buildings

MDPI

Article

Residual Learning Enhanced Grey-Box Modelling for Indoor Temperature Prediction and IEQ Assessment

Constantin Cilibiu *, Horatiu Calin Albu and Ancuta Coca Abrudan

Building Services Engineering Department, Faculty of Building Services Engineering, Technical University of Cluj-Napoca, 400114 Cluj-Napoca, Romania; horatiu.albu@insta.utcluj.ro (H.C.A.); ancuta.abrudan@insta.utcluj.ro (A.C.A.)

* Correspondence: constantin.cilibiu@insta.utcluj.ro; Tel.: +40-724-311-138

Abstract

The increasing demand for the energy-efficient and occupant-centred operation of educational buildings requires accurate and interpretable models capable of predicting indoor environmental conditions under real operating constraints. This study proposes a residual learning-enhanced grey-box modelling framework for predicting indoor air temperature and assessing indoor environmental quality indicators in a KNX-enabled educational building operating under simple thermostatic heating control. The approach combines a reduced-order discrete-time RC thermal model with a data-driven machine learning component trained to model the next-step residual between measured and simulated indoor temperatures. High-resolution KNX monitoring data were recorded at a 5 min sampling interval over three consecutive months (October–December) during the heating season. Using a chronological 70/30 train–test split, the identified RC grey-box model achieved a pooled test RMSE of 0.269 °C, an MAE of 0.126 °C, and an R^2 of 0.987. The proposed hybrid formulation achieved RMSE = 0.343 °C, MAE = 0.106 °C, and R^2 = 0.978 across 62,456 test samples. While the pooled RMSE remains influenced by occasional larger deviations in a small number of rooms, the hybrid model yields a consistent reduction in absolute error (≈16% MAE reduction) and reduced inter-room variability compared to the physics-based baseline. These results indicate that residual learning can enhance predictive robustness under decentralized thermostatic operation and limited sensing, while preserving physical interpretability. The proposed framework provides a practical and scalable solution for indoor temperature prediction and IEQ assessment in educational buildings using existing KNX automation data.

Keywords: residual-based hybrid modelling; grey-box modelling; indoor air temperature prediction; indoor environmental quality; KNX building automation; educational buildings; machine learning

Check for updates

Academic Editor: Constantinos A. Balaras

Received: 28 January 2026
Revised: 19 February 2026
Accepted: 24 February 2026
Published: 1 March 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The global pursuit of sustainable and energy-efficient building operation has intensified in recent years due to the combined challenges of climate change, rising energy demand, and an increasing emphasis on occupant well-being. Buildings account for nearly 40% of global energy consumption and a substantial share of CO_2 emissions, making the performance assessment and informed operation of heating, ventilation, and air conditioning (HVAC) systems a critical priority for achieving environmental objectives and ensuring adequate indoor environmental quality (IEQ) [1,2].

Accurate prediction and assessment of key IEQ parameters—such as indoor air temperature, CO_2 concentration, and volatile organic compounds (VOC)—are essential for maintaining thermal comfort and indoor air quality while minimizing energy use [3]. However, indoor environments are inherently dynamic, being strongly influenced by occupant behaviour, weather variability, and ventilation practices, which poses significant challenges to conventional model-based HVAC analysis and control approaches [4].

Traditional physics-based and grey-box models, grounded in thermodynamic principles, provide interpretability and physical consistency but often require extensive calibration and detailed knowledge of building properties [5,6]. Their predictive accuracy may deteriorate under highly variable occupancy and ventilation conditions, which are typical for educational buildings. In contrast, purely data-driven machine learning (ML) models can capture nonlinear relationships from operational data but may lack physical interpretability and robustness when applied outside their training domain [7]. This trade-off between predictive accuracy and physical consistency has motivated growing research interest in hybrid modelling approaches and related physics-informed machine learning formulations discussed in the literature [8].

Recent studies have demonstrated the potential of hybrid modelling frameworks for predicting building thermal dynamics, energy consumption, and indoor air quality [9–11]. Nevertheless, many reported applications rely on simulated environments, centralized HVAC systems, or extensive sensing infrastructures. In practice, numerous educational buildings operate under simple thermostatic or rule-based HVAC control strategies and provide limited access to advanced building management systems. Consequently, there remains a need for hybrid modelling approaches that are validated using real operational data from intelligent automation systems, such as KNX, under realistic constraints.

In this context, the present study proposes a hybrid physics-based and machine learning framework for predicting indoor air temperature and assessing indoor environmental quality in a KNX-enabled educational building located at the Faculty of Building Services Engineering, Cluj-Napoca, Romania. The proposed approach integrates a reduced-order physics-based thermal model with a data-driven component that models residual effects associated with occupancy-driven internal gains and other unmodeled disturbances under real operating conditions.

1.1. Related Work

The prediction and performance assessment of indoor environmental quality and HVAC-related energy use have become key research areas in building performance engineering. Traditionally, two main modelling paradigms have been employed: physics-based models (white-box or grey-box) and machine learning models (black-box). Each presents unique advantages and limitations in capturing building thermal dynamics and occupant-driven variability [1,3,12].

1.2. Physics-Based and Grey-Box Approaches

Physics-based models rely on thermodynamic equations describing heat and mass transfer within building envelopes and HVAC systems [4]. While these models offer interpretability and physical consistency, their calibration requires extensive parameter identification and detailed knowledge of the building geometry and materials [13]. Simplified "grey-box" models, such as resistance-capacitance (RC) networks, have been widely adopted for their balance between physical realism and computational efficiency [14,15]. Gouda et al. [5] demonstrated early on that reduced-order thermal models could effectively approximate dynamic indoor temperature profiles, though their accuracy deteriorates under highly variable occupancy and ventilation patterns. In subsequent studies, short-

horizon temperature prediction errors for calibrated reduced-order RC models have typically been reported within a few tenths of a degree Celsius, depending on the identification strategy and boundary conditions [5,14–16].

Recent grey-box modelling approaches, as reviewed by Zhang et al. [6] and further advanced through physics-based and hybrid modelling strategies demonstrated by Drgoňa et al. [16], have shown improved generalization by combining first-principles equations with statistical calibration. However, these methods still struggle to incorporate behavioural factors such as intermittent occupancy, which strongly influence IEQ in educational buildings [17].

Reported short-horizon indoor temperature prediction errors for low-order RC models typically fall within sub-degree ranges (often below 0.5 °C RMSE for one-step-ahead prediction), with performance strongly dependent on calibration quality, occupancy variability, and boundary condition accuracy [6,14–16].

1.3. Data-Driven and Machine Learning Approaches

Machine learning models have gained prominence due to their ability to learn complex, nonlinear relationships directly from operational data. Techniques such as random forest, support vector regression (SVR), and neural network-based models have been applied for short-term temperature and CO_2 concentration forecasting [7,18–20]. For instance, Fan et al. [7] demonstrated that deep learning algorithms can accurately predict cooling loads with minimal feature engineering, reporting coefficients of determination above 0.9 and mean absolute percentage errors in the single-digit to low double-digit range for short-term prediction under controlled datasets. Similarly, artificial neural network-based models have been applied to model building thermal dynamics with high predictive accuracy, while often requiring substantial training data and exhibiting limited adaptability to changing operating conditions [7,18].

Despite their predictive power, purely data-driven models often suffer from limited physical interpretability and reduced robustness outside their training domain, which may compromise their reliability for building performance assessment under varying operational conditions [21]. This issue has motivated the emergence of hybrid grey-box approaches, as well as physics-informed machine learning models reported in the literature.

Purely data-driven approaches frequently report high coefficients of determination ($R^2 > 0.9$) under stable operating conditions; however, their performance may degrade significantly when extrapolating beyond the training domain or under sparse data scenarios [7,18].

1.4. Hybrid Approaches

Hybrid modelling approaches combining grey-box structures with data-driven components have attracted increasing research interest, alongside physics-informed machine learning formulations reported in the literature [8,22]. In parallel, learning-based control and optimization approaches for building HVAC systems have also attracted significant research interest, demonstrating the potential of artificial intelligence to enhance building performance under complex operational conditions [23]. In these models, domain knowledge from thermodynamics is embedded within ML architectures, guiding the learning process and improving extrapolation under unseen conditions. Karniadakis et al. [8] formalized the concept of physics-informed machine learning, demonstrating its ability to integrate governing equations into neural networks for scientific modelling.

