be simplified in form by reducing thermal zones or equivalent processes. In addition, they can reproduce the average temperature patterns of complex models with high fidelity, provided that the consistency of the thermal zones and control over the variables are maintained.

The cross-analysis of the results based on quantitative indicators (NMBE and CV(RMSE)) and adjustments by polynomial curves add robustness to the validation of the proposed metamodels and represent a methodological contribution to future studies. The ASHRAE indices allowed the variability between models to be assessed, while the polynomial curves enabled the visualization of thermal behavior over time, highlighting seasonal patterns and distortions that could go unnoticed in purely numerical analyses.

The use of averages per building region, defined by floor and orientation, was a possible strategy to enable comparison between the original model and the metamodels, given the difference in the number of thermal zones. In addition, it can serve as a basis for future reintegration of zones from the metamodels, making it possible, for example, to deepen the detail only in identified critical areas. However, it should be noted that this strategy represents an additional simplification in the model and, as such, may attenuate relevant internal variations that would occur between individual zones.

The hypothesis that refining the *timestep* parameter could add robustness to the best-performing metamodel was discarded. This is because the improvements in results were negligible compared to the increase in computational time required for the simulations.

However, some limitations must be acknowledged. The absence of modeling of occupancy, lighting, and internal loads restricts the extrapolation of results for absolute thermal comfort and energy consumption analyses. Despite being naturally ventilated and relying on ventilation modeling via *AirflowNetwork*, calibration was conducted based only on air temperature analyses. In addition, the comparison was made using simulated reference data rather than directly measured data, which directly impacts its reliability when compared to the actual behavior of the building. For future work, it is recommended that, in addition to analyzing the air exchanges of the models, comparisons be made with measured data, if available, increasing the level of fidelity of the metamodel and exploring its potential in relation to energy issues.

Thus, in addition to presenting a validated metamodel for representing PVB, this study also contributes a methodological strategy for advancing energy simulations in the field of architecture, promoting more accessible and replicable solutions in academic and professional environments.

Acknowledgments

This work was carried out with support from the National Council for Scientific and Technological Development (CNPq) - Project Process 406426/2022-8, and with partial support from the Ibero-American Program of Science and Technology for Development - CYTED through the TRAPECIO Network. In addition, financial resources from other sources include the doctoral scholarship awarded to Ligiana Pricila Guimarães Fonseca by Capes Brasil (funding code 001) through the Graduate Program in Architecture and Urbanism PPG.au/UFV and international mobility. Finally, we would like to thank CNPq under the Institutional Scientific Initiation Program PIBIC/CNPq 2023-2024 and 2024-2025.

References

ALI, Usman; SHAMSI, Mohammad Haris; HOARE, Cathal; MANGINA, Eleni; O'DONNELL, James. Review of urban building energy modeling (UBEM) approaches, methods, and tools using qualitative and quantitative analysis. **Energy and Buildings**, v. 246, 2021. DOI: https://doi.org/10.1016/j.enbuild.2021.111073.

ARENDT, Krzysztof, *et al*. "Comparative analysis of white-, gray-and black-box models for thermal simulation of indoor environment: Teaching building case study." *In:* Building Performance Analysis Conference and SimBuild: Co-organized by ASHRAE and IBPSA-USA. ASHRAE, 2018.

ASHRAE. **ASHRAE Guideline 14-2014** (Reaffirmed 2018): Measurement of Energy, Demand, and Water Savings. Atlanta: ASHRAE, 2018.

ASHRAE – American Society of Heating, Refrigerating and Air-Conditioning Engineers. **ASHRAE Guideline 14-2023***:* Measurement of Energy, Demand, and Water Savings. Atlanta: ASHRAE, 2023.

BOUCHLAGHEM, N. M.; LETHERMAN, K. M. Numerical optimization applied to the thermal design of buildings. **Building and environment**, v. 25, n. 2, p. 117-124, 1990.

COAKLEY, Daniel; RAFTERY, Paul; KEANE, Marcus. A review of methods to match building energy simulation models to measured data. **Renewable and Sustainable Energy Reviews**, v. 37, p. 123–141, 2014. DOI: https://doi.org/10.1016/j.rser.2014.05.007.

CUI, C.; HU, M.; WEIR, J. D.; WU, T. A recommendation system for meta-modeling: a meta-learning based approach. **Expert Systems with Applications**, v. 46, p. 33–44, 2016. DOI: 10.1016/j.eswa.2015.10.021.

DONOVAN, O.; PAUL; MURPHY, M. D. Predicting air temperatures in a naturally ventilated nearly zero energy building: Calibration, validation, analysis and approaches. **Applied Energy**, v. 250, p. 991–1010, 2019. Available at: <https://doi.org/10.1016/j.apenergy.2019.04.082>.

FOUCQUIER, A.; *et al.* State of the art in building modeling and energy performance prediction: A review. **Renewable and Sustainable Energy Reviews**, v. 23, p. 272-288, 2013.

GIL, María del Pilar Casatejada. **Simplificações na modelagem de habitações de interesse social no programa de simulação de desempenho térmico EnergyPlus**, 2017. Dissertation (Master's Degree in Architecture, Urbanism, and Technology) - Institute of Architecture and Urbanism, University of São Paulo, São Carlos, 2017. DOI: 10.11606/D.102.2018.tde-12012018-103257. Accessed on: Apr. 4, 2025. Portuguese.

IPCC, Working Group II. Technical Summary. *In:* Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate ChangeClimate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems. Cambridge; New York. DOI: 10.1017/9781009325844.002.

KAVGIC, Miroslava *et al.* A review of bottom-up building stock models for energy consumption in the residential sector. **Building and environment**, v. 45, n. 7, p. 1683-1697, 2010.

LOPES, Adriano Felipe Oliveira; SILVA, Caio Frederico e; AMORIM, Cláudia Naves David; BATISTA, Juliana Oliveira. Avaliação do desempenho térmico de ambiente escolar padronizado, em contexto climático brasileiro, por meio de simulação termoenergética. **PARC: Pesquisa em Arquitetura e Construção**, Campinas, SP, v. 14, n. 00, e023030, 2023. DOI: 10.20396/parc.v14i00.8670652. Available at: https://periodicos.sbu.unicamp.br/ojs/index.php/parc/article/view/8670652. Accessed on: Oct. 17, 2025. Portuguese.

LUCARELLI, Caio de Carvalho; OLIVEIRA, Matheus Menezes; CARLO, Joyce Correna. Comparative analysis of Viçosa's weather files: simulation adequacy for urban microclimate. *In:* PASSIVE AND LOW ENERGY ARCHITECTURE CONFERENCE – PLEA, 37., 2022, Santiago. **Proceedings of the Passive and Low Energy Architecture Conference 2022**. Santiago: PLEA, 2022. Available at: https://www.researchgate.net/publication/371444251.

MACIEL, Thalita dos Santos; LEITZKE, Rodrigo Karini; DUARTE, Carolina de Mesquita; SCHRAMM, Fábio Kellermann; CUNHA, Eduardo Grala da. Otimização termoenergética de uma edificação escolar: discussão sobre o desempenho de quatro algoritmos evolutivos multiobjetivo. **Ambiente Construído**, Porto Alegre, v. 21, n. 4, p. 67–86, Oct./Dec. 2021. DOI: https://doi.org/10.1590/s1678-86212021000400567. Portuguese

OLINGER, M. S.; MELO, A. P.; NEVES, L. O.; LAMBERTS, R. Surrogate model development for naturally ventilated office buildings. *In:* BUILDING SIMULATION 2019: 16th International Conference of the International Building Performance Simulation Association (IBPSA), Rome, Italy, Sept. 2–4, 2019. **Proceedings of Building Simulation**

2019. Rome: International Building Performance Simulation Association, 2019. p. 1396–1403. DOI: https://doi.org/10.26868/25222708.2019.210542.

SILVA, OLINGER, Marcelo Salles; MELO, Ana Paula; LAMBERTS, Roberto. Developing a surrogate model for naturally ventilated cellular offices in Brazil. **Building and Environment**, v. 233, p. 110075, 2023.

ØSTERGÅRD, Torben; JENSEN, Rasmus L.; MAAGAARD, Steffen E. Early Building Design: Informed decision-making by exploring multidimensional design space using sensitivity analysis. **Energy and Buildings**, v. 142, p. 8-22, 2017.

SAKIYAMA, N. R. M.; MAZZAFERRO, L.; CARLO, J. C.; BEJAT, T.; GARRECHT, H. Natural ventilation potential from weather analyses and building simulation. **Energy and Buildings**, v. 231, 2021. DOI: https://doi.org/10.1016/j.enbuild.2020.110596.