Applications in building energy systems are expanding rapidly. Gokhale et al. [24] reported measurable reductions in energy consumption and improved comfort stability when integrating physics-based constraints within data-driven control frameworks, with

reported savings in the order of low double-digit percentages depending on operational scenarios. Similarly, other physics-informed machine learning studies have reported relative prediction error reductions typically ranging between 5% and 20% compared to standalone data-driven or grey-box baselines, particularly for short-term forecasting tasks and sensor-limited configurations [24,25].

In this paper, we use the term residual-based hybrid grey-box/ML modelling to describe an identified discrete-time RC grey-box thermal formulation augmented by data-driven residual correction. This should be distinguished from physics-informed neural network formulations in which governing equations are embedded directly into the learning architecture.

More recent studies have highlighted the importance of accurate and well-identified building thermal models for advanced HVAC control applications, including model predictive control (MPC) [26]. However, the present study focuses on prediction and performance assessment rather than implementing MPC. While these studies report substantial energy savings and improved comfort, their implementation typically relies on centralized control architectures and extensive system integration.

In contrast, many educational buildings operate under simple thermostatic or rule-based HVAC control strategies, with limited access to advanced control layers. In such contexts, hybrid ML models can provide significant value by improving the prediction and assessment of indoor environmental quality and thermal dynamics using real operational data.

In educational buildings, where occupant behaviour and ventilation patterns vary significantly, such predictive capabilities are particularly relevant. Rus et al. [9] demonstrated how pandemic-related ventilation protocols impacted students' thermal comfort, while Rus et al. [10] highlighted the interplay between IEQ, energy performance, and environmental impacts in Romanian university buildings. These findings motivate the present study, which focuses on hybrid grey-box/ML-based IEQ prediction using real KNX automation data from an educational building operating under simple HVAC control.

While residual-based hybrid modelling approaches are well-established in the literature, their empirical validation under decentralized KNX-operated educational buildings with simple thermostatic control and limited sensing remains limited. The present work contributes by assessing the consistency of such an approach across multiple rooms under realistic operational constraints.

The main contributions of this study are (i) the implementation and evaluation of a residual-based hybrid grey-box and machine learning framework using real KNX automation data; (ii) a comparative assessment against identified grey-box and purely data-driven baselines under decentralized thermostatic control; and (iii) an analysis of model robustness across lecture rooms with heterogeneous occupancy patterns.

Despite the growing body of research on hybrid and data-driven building modelling, limited empirical evidence is available regarding the performance of residual-based hybrid grey-box approaches under real decentralized control conditions and restricted sensing infrastructures typical of many retrofitted educational buildings. In practice, such buildings often operate with simple thermostatic logic and without comprehensive ventilation monitoring, which limits the direct applicability of advanced control-oriented modelling frameworks reported in the literature.

In this context, the objective of the present study is not to propose a new hybrid modelling methodology, but to evaluate the practical effectiveness and robustness of a residual-based hybrid grey-box and machine learning framework when applied to a KNX-operated educational building characterized by simple room-level control and limited sensing.

More specifically, this study addresses the following questions:

(i) Can residual-based hybrid modelling provide statistically meaningful improvements over an identified RC grey-box model under decentralized thermostatic operation?
(ii) How consistent are such improvements across rooms with heterogeneous occupanc patterns?
(iii) To what extent can KNX-derived operational data support integrated indoor environmental quality assessment under incomplete ventilation information?

By explicitly framing these questions, the study aims to provide a realistic evaluation of hybrid modelling under practical constraints rather than advancing a new methodological paradigm.

Hybrid and physics-informed approaches commonly report relative error reductions in the single-digit to low double-digit percentage range compared to standalone grey-box or black-box baselines, particularly for short-term prediction horizons and sensor-limited environments [8,24,25].

While physics-informed neural networks embed governing equations directly into the training loss, the present work adopts a residual-based hybrid grey-box/ML formulation in which physical structure is preserved explicitly and corrected through data-driven residual learning.

2. Materials and Methods

2.1. Case Study and KNX Data Acquisition

The study was conducted in an educational building belonging to the Faculty of Building Services Engineering, Cluj-Napoca, Romania, equipped with a KNX-based building automation system. The monitored spaces consist of lecture rooms characterized by highly variable occupancy schedules, a common feature in educational buildings, particularly under free-running or naturally ventilated conditions [17].

The heating system is based on room-level hydronic fan coil units operating under simple thermostatic control.

Basic geometric characteristics of the monitored lecture rooms, including floor area, window area, and façade orientation, are provided in Supplementary Table S2.

The KNX system provides time-series measurements of indoor air temperature, relative humidity, CO_2 concentration, volatile organic compound index, and HVAC actuator signals. Data were recorded at a 5 min sampling interval over three consecutive months (October–December) during the heating period, corresponding to active teaching weeks, capturing representative indoor environmental dynamics under real operating conditions [10].

The KNX indoor environmental sensors used in the monitored rooms correspond to commercially available multi-sensor units typically deployed in educational buildings. According to manufacturer specifications, the temperature measurement accuracy is $\pm 0.3\ ^{\circ}C$, while CO_2 concentration accuracy is $\pm$(30–50 ppm + 3% of reading), depending on calibration conditions. Such uncertainty levels are consistent with standard NDIR-based CO_2 sensing technology commonly used in building automation systems.

These measurement uncertainties should be considered when interpreting short-term CO_2 variability and prediction errors; however, they do not materially affect the comparative evaluation between modelling approaches.

Outdoor air temperature was not directly measured on site. Instead, hourly outdoor temperature values were obtained from the national climatic annex corresponding to the building location (Cluj-Napoca). Such data are commonly used in building performance studies when local weather station measurements are unavailable and have been shown to provide sufficiently representative boundary conditions for thermal modelling [6].

Although climatic annex data do not replace on-site weather station measurements, they provide a standardized and widely adopted boundary reference in building performance studies when local meteorological monitoring is unavailable.

2.2. HVAC Operation Context

The HVAC system serving the monitored rooms consists of hydronic fan coil units operating under local thermostatic (on–off) control, a control strategy widely adopted in existing educational and office buildings [4]. In each room, the fan coil unit is activated when the indoor air temperature falls below a predefined setpoint and deactivated once the setpoint is reached.

No centralized supervisory control or model predictive control is implemented. Consequently, the proposed hybrid grey-box/ML framework is not used to directly actuate the HVAC system, but rather to improve the prediction and assessment of indoor environmental quality under simple control regimes, as commonly considered in building performance assessment studies [21].

The heating actuator signal $u_{hvac}(t)$ recorded in the dataset corresponds to the binary state $(0/1)$ transmitted on the KNX bus by the local room thermostat. Any internal hysteresis or deadband logic around the temperature setpoint is executed locally within the KNX controller and is not explicitly reconstructed in this study. Instead, the measured actuator state is directly used within the grey-box formulation, ensuring that the modelling framework reflects the actual operational logic implemented in the building.

The fan operation of the fan coil units only involves air recirculation within the room and does not introduce outdoor air or provide mechanical ventilation, a characteristic of such terminal units widely reported in the HVAC literature [4].

The KNX thermostatic controllers implement an internal two-point control logic based on the deviation between measured indoor air temperature and the predefined setpoint. While the detailed controller configuration parameters (e.g., exact hysteresis width or switching thresholds) are not directly accessible within the recorded dataset, the resulting binary actuator state is fully captured through the KNX group objects.

Therefore, the modelling framework does not reconstruct the internal control algorithm. Instead, it directly uses the measured actuator signal $u_{hvac}(t)$, which implicitly embeds the implemented hysteresis and switching behaviour. This approach ensures that the thermal model reflects the actual operational behaviour of the installed KNX system.

Regarding data acquisition, the monitored variables were logged at fixed 5 min intervals using cyclic sampling of KNX group objects. The recorded values correspond to the last transmitted state available at each logging timestamp rather than averaged values, ensuring consistency with the discrete-time modelling resolution adopted in this study.

2.3. Data Preprocessing and Feature Construction

Occupancy was not directly measured. Instead, an occupancy proxy was inferred from CO_2 concentration dynamics and teaching schedules, an approach commonly adopted in educational building studies where direct occupancy sensing is unavailable [19].

Hourly outdoor air temperature values obtained from the national climatic annex were used directly as boundary conditions for the thermal model. Given the relatively slow dynamics of outdoor temperature variations compared to indoor thermal response, hourly outdoor temperature values were mapped to the 5 min grid using a piecewise-constant (zero-order hold) representation, i.e., each hourly value was held constant over the corresponding 5 min samples. This approach avoids introducing artificial variability and is consistent with common practices in reduced-order building thermal modelling [6,14–16].

Prior to model development, the KNX time-series data were preprocessed following standard practices in building performance analytics [20]. Short gaps in the data were handled using linear interpolation, while longer gaps were excluded from the analysis. All continuous variables were standardized using z-score normalization to ensure numerical stability during model training [27].

2.4. Discrete-Time RC Grey-Box Model

The grey-box reference model used in this study is implemented as a discrete-time RC grey-box formulation identified directly from measured KNX data. Rather than a detailed multi-node thermal network, the model captures dominant thermostat-driven dynamics and effective thermal inertia at the 5 min sampling resolution used in the study.

The discrete-time formulation adopted in this study is derived from the classical first-order lumped thermal energy balance of a single-zone building, expressed as:

$$C\frac{dT_{in}(t)}{dt} = \frac{T_{out}(t) - T_{in}(t)}{R} + Q_{hvac}(t) + Q_{int}(t) \quad (1)$$

where C represents the effective lumped thermal capacitance of the room, R denotes the overall thermal resistance between indoor and outdoor environments (including envelope and infiltration effects), $Q_{hvac}(t)$ represents the heating input, and $Q_{int}(t)$ accounts for internal gains and unmodeled disturbances.

Assuming a sampling interval $\Delta t = 5$ minutes and applying a forward Euler discretization, the continuous formulation can be expressed in discrete-time form. In the present grey-box identification approach, the resulting discrete coefficients are directly estimated from data, yielding the following parametric representation.

In discrete form, the model is defined as:

$$T_{in}(t+1) = \beta_0\, T_{in}(t) + \beta_1\, (T_{out}(t) - T_{in}(t)) + \beta_2\, u_{hvac}(t)(T_{set}(t) - T_{in}(t)) + \beta_3 \quad (2)$$

where $T_{in}(t)$ is the measured indoor air temperature at time step t, $T_{out}(t)$ is the outdoor air temperature used as an exogenous boundary condition, $u_{hvac}(t)$ is the binary heating actuator state (0/1), and $T_{set}(t)$ denotes the thermostat setpoint. The term $\beta_1\, (T_{out}(t) - T_{in}(t))$ represents the effective heat exchange with the outdoor environment (envelope/infiltration lumped effect), while $\beta_2\, u_{hvac}(t)(T_{set}(t) - T_{in}(t))$ captures the dominant thermostat-driven heating influence at the adopted sampling resolution. β_3 is a constant term accounting for average unmodeled gains and offsets.

This discrete-time formulation corresponds to a first-order lumped RC behaviour with an exogenous outdoor boundary condition and a thermostat-driven heating term at the adopted sampling resolution.

The parameter vector β is estimated independently for each monitored room using ordinary least squares (OLS) applied to the training segment of the chronologically split dataset. The estimation minimizes the squared difference between measured and predicted next-step indoor temperature. No additional regularization or parameter constraints are imposed.

The identified discrete-time coefficients were verified to ensure physically consistent and stable thermal dynamics. In particular, the effective autoregressive coefficient associated with indoor temperature remained within the stability range ($|\beta_0| < 1$) for all rooms, consistent with first-order lumped thermal behaviour.

2.5. Machine Learning Residual Modelling

All machine learning models were implemented in Python 3.12.12 using the scikit-learn library (version 1.6.1), together with NumPy (version 2.0.2) and Pandas (version 2.2.2), within a Google Colab runtime environment.

For the temperature benchmarks reported in Tables 1 and S1, a HistGradientBoostingRegressor was used both for the Pure ML baseline and for the residual model in the hybrid formulation.

Table 1. Predictive performance comparison (pooled test dataset across all monitored rooms).

Model	RMSE [°C]	MAE [°C]	R^2 [-]	CV-RMSE [%]
Physics (RC grey-box)	0.269	0.126	0.987	1.384
Pure ML	0.377	0.153	0.974	1.934
Hybrid (RC + residual ML)	0.343	0.106	0.978	1.760

The Hybrid model estimates the residual between the measured next-step temperature and the discrete-time RC grey-box model prediction and applies an additive correction.

The input variables used in the reported temperature benchmarks were:

- Indoor air temperature $T_{in}(t)$;
- HVAC heating actuator state $u_{hvac}(t)$;
- Thermostat setpoint $T_{set}(t)$;
- CO_2 concentration;
- VOC index;
- Relative humidity;
- Window state (binary);
- Hour of day;
- Day of week;
- Weekend indicator.

For the hybrid residual model, the RC grey-box prediction was additionally included as an input feature.

Due to simplified assumptions inherent to the reduced-order physical model and the presence of unmodeled influences—such as occupancy-driven internal heat gains, manual ventilation actions, and other stochastic disturbances—systematic discrepancies between measured and simulated indoor air temperatures are expected [5].

For consistency with the one-step-ahead RC formulation introduced in Equation (2), the residual is computed for the next-step prediction target as:

$$\varepsilon(t+1) = T_{\text{in}}(t+1) - T_{\text{phys}}(t+1 \mid t) \tag{3}$$

where $T_{\text{phys}}(t+1 \mid t)$ denotes the indoor temperature predicted by the RC grey-box model at time step t+1 using information available at time t.

Residual learning constitutes a common strategy in hybrid grey-box and machine learning frameworks, whereby data-driven models are trained to approximate this next-step residual component rather than the full system dynamics. This approach enables the correction of systematic model errors associated with unmodeled effects, while preserving the structural physical consistency of the underlying thermal formulation [22].

Tree-based gradient boosting was selected for the reported benchmarks due to its favourable balance between predictive performance and robustness for the available dataset [7,18,24].

For supplementary aggregated visual analyses, a histogram-based gradient boosting variant was employed for computational efficiency and robustness to missing values, without affecting the main quantitative benchmarks.

The boosting models used the following fixed hyperparameters in the reported benchmarks: maximum tree depth of 3, learning rate of 0.05, and 400 boosting iterations. The same configuration was applied consistently across rooms to ensure comparability.

The Pure ML baseline uses the complete set of KNX-derived operational variables listed above, without imposing any physical structure. Outdoor temperature is only used within the RC grey-box component and is not included as an explicit feature in the KNX-only Pure ML benchmark.

All temperature benchmarks reported in Table 1 and Supplementary Table S1 use the HistGradientBoostingRegressor with fixed hyperparameters. The ExtraTreesRegressor is employed exclusively for the exploratory CO_2 prediction experiment presented in Section 3.8. Any reference to robustness to missing values in the supplementary visualizations refers to the histogram-based gradient boosting implementation used for temperature modelling.

2.6. Hybrid Prediction Formulation

The final hybrid temperature prediction is obtained by augmenting the RC grey-box model output with a machine learning-estimated residual term, according to:

$$T_{\mathrm{hyb}}(t) = T_{\mathrm{phys}}(t) + f_{\mathrm{ML}}(\mathrm{x}_t) \quad (4)$$

where $T_{\mathrm{hyb}}(t)$ denotes the hybrid predicted indoor air temperature at time t, $T_{\mathrm{phys}}(t)$ is the temperature predicted by the RC grey-box model and $f_{\mathrm{ML}}(\mathrm{x}_t)$ represents the residual correction estimated by the machine learning model. The input vector includes the KNX-derived variables listed in Section 2.5 (temperature, actuator state, setpoint, IEQ indicators, window state, and time features). For the residual model, the RC grey-box prediction is additionally provided as an input feature.

This additive hybrid formulation is widely adopted in hybrid learning frameworks, as it preserves the dominant system dynamics governed by physical laws while enabling data-driven adaptation to occupancy-related effects and other unmodeled disturbances [8].

2.7. Model Training and Evaluation

Model training followed a strictly time-ordered data split. For each monitored room, the October–December dataset was arranged chronologically, with the first 70% of the samples used for model identification and fitting, and the remaining 30% reserved as an unseen test segment. This approach preserves temporal dependencies and avoids information leakage, as recommended for time-series modelling applications [28].

The discrete-time RC grey-box model coefficients were estimated exclusively on the training segment using ordinary least squares. The Pure ML and Hybrid residual models were also fitted only on the training portion of the data. Model performance was evaluated using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2), which are widely adopted metrics for predictive performance assessment in building energy studies [29]. All performance indicators were computed exclusively on the held-out test segment.

No random shuffling or cross-validation was applied in the temperature benchmarking workflow, ensuring that evaluation reflects genuine forward prediction under realistic operating conditions.