SÁNCHEZ-ZABALA, Víctor F.; GÓMEZ-ACEBO, Tomás. Building energy performance metamodels for district energy management optimization platforms. **Energy Conversion and Management**: X, v. 21, 2024. DOI: https://doi.org/10.1016/j.ecmx.2023.100512.

SHI, Xing; TIAN, Zhichao; CHEN, Wenqiang; SI, Binghui; JIN, Xing. A review on building energy efficient design optimization from the perspective of architects. **Renewable and Sustainable Energy Reviews**, v. 65, p. 872-884, 2016. DOI: https://doi.org/10.1016/j.rser.2016.07.050.

SILVA, Arthur Santos; GHISI, Enedir. Uncertainty analysis of the computer model in building performance simulation. **Energy and Buildings**, v. 76, p. 258-269, 2014.

SOUZA, Pedro Carmo e. **Effect of natural ventilation on airborne disease infection risk in lecture halls**. 2024. 111 f. Dissertation (Master's Degree in Architecture and Urbanism) - Federal University of Viçosa, Viçosa. 2024. Portuguese.

SUN, Han; BURTON, Henry; HUANG, Honglan. Machine learning applications for building structural design and performance assessment: state-of-the-art review. **Journal of Building Engineering**, v. 33, 2020. DOI: https://doi.org/10.1016/j.jobe.2020.101816.

TOULOUPAKI, Eleftheria; THEODOSIOU, Theodoros. Optimization of building form to minimize energy consumption through parametric modeling. **Procedia Environmental Sciences**, v. 38, p. 509–514, 2017. DOI: https://doi.org/10.1016/j.proenv.2017.03.114.

U.S. DEPARTMENT OF ENERGY. **EnergyPlus Engineering Reference: The Reference to EnergyPlus Calculations**. Version 25.1.0. Washington, D.C.: U.S. Department of Energy, 2025. Available at: https://energyplus.net/documentation. Accessed on: Apr. 4, 2025.

VEIGA, Rodolfo Kirch; ELI, Letícia Gabriela; KRELLING, Amanda F.; OLINGER, Marcelo Salles; others. Development of a metamodel to assess building thermal performance s for naturally ventilated residential buildings. **Proceedings of Building Simulation 2021**, 2021.

WESTERMANN, Paul; EVINS, Ralph. Surrogate modeling for sustainable building design–A review. **Energy and buildings**, v. 198, p. 170-186, 2019.

WMO - World Meteorological Organization. **State of the Global Climate 2022**. Geneva, Switzerland: WMO.

WONG, Cyrus Ho Hin; CAI, Meng; REN, Chao; HUANG, Ying; LIAO, Cuiping; YIN, Shi. Modelling building energy use at urban scale: a review on their account for the urban environment. **Building and Environment**, v. 205, 2021. DOI: https://doi.org/10.1016/j.buildenv.2021.108235.

WORTMANN, Thomas; CICHOCKA, Judyta; WAIBEL, Christoph. Simulation-based optimization in architecture and building engineering - Results from an international user survey in practice and research. **Energy and Buildings**, v. 259, 2022. DOI: https://doi.org/10.1016/j.enbuild.2022.111863.

XIA, D.; WU, Z.; ZOU, Y. Developing a bottom-up approach to assess energy challenges in urban residential buildings of China. **Frontiers of Architectural Research**, [S.l.]: Elsevier/KeAi, 2025. DOI: https://doi.org/10.1016/j.foar.2025.03.006.

YANG, Song; TIAN, Wei; CUBI, Eduard; MENG, QingXin; LIU, YunLiang; WEI, Lai. Comparison of sensitivity analysis methods in building energy assessment. **Procedia Engineering**, v. 146, p. 174–181, 2016. DOI: https://doi.org/10.1016/j.proeng.2016.06.369.

YU, Xingji; GEORGES, Laurent; KNUDSEN, Michael D.; SARTORI, Igor; IMSLAND, Lars. Investigation of the model structure for low-order grey-box modeling of residential buildings. *In:* INTERNATIONAL BUILDING PERFORMANCE SIMULATION ASSOCIATION CONFERENCE, 16., 2019, Rome. **Proceedings of the International Building Performance Simulation Association Conference**. Rome: IBPSA, 2019. DOI: https://doi.org/10.26868/25222708.2019.211209.