All data processing and model implementation were carried out in Python using NumPy, Pandas, and scikit-learn within a Google Colab runtime environment. Due to the

reduced model complexity and moderate dataset size, all room-level models were trained and evaluated within a single interactive session. Training and evaluation were completed within typical Colab execution times for all rooms.

2.8. Model Scope and Assumptions

This study focuses on the prediction of indoor air temperature and the assessment of indoor environmental quality indicators in an educational building operating under simple thermostatic heating control.

Indoor air temperature dynamics are represented using a reduced-order physics-based thermal model, while unmodeled effects—such as occupancy-driven internal gains, uncontrolled infiltration, and other disturbances—are implicitly captured through a data-driven residual learning component [5,14,22].

Indoor air quality indicators (e.g., CO_2 concentration) are analyzed as occupancy-related variables and are not used for ventilation control or airflow estimation.

The proposed framework is intended for prediction and performance assessment rather than real-time HVAC control, and it is designed to operate under limited sensing and control infrastructure, as commonly encountered in existing educational buildings [4,9,10].

3. Results and Discussion

3.1. Overview of Model Performance Under Real Operating Conditions

The performance of the proposed hybrid grey-box and machine learning model was evaluated using real operational data collected from October to December during the heating season. The analyzed period corresponds to active teaching weeks, characterized by highly variable occupancy patterns and intermittent HVAC operation governed by simple thermostatic control.

To provide proper context for the reported error magnitudes, performance metrics were computed over the pooled test dataset (n = 62,456 samples across all monitored rooms). The mean indoor temperature during the test period was 19.47 °C. The indoor temperature ranged between 17.2 °C and 22.8 °C during the test period, with a standard deviation of 0.94 °C. The pooled RMSE values were 0.27 °C for the RC grey-box model, 0.38 °C for the Pure ML model, and 0.34 °C for the hybrid model, corresponding to CV-RMSE values between 1.38% and 1.93%.

The higher RMSE observed in Room 01 (1.04 °C) reflects atypical operational patterns during the monitoring period and contributes significantly to the pooled error metric.

The higher pooled RMSE of the hybrid model is primarily driven by a limited number of larger residual deviations observed in Room 01, whereas median-based and MAE metrics indicate improved typical predictive accuracy across rooms. This highlights the sensitivity of RMSE to rare but higher-amplitude deviations under decentralized thermostatic operation and reinforces the complementary interpretation of RMSE and MAE in this context.

These values are lower than those typically reported in multi-step forecasting or mechanically ventilated scenarios; however, they are consistent with the one-step-ahead prediction horizon and the inertia-dominated thermostatic operation of the monitored spaces.

To provide a comprehensive assessment, the hybrid grey-box/ML model was benchmarked against:

1. An identified reduced-order RC grey-box model (baseline), and
2. A purely data-driven machine learning model trained directly on measured inputs and outputs.

The purely data-driven baseline was configured using the same measured inputs available to the hybrid grey-box/ML framework, but without any explicit physical structure, to

reflect realistic data-driven deployment scenarios under limited sensing and decentralized building automation.

Model accuracy was quantified using the root mean square error, mean absolute error, and the coefficient of determination (R^2), computed on an unseen test dataset representing 30% of the monitored period.

Figure 1 illustrates a representative one-week comparison between the measured indoor air temperature and the corresponding RC grey-box and hybrid grey-box/residual ML model predictions for Room 301 under real operating conditions. The model predictions correspond to one-step-ahead forecasts and are shifted by one sampling interval for visual alignment with the measured temperature. The upper panel shows the absolute indoor temperature evolution, while the lower panel presents the temperature deviation relative to the thermostat setpoint ($\Delta T = T_{\text{in}} - T_{\text{set}}$).

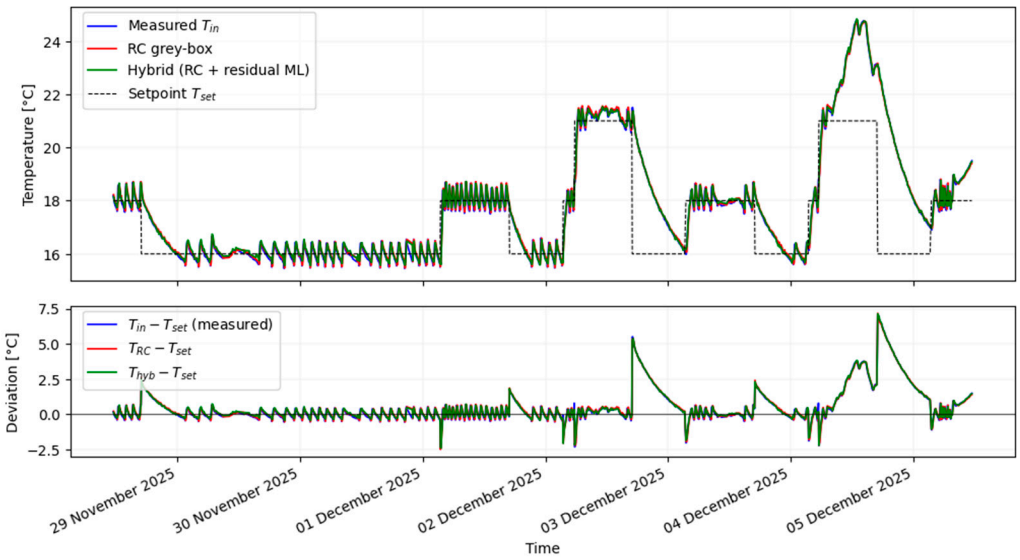


Figure 1. Comparison between measured indoor temperature and model predictions (RC grey-box and hybrid grey-box/residual ML) for Room 301 during a representative test week (held-out test segment).

The time-series analysis indicates that the RC grey-box model successfully captures the dominant thermal dynamics and the overall indoor temperature trend driven by HVAC operation and building thermal inertia. The limited absolute temperature variability during the monitored heating period (typically within a narrow band around the thermostat setpoint) partially explains the high R^2 values observed. Nevertheless, systematic deviations occur during periods of high occupancy and abrupt internal heat gains, which are not explicitly represented in the simplified physical formulation.

In contrast, the hybrid grey-box/ML model reduces typical absolute deviations (MAE), while RMSE remains sensitive to occasional larger errors in a small number of rooms (notably Room 01). The improvement is particularly pronounced during lecture hours, when occupancy-driven disturbances dominate the indoor thermal behaviour, highlighting the added value of combining physics-based structure with data-driven residual learning.

While Figure 1 provides insight into model behaviour for a representative lecture room, a broader quantitative evaluation across all monitored spaces is required to assess robustness and generalization under heterogeneous operating conditions.

3.2. Quantitative Performance Comparison

To complement the qualitative time-series analysis presented in Section 3.1, a quantitative performance evaluation was conducted across all monitored lecture rooms. The

objective of this analysis is to assess the generalization capability of the proposed hybrid grey-box model under diverse operational conditions, rather than focusing on isolated representative cases.

Model performance was evaluated using three widely adopted statistical indicators: the root mean square error, the mean absolute error, and the coefficient of determination (R^2). These metrics were computed on an unseen test dataset corresponding to 30% of the available measurements for each room, ensuring a fair and consistent comparison across models.

The quantitative evaluation focuses on indoor temperature dynamics driven by HVAC operation and internal disturbances. Outdoor temperature was incorporated within the RC grey-box formulation as a boundary condition. It was not directly included as an explicit feature in the purely data-driven ML baseline, to maintain consistency with the available KNX-derived operational variables.

Table 1 summarizes the aggregated performance metrics obtained for the physics-based RC model, the purely data-driven machine learning model, and the proposed hybrid grey-box approach. Results are reported as pooled performance metrics computed over the aggregated test dataset across all monitored rooms, providing a global assessment of predictive capability under heterogeneous operating conditions.

While the hybrid model reduces typical absolute errors (lower MAE), the pooled RMSE remains sensitive to occasional larger deviations in a small number of rooms (Table S1), which motivates reporting both RMSE and MAE.

Detailed per-room performance metrics are provided in Table S1 (Supplementary Materials), allowing a transparent assessment of inter-room variability and room-specific behaviour.

The Pure ML baseline shows higher inter-room variability, with a small number of rooms exhibiting substantially larger errors (Table S1), likely due to sparse data coverage or atypical operating patterns.

Notably, certain rooms exhibit markedly different behaviour across modelling approaches. For instance, in Room 324, the purely data-driven model shows substantially higher prediction error compared to both the RC and hybrid formulations. This discrepancy is likely associated with irregular sensing patterns or sparse effective excitation in that zone, conditions under which the physically structured hybrid model remains comparatively stable.

Overall, the physics-based RC model provides a stable baseline performance, achieving reasonable accuracy across all rooms due to its explicit representation of dominant thermal dynamics and HVAC control effects. However, its predictive capability is consistently limited by unmodeled internal heat gains and occupancy-related disturbances, resulting in higher RMSE and MAE values, particularly during periods of intensive use.

The purely data-driven model achieves improved short-horizon accuracy compared to the physics-based baseline in several rooms; nevertheless, its performance exhibits higher variability and sensitivity to data sparsity and operational changes. This behaviour indicates reduced robustness when extrapolating beyond previously observed patterns, especially under irregular occupancy schedules.

In contrast, the hybrid grey-box model provides modest but consistent improvements over the grey-box baseline while maintaining stronger physical coherence. Although the analyzed heating-season dataset reflects inertia-dominated indoor temperature dynamics under thermostatic operation, the hybrid formulation reduces inter-room variability and improves stability relative to the purely data-driven approach.

On average, the residual-based hybrid framework achieves a consistent reduction in MAE and improved error stability across rooms, while RMSE remains influenced by a lim-

ited number of larger deviations. The standard deviation of the error metrics across rooms is also reduced, highlighting enhanced robustness under heterogeneous operating conditions.

These results suggest that embedding physical structure within the learning process enables the model to maintain consistency with first-principles thermal behaviour while compensating for unmodeled dynamics through residual learning. Consequently, the proposed hybrid approach provides a structured and scalable framework for indoor temperature prediction in real educational buildings, supporting future control-oriented and diagnostic applications.

3.3. Residual Learning Behaviour and Interpretability

To further analyze the contribution of the machine learning component, the distribution of prediction errors (residuals) across all monitored rooms was examined.

Figure S1 (Supplementary Materials) presents boxplots of RMSE and MAE for the physics-based and hybrid grey-box models, highlighting both error magnitude and dispersion.

The residuals associated with the physics-based model exhibit a clear bias and increased variance during occupied periods, indicating systematic underestimation of internal heat gains. In contrast, the residuals of the hybrid grey-box framework are centred around zero and show reduced dispersion, demonstrating that the data-driven component effectively compensates for unmodeled disturbances.

Importantly, while occupancy patterns inferred from CO_2 dynamics affect indoor environmental quality indicators, correlation analysis performed on the pooled dataset did not reveal a statistically significant linear relationship between temperature prediction residuals and CO_2 concentration. This indicates that short-term residual variability cannot be directly attributed to occupancy-driven ventilation effects under the investigated operating conditions.

This behaviour highlights one of the key advantages of residual-based hybrid modelling: physical interpretability is preserved, while data-driven correction is applied only where necessary, avoiding the risk of overfitting or physically inconsistent predictions.

Detailed performance metrics (RMSE, MAE, and R^2) for each room are presented in Supplementary Table S1, highlighting room-level differences and confirming the robustness of the hybrid model across spatially diverse conditions.

To examine whether occupancy-related disturbances contribute to the residual structure of the grey-box model, a correlation analysis was performed between model residuals and measured CO_2 concentration across the pooled test dataset (n = 62,456 samples). The Pearson correlation coefficient between the hybrid residual and CO_2 was r = 0.003 (p = 0.44), while for the RC grey-box residual it was r = 0.007 (p = 0.072). These values indicate no statistically significant linear relationship. This suggests that, at the adopted 5 min sampling resolution and under thermostatic control conditions, short-term temperature prediction residuals are not directly explained by CO_2 variability. Consequently, CO_2 and VOC indicators are interpreted as complementary IEQ descriptors rather than explanatory variables of the residual-based modelling framework.

3.4. Impact of Simplified HVAC Control and Limited Sensing

The analyzed building operates under simple thermostatic on–off control, without centralized supervision or mechanical ventilation. Under such conditions, indoor thermal dynamics are strongly influenced by uncontrolled factors, including occupant behaviour and manual window operation, which are not directly measured.

Despite these limitations, the hybrid grey-box model demonstrated robust predictive performance under heterogeneous occupancy patterns, indicating that meaningful

improvements in IEQ prediction can be achieved even in buildings with limited sensing and control infrastructure.

In addition, the aggregated relationship between predicted and measured indoor temperature was analyzed across all rooms.

While aggregated distributions offer a global performance overview, they also partially mask room-level improvements achieved by the hybrid grey-box framework. Figure S2 (Supplementary Materials) presents a two-dimensional kernel density comparison between predicted and measured indoor temperatures for both the physics-based and hybrid models.

The hybrid grey-box framework shows a tighter concentration along the 1:1 line, consistent with the aggregated metrics; however, the improvement is modest when evaluated on fully aggregated distributions.

These findings are particularly relevant for educational buildings, where advanced HVAC control strategies are often infeasible due to infrastructural constraints. In such contexts, hybrid models can serve as effective tools for performance assessment and decision support without requiring extensive system upgrades.

3.5. Discussion in Relation to the Existing Literature

The obtained results align with and extend recent findings in the field of physics-based machine learning for building applications. While previous studies have demonstrated the potential of hybrid models in simulated environments or under centralized control architectures [24,25], the present work demonstrates that comparable robustness in predictive accuracy and generalization can be achieved in decentralized, sensor-limited environments.

Compared to grey-box-only approaches [6,14–16], the proposed framework achieves superior predictive accuracy under highly variable occupancy conditions. At the same time, it avoids the black-box limitations associated with purely data-driven machine learning models [7,18], ensuring physical plausibility and improved generalization.

Unlike advanced model predictive control schemes, which require centralized control architectures, real-time optimization routines, and detailed system identification, the proposed hybrid grey-box model is better suited for buildings with decentralized, rule-based control and limited sensor infrastructure. In this study, the KNX-based automation system provided only coarse-grained actuator and temperature data, making MPC impractical. The residual learning component in the ML model allows for adaptive correction of model discrepancies without requiring full observability or controllability, thus offering a viable alternative in retrofitted or minimally instrumented educational environments.

These findings are consistent with recent advances in hybrid modelling for real-world building applications [24,25], and further support the feasibility of deploying such methods under constrained sensing and control conditions.

A key contribution of this study lies in demonstrating that hybrid grey-box frameworks can deliver statistically and operationally meaningful improvements even under simplified HVAC control and incomplete sensing—a scenario representative of a large fraction of existing educational buildings, where advanced HVAC control strategies are infeasible due to infrastructural and budgetary constraints.

3.6. Indoor Air Quality Assessment Using CO_2 Indicators

CO_2 and VOC measurements were not available for all rooms included in the temperature analysis. Consequently, the IEQ-related tables include only rooms for which complete air quality data were recorded.

Carbon dioxide concentration is widely used as a proxy indicator for occupancy-related indoor air quality and ventilation adequacy in educational buildings. Numerous studies have demonstrated that elevated CO_2 levels in classrooms are associated with

insufficient air exchange and can negatively affect occupants perceived air quality and cognitive performance in educational environments [17,30].

Table 2 summarizes descriptive CO_2 indicators for each monitored lecture room, including the median, 95th percentile (P95), maximum concentration, and the total number of samples. Results are reported for the full monitoring period using 5 min aggregated data.

Table 2. CO_2 indicators for each monitored lecture room, median, 95th percentile (P95), maximum concentration and the total number of samples.

Room	CO_2 Median [ppm]	CO_2 P95 [ppm]	CO_2 Max [ppm]	Time % GT1000	Time % GT1400	Time % GT2000	N Samples
Overall	625	1897	7224	22	11	4	70,682
1	654	1742	3494	16	7	3	5074
3	1048	1677	2527	58	9	2	301
13	1018	2083	3092	52	17	6	1384
15	687	1999	3405	27	15	5	2351
16	532	1233	2569	14	3	0	4126
101	613	1018	1583	6	1	0	3036
107	724	2844	5240	36	23	13	4367
109	570	1645	3333	15	8	1	3819
110	646	1722	3868	20	9	3	3980
201	1113	4435	7224	54	40	26	1571
202	670	1971	3889	25	13	5	3356
203	742	2786	7071	32	22	11	3664
204	799	1834	4004	35	15	3	4486
205	820	1970	3935	33	15	5	4274
206	685	1998	2555	21	10	5	343
207	660	1463	3464	14	5	2	1330
208	605	2264	4457	16	11	6	2747
209	790	1863	3433	34	15	4	3657
210	688	1612	2426	22	10	1	656
301	492	1705	3556	11	7	3	6530
302	518	1098	1807	7	2	0	4500
324	497	1211	2627	10	3	0	5130

While median CO_2 levels remain close to outdoor background values for most rooms, elevated P95 and maximum concentrations are observed in several lecture rooms. Across the monitored spaces, approximately one fifth of the occupied time is characterized by CO_2 concentrations exceeding commonly referenced comfort thresholds, indicating recurrent periods of insufficient air exchange during teaching activities.

Figure 2 illustrates the relative severity of CO_2 exceedance across lecture rooms by ranking the percentage of time with concentrations above 1000 ppm. The figure highlights a strong heterogeneity between rooms, reflecting differences in occupancy density, usage patterns, and the effectiveness of natural ventilation strategies.

While elevated CO_2 concentrations indicate increased occupancy and potential internal gains, correlation analysis (Section 3.3) did not demonstrate a statistically significant linear relationship between CO_2 levels and temperature prediction residuals. This suggests that ventilation-related thermal effects are not directly identifiable through CO_2 variability at the adopted temporal resolution.

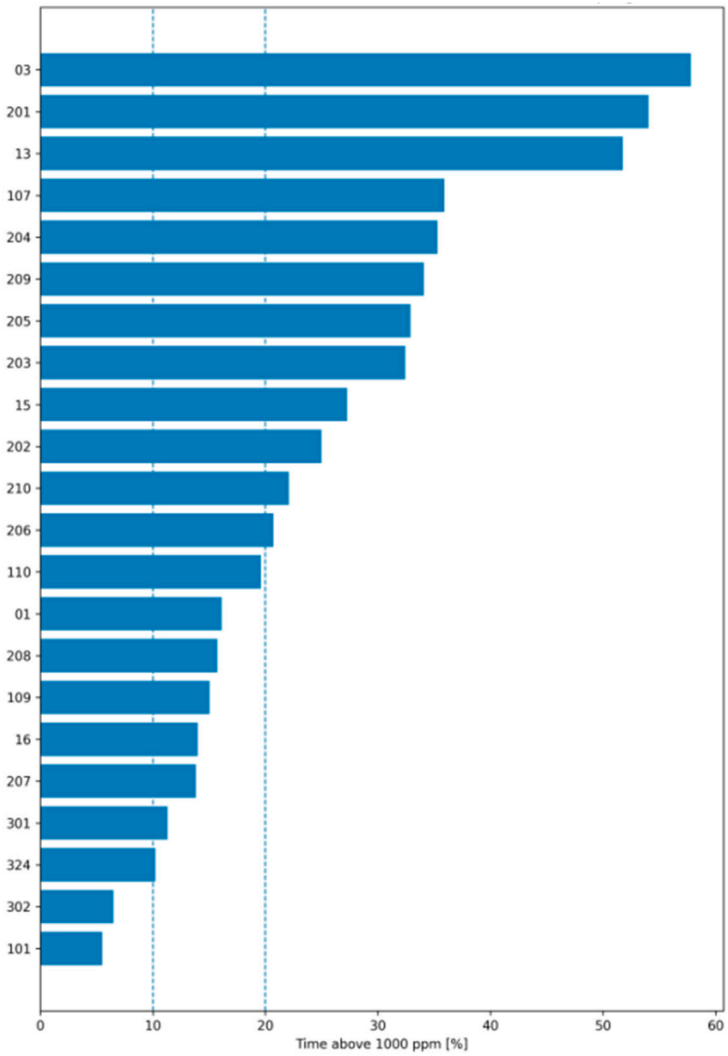


Figure 2. CO_2 exceedance across lecture rooms by ranking the percentage of time with concentrations above 1000 ppm (5 min sampling).

3.7. Complementary IEQ Analysis Using VOC Index

Volatile organic compound measurements are increasingly employed as complementary indicators of indoor air quality, particularly in naturally ventilated educational buildings where pollutant sources are strongly linked to occupancy and indoor activities. Recent reviews highlight the relevance of VOC-related metrics for capturing variations in indoor pollutant loads and supporting comparative IEQ assessments across rooms and usage patterns, despite sensor-specific limitations [31].

Unlike CO_2, VOC index values are sensor-specific and do not directly correspond to absolute concentration limits; therefore, the analysis focuses on descriptive statistics and distributional patterns rather than threshold exceedances.

Table 3 reports the median, P95, and maximum VOC index values for each lecture room. The results indicate recurring VOC peaks during occupied periods, suggesting increased indoor pollutant loads associated with human presence, teaching activities, and material emissions.

Figure 3 presents boxplots of the VOC index distribution across lecture rooms. The interquartile ranges and upper whiskers reveal substantial variability between spaces, while the overall distribution confirms the episodic nature of elevated VOC levels in naturally ventilated educational environments.

Table 3. Median, P95, and maximum VOC index values for each lecture room.

Room	VOC Median Index	VOC P95 Index	VOC Max Index	N Samples
Overall	1240	3008	4398	56,021
1	1160	2519	4352	4482
3	1248	2404	3686	250
13	1280	3533	4360	1007
15	1288	3615	4360	1986
16	1296	2663	4352	3420
101	1208	1696	4383	1503
107	1240	3528	4398	3262
109	1224	2510	4326	2745
110	1248	3169	4398	3034
201	1327	3692	4278	1077
202	1184	2199	4326	2709
203	1224	2648	4375	2670
204	1320	2662	4390	3540
205	1296	2655	4352	3602
206	1216	2287	2632	221
207	1208	2528	4342	1061
208	1200	3283	4352	2198
209	1264	2583	4111	2538
210	1200	2354	3983	548
301	1264	3607	4375	6113
302	1224	2655	4398	3577
324	1264	3528	4398	4478

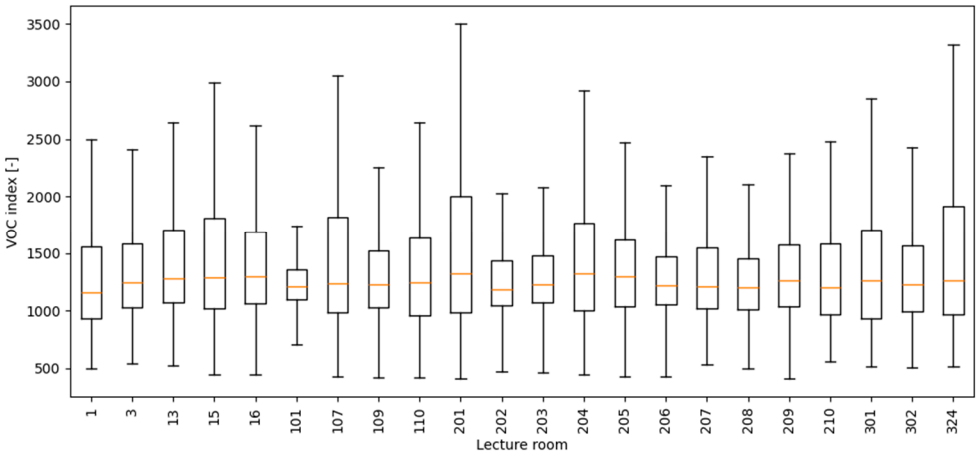


Figure 3. VOC index distribution across monitored lecture rooms (5 min sampling).

3.8. Short-Term CO_2 Prediction Using Machine Learning

This analysis is presented as a complementary extension to the temperature-focused hybrid grey-box framework, aiming to illustrate the broader IEQ assessment potential of KNX-derived data rather than to introduce a separate predictive modelling contribution. Beyond descriptive indoor air quality assessment, short-term prediction of CO_2 concentration can provide additional insights into occupancy-driven dynamics and support proactive building operation strategies. In educational buildings characterized by highly variable occupancy patterns and limited ventilation control, short-horizon CO_2 forecasting represents a valuable analytical extension of IEQ monitoring frameworks [19,30].

Using the same KNX-derived time-series dataset employed for the temperature analysis, a data-driven machine learning approach was applied to predict indoor CO_2 concentration at a 15 min forecasting horizon. The model was trained independently for each monitored room and evaluated using forward-chaining time-series cross-validation (Time Series Split) to preserve temporal dependencies. A persistence model, assuming constant CO_2 concentration over the prediction horizon, was used as a baseline reference.

Table 4 summarizes the aggregated prediction performance across all monitored rooms, reporting root mean square error and mean absolute error for the machine learning model and the persistence baseline. Results are expressed as mean ± standard deviation to reflect inter-room variability.

Table 4. Prediction performance across all monitored rooms.

Model	Horizon [Min]	RMSE (ML) [ppm]	MAE (ML) [ppm]	RMSE (Persistence) [ppm]	MAE (Persistence) [ppm]	N Rooms
Extra Trees	15	565.8 ± 212.6	311.3 ± 115.4	185.7 ± 85.7	84.2 ± 39.4	20

The results indicate that short-term CO_2 prediction under limited sensing conditions is strongly inertia-dominated, with the persistence baseline outperforming the data-driven model at the 15 min horizon. This finding suggests that, in the absence of explicit ventilation rate or occupancy measurements, additional model complexity does not necessarily translate into improved short-term forecasting accuracy.

Data-driven CO_2 prediction approaches have been widely explored in educational buildings [19,31]; however, many reported improvements rely on richer sensing infrastructures or additional occupancy-related inputs. In the present decentralized KNX setting, the persistence model remains a competitive and robust short-term reference.

When combined with the temperature-focused hybrid framework, this exploratory CO_2 analysis broadens the IEQ monitoring perspective, while highlighting that strong baselines (e.g., persistence) may remain difficult to outperform under limited sensing.

3.9. Practical Implications and Limitations

From a practical perspective, the combined analysis of temperature prediction, CO_2 dynamics, and VOC index variability demonstrates that meaningful IEQ assessment can be achieved using existing building automation data, even in the absence of mechanical ventilation and advanced sensing infrastructure. CO_2 and VOC indicators provide complementary information on occupancy patterns and indoor air quality variability. However, correlation analysis performed in Section 3.3 did not reveal a statistically significant linear relationship between CO_2 concentration and temperature prediction residuals, indicating that short-term thermal residuals cannot be directly attributed to occupancy-driven ventilation effects under the investigated conditions.

The proposed hybrid modelling approach can be deployed as a non-intrusive performance assessment tool, supporting identification of rooms with recurrent comfort deviations, evaluation of operational strategies, and prioritization of retrofit or ventilation improvement measures.

Nevertheless, several limitations must be acknowledged. First, ventilation effects are not explicitly modelled and are instead implicitly captured through the residual learning component, an approach commonly adopted in hybrid modelling when detailed airflow data are unavailable [22]. While this enables practical implementation under limited sensing conditions, it restricts the direct quantification of airflow rates and their physical interpretation.

It is also important to note that the achieved prediction errors are on the same order of magnitude as the nominal sensor accuracy (± 0.3 °C for indoor temperature measurements). This suggests that further reduction in absolute prediction error may be partially constrained by measurement uncertainty rather than model structure alone.

The analyzed dataset reflects a heating season segment characterized by relatively stable thermostatic operation and limited high-frequency disturbances. Consequently, short-term prediction remains strongly influenced by thermal inertia, which should be considered when interpreting short-horizon predictive performance results.

The reported performance corresponds to one-step-ahead prediction under heating-season conditions. Extension to longer prediction horizons or cooling-dominated periods may require additional disturbance modelling or regularization strategies.

Furthermore, although ventilation-related effects are implicitly represented within the residual term of the grey-box model, the absence of direct airflow measurements prevents their quantitative isolation. The lack of statistically significant correlation between model residuals and CO_2 concentration suggests that residual variability cannot be straightforwardly interpreted as occupancy-driven ventilation dynamics. This highlights the methodological limitation of relying solely on decentralized building automation data without dedicated airflow sensing.

The use of hourly outdoor temperature data may limit the representation of short-term outdoor fluctuations. However, given the first-order thermal dynamics represented by the RC grey-box model and the adopted time resolution, such effects are expected to have limited impact on the reported predictive performance.

The absence of explicit ventilation rate measurements and the sensor-specific nature of the VOC index limit the possibility of quantitative exposure assessment. Consequently, the proposed framework is intended for performance assessment and decision support rather than regulatory compliance evaluation.

Despite these limitations, the results indicate that the proposed residual-based hybrid framework maintains robust predictive capability for the studied KNX-operated educational building.

In addition, the demonstrated feasibility of short-term CO_2 prediction highlights the potential for early detection of abnormal occupancy or ventilation patterns, supporting future predictive maintenance and diagnostic applications in KNX-based educational buildings.

Future work will focus on extending the model to incorporate real-time weather measurements, explicit airflow estimation, and integration with supervisory control strategies.

The present study was conducted under temperate continental climatic conditions characterized by moderate heating-season variability. In significantly different climates (e.g., hot-arid or subtropical regions such as parts of South Africa), building envelope dynamics, solar gains, and ventilation patterns may differ substantially. While the residual-based hybrid framework itself is not climate-specific, recalibration of the grey-box parameters and potential inclusion of additional boundary variables (e.g., solar radiation) would be required for reliable transferability.

4. Conclusions

4.1. Summary and Concluding Remarks

This study evaluated a residual-based hybrid grey-box and machine learning framework for indoor temperature prediction in an educational building operating under simple thermostatic heating control and limited sensing infrastructure. The model was assessed using real KNX automation data collected during the heating season, reflecting heterogeneous occupancy patterns and decentralized operation.

The identified discrete-time RC grey-box model provided a stable and physically interpretable baseline across all monitored rooms, capturing the dominant inertia-driven thermal dynamics. The addition of a data-driven residual component resulted in modest but consistent improvements in predictive performance over the grey-box baseline. On the pooled test dataset (n = 62,456 samples), RMSE values ranged between 0.27 °C and 0.38 °C (CV-RMSE: 1.38–1.93%), indicating incremental rather than transformative gains.

In one room (Room 01), the hybrid model exhibited larger prediction errors compared to both the RC and purely data-driven baselines, indicating that residual correction does not universally guarantee improvement and may amplify errors under unstable sensing or control behaviour.

The main benefit of the hybrid formulation lies not in large absolute error reductions, but in reducing systematic bias and improving robustness across rooms with heterogeneous operating conditions. The hybrid framework maintained physical interpretability while compensating for unmodeled disturbances through residual learning.

Correlation analysis between model residuals and CO_2 concentration did not reveal statistically significant linear relationships. Therefore, temperature prediction residuals cannot be directly attributed to occupancy-derived ventilation effects under the investigated temporal resolution. CO_2 and VOC indicators are thus interpreted as complementary IEQ descriptors rather than explanatory drivers within the temperature modelling framework.

The analyzed heating-season dataset reflects inertia-dominated indoor temperature dynamics under decentralized thermostatic control. Therefore, the reported one-step-ahead accuracy should be interpreted within this operational context rather than as a multi-horizon forecasting benchmark.

The primary contribution of this work is empirical: demonstrating that residual-based hybrid grey-box modelling can be applied consistently across multiple rooms in a decentralized KNX-operated educational building without advanced control architectures or comprehensive sensing. While residual-based hybrid modelling approaches are well established in the literature, their validation under realistic, sensor-limited educational building conditions remains limited.

Overall, the results indicate that hybrid grey-box/ML modelling represents a practical and scalable approach for performance assessment in existing educational buildings, where simplicity, interpretability, and robustness are more critical than marginal gains in predictive accuracy.

4.2. Future Work

While the present study focuses on indoor temperature prediction and model performance assessment, the proposed hybrid framework naturally opens several promising research directions.

The first extension concerns the integration of additional data sources and higher-resolution boundary conditions, enabling the investigation of short-term thermal responses and the interaction between indoor dynamics and external forcing. Such extensions would further enhance model fidelity and broaden applicability across different climatic contexts.

More importantly, the results presented in this work lay the foundation for a dedicated investigation into predictive maintenance and fault detection of KNX-based HVAC systems. The availability of long-term, room-level measurements across multiple indoor environmental variables enables the identification of abnormal operational patterns, such as persistent deviations between measured indoor temperature and thermostat setpoints, delayed system response, or asymmetric heating behaviour across similar rooms.

Building on the predictive capabilities demonstrated in this article, future research will explore the use of hybrid and data-driven models to detect early signs of system

degradation, sensor drift, actuator malfunction, or control inefficiencies. In this context, prediction residuals and setpoint tracking errors can be exploited as diagnostic indicators, enabling a transition from reactive to predictive maintenance strategies in building automation systems.

Such an approach is particularly relevant for KNX-controlled educational buildings, where large numbers of decentralized terminal units operate under simple control logic and manual intervention is often required to address comfort complaints. A predictive maintenance framework informed by hybrid modelling could support facility managers by prioritizing interventions, reducing downtime, and improving overall system reliability without requiring intrusive hardware upgrades.

Finally, future work will investigate the integration of the proposed predictive and diagnostic framework into decision-support tools for facility management, bridging the gap between data-driven building analytics and practical operational workflows. This direction represents a natural continuation of the present study and contributes to the broader objective of enhancing the resilience, efficiency, and sustainability of existing building stocks.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/buildings16050964/s1, Table S1: (predictive performance metrics for individual lecture rooms), Table S2: (basic geometric characteristics of the monitored lecture rooms), Figure S1: (Distribution of room-level predictive performance across monitored lecture rooms), Figure S2: (Aggregated error distribution across all monitored rooms on the pooled test dataset), and supplementary methodological details.

Author Contributions: Conceptualization, C.C.; methodology, C.C.; software, C.C.; validation, C.C. and H.C.A.; formal analysis, C.C.; investigation, C.C.; resources, H.C.A. and A.C.A.; data curation, C.C.; writing—original draft preparation, C.C.; writing—review and editing, C.C., H.C.A. and A.C.A.; visualization, C.C.; supervision, H.C.A.; project administration, C.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding. The article processing charge (APC) was not funded by any external organization.

Data Availability Statement: The data presented in this study originates from a real operational educational building and is not publicly available due to privacy and operational constraints. The datasets can be made available from the corresponding author upon reasonable request.

Acknowledgments: The authors acknowledge the technical and administrative support provided during the KNX monitoring and data acquisition process.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial Neural Network
CO_2	Carbon Dioxide
HVAC	Heating, Ventilation, and Air Conditioning
IEQ	Indoor Environmental Quality
KNX	Building automation and control standard
MAE	Mean Absolute Error
ML	Machine Learning

MPC	Model Predictive Control
RC	Resistance–Capacitance
RMSE	Root Mean Square Error
SVR	Support Vector Regression
VOC	Volatile Organic Compounds

References

1. Pérez-Lombard, L.; Ortiz, J.; Pout, C. A review on buildings energy consumption information. *Energy Build.* **2008**, *40*, 394–398. [CrossRef]
2. International Energy Agency (IEA). *Energy Efficiency 2023*; IEA: Paris, France, 2023; Available online: https://www.iea.org/reports/energy-efficiency-2023 (accessed on 15 January 2026).
3. Fanger, P.O. *Thermal Comfort: Analysis and Applications in Environmental Engineering*; Danish Technical Press: Copenhagen, Denmark, 1970.
4. Wang, S.; Ma, Z. Supervisory and optimal control of building HVAC systems: A review. *HVAC&R Res.* **2008**, *14*, 3–32. [CrossRef]
5. Gouda, M.M.; Danaher, S.; Underwood, C.P. Building thermal model reduction using nonlinear constrained optimization. *Build. Environ.* **2002**, *37*, 1255–1265. [CrossRef]
6. 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. [CrossRef]
7. Fan, C.; Xiao, F.; Zhao, Y. A short-term building cooling load prediction method using deep learning algorithms. *Appl. Energy* **2017**, *195*, 222–233. [CrossRef]
8. Karniadakis, G.E.; Kevrekidis, I.G.; Lu, L.; Perdikaris, P.; Wang, S.; Yang, L. Physics-informed machine learning. *Nat. Rev. Phys.* **2021**, *3*, 422–440. [CrossRef]
9. Rus, T.; Moldovan, R.; Albu, H.; Beu, D. Impact of pandemic safety measures on students' thermal comfort—Case study: Romania. *Buildings* **2023**, *13*, 794. [CrossRef]
10. Rus, T.; Moldovan, R.-P.; Pop, M.I.; Moldovan, A.-M. Assessing the interplay of indoor environmental quality, energy use, and environmental impacts in educational buildings. *Appl. Sci.* **2025**, *15*, 3591. [CrossRef]
11. Moldovan, R.-P.; Rus, T.; Beu, D.; Albu, H.; Domnița, F.; Moldovan, A.-M. From assessment to action: A strategic approach to carbon management for climate neutrality in higher education. *Sustain. Futures* **2025**, *10*, 101354. [CrossRef]
12. Olesen, B.W.; Parsons, K.C. Introduction to thermal comfort standards and to the proposed new version of EN ISO 7730. *Energy Build.* **2002**, *34*, 537–548. [CrossRef]
13. Afram, A.; Janabi-Sharifi, F. Review of modeling methods for HVAC systems. *Appl. Therm. Eng.* **2014**, *67*, 507–519. [CrossRef]
14. Bacher, P.; Madsen, H. Identifying suitable models for the heat dynamics of buildings. *Energy Build.* **2011**, *43*, 1511–1522. [CrossRef]
15. Reynders, G.; Diriken, J.; Saelens, D. Quality of grey-box models and identified parameters as a function of the accuracy of input and observation signals. *Energy Build.* **2014**, *82*, 263–274. [CrossRef]
16. Drgoňa, J.; Picard, D.; Kvasnica, M.; Helsen, L. Approximate model predictive building control via machine learning. *Appl. Energy* **2018**, *218*, 199–216. [CrossRef]
17. Corgnati, S.P.; Ansaldi, R.; Filippi, M. Thermal comfort in Italian classrooms under free-running conditions during mid-seasons: Assessment through objective and subjective approaches. *Build. Environ.* **2009**, *44*, 785–792. [CrossRef]
18. Ben-Nakhi, A.; Mahmoud, M.A. Cooling load prediction for buildings using general regression neural networks. *Energy Convers. Manag.* **2004**, *45*, 2127–2141. [CrossRef]
19. Dai, Z.; Yuan, Y.; Zhu, X.; Zhao, L. A method for predicting indoor CO_2 concentration in university classrooms: An RF–TPE–LSTM approach. *Appl. Sci.* **2024**, *14*, 6188. [CrossRef]
20. Amasyali, K.; El-Gohary, N.M. A review of data-driven building energy consumption prediction studies. *Renew. Sustain. Energy Rev.* **2018**, *81*, 1192–1205. [CrossRef]
21. Afram, A.; Janabi-Sharifi, F. Theory and applications of HVAC control systems—A review of model predictive control (MPC). *Build. Environ.* **2014**, *72*, 343–355. [CrossRef]
22. Ma, Z.; Jiang, G.; Hu, Y.; Chen, J. A review of physics-informed machine learning for building energy modeling. *Appl. Energy* **2025**, *381*, 125169. [CrossRef]
23. Wei, T.; Wang, Y.; Zhu, Q. Deep reinforcement learning for building HVAC control. In Proceedings of the 54th Annual Design Automation Conference (DAC '17), San Francisco, CA, USA, 18–22 June 2017; Volume 22, pp. 1–6. [CrossRef]
24. Gokhale, G.; Claessens, B.; Develder, C. Physics-informed neural networks for control-oriented thermal modeling of buildings. *Appl. Energy* **2022**, *314*, 118852. [CrossRef]
25. Shan, R.; Ning, H.; Xu, Q.; Su, X.; Guo, M.; Jia, X. Physics-informed and explainable graph neural networks for generalizable urban building energy modeling. *Appl. Sci.* **2025**, *15*, 8854. [CrossRef]

26. Privara, S.; Cigler, J.; Váňa, Z.; Oldewurtel, F.; Sagerschnig, C.; Zacekova, E. Building modeling as a crucial part for building predictive control. *Energy Build.* **2013**, *56*, 8–22. [CrossRef]
27. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning*; Springer: New York, NY, USA, 2009. [CrossRef]
28. Bergmeir, C.; Hyndman, R.J.; Benítez, J.M. Bagging exponential smoothing methods using STL decomposition and Box–Cox transformation. *Int. J. Forecast.* **2016**, *32*, 303–312. [CrossRef]
29. ASHRAE. *ASHRAE Guideline 14: Measurement of Energy, Demand, and Water Savings*; ASHRAE: Atlanta, GA, USA, 2014.
30. Wargocki, P.; Porras-Salazar, J.A.; Contreras-Espinoza, S.; Bahnfleth, W. The relationships between classroom air quality and children's performance in school. *Build. Environ.* **2020**, *173*, 106749. [CrossRef]
31. Branco, P.T.B.S.; Sousa, S.I.V.; Dudzińska, M.R.; Ruzgar, D.G.; Mutlu, M.; Panaras, G.; Papadopoulos, G.; Saffell, J.; Scutaru, A.M.; Struck, C.; et al. A review of relevant parameters for assessing indoor air quality in educational facilities. *Environ. Res.* **2024**, *261*, 119713. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.