Physics-Informed vs. Deep Learning: Indoor Temperature Prediction with Different Data Availability

Physics-Informed vs. Deep Learning: Indoor Temperature Prediction with Different Data Availability

Maria Adelaide Loffa*
Politecnico di Torino
Turin, Italy
maria.loffa@polito.it

Enrico Macii
Politecnico di Torino
Turin, Italy
enrico.macii@polito.it

Edoardo Patti
Politecnico di Torino
Turin, Italy
edoardo.patti@polito.it

Lorenzo Bottaccioli
Politecnico di Torino
Turin, Italy
lorenzo.bottaccioli@polito.it

Abstract

Reducing energy consumption in the building sector is crucial for global sustainability. Achieving energy efficiency requires advanced technologies and robust building models that address uncertainties and environmental variations. White-box models, based on physical principles, provide interpretability but require significant expertise and resources. In contrast, black-box models rely on historical data, lacking interpretability and being dataset-dependent. To bridge this gap, Scientific Machine Learning integrates physical insights into machine learning frameworks, ensuring interpretability while maintaining the accuracy and computational efficiency of models like neural networks.

This work incorporates physical knowledge through Ordinary Differential Equations into the neural network framework, developing a physics-informed model to predict indoor air temperature based on external conditions, system thermal powers, and building internal gains. Datasets from four cities with diverse climatic conditions were used, and the model was trained on varying amounts of data, from two weeks to two years. This approach offers a novel exploration of model performance under different data availability and multiple scenarios. A comparative analysis with a Long Short-Term Memory neural network shows that, especially with limited training data, the Physics-Informed Neural Network outperforms the conventional model, with a Mean Absolute Error up to 0.69°C lower. This advantage is due to the incorporation of physics-based constraints, reducing reliance on large datasets. Additionally, the Physics-Informed Neural Network demonstrates stable accuracy across seasonal and uncontrolled dynamics conditions, highlighting its potential for temperature prediction and building control applications.

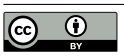
CCS Concepts

• **Computing methodologies** → **Model verification and validation**.

Keywords

physics-informed, neural network, building, modeling, Ordinary Differential Equation

ACM Reference Format:
Maria Adelaide Loffa, Enrico Macii, Edoardo Patti, and Lorenzo Bottaccioli. 2025. Physics-Informed vs. Deep Learning: Indoor Temperature Prediction with Different Data Availability. In *The 16th ACM International Conference on Future and Sustainable Energy Systems (E-ENERGY '25), June 17–20, 2025, Rotterdam, Netherlands.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3679240.3734642



E-ENERGY '25, Rotterdam, Netherlands

ACM ISBN 979-8-4007-1125-1/25/06
https://doi.org/10.1145/3679240.3734642

1 Introduction

Climate change, driven by the accumulation of greenhouse gases (GHGs) like CO_2, leads to global temperature rise, necessitating significant emission reductions. Europe aims to cut emissions by 55% by 2030 and achieve net-zero by 2050 [3]. With energy consumption contributing over three-quarters of global emissions, buildings account for 17.5% [29], making them a crucial target for emission reductions. Energy demand in buildings is projected to rise due to increased appliance use and air conditioning [19]. To mitigate this, energy-efficient buildings with advanced control strategies are essential. Control strategies ensure stable indoor environments while minimizing energy consumption and managing transitions between operational states [16]. Traditional approaches often struggle to adapt to changing external conditions and varying internal loads, reducing overall system efficiency. Advanced control strategies provide a more adaptive solution by considering external factors, occupancy patterns, and energy pricing to optimize energy use dynamically [35]. However, their effectiveness can be undermined by data scarcity, which is common in both older buildings lacking sensor installations and new constructions without historical data. Developing accurate building models is essential to predict performance, simulate scenarios, and implement reliable control solutions. For this purpose, the development of an accurate and reliable building model becomes fundamental [8].

Building models are classified as white-box, black-box, and grey-box. White-box models rely on the system's physics, described by governing ODEs and PDEs, providing interpretable solutions and generalizable results. While useful with limited data, they require significant time, expertise, and computational resources [30], as seen with EnergyPlus, a building energy simulation tool [11]. In contrast, black-box models are data-driven, relying on large datasets for predictions without expert knowledge, so they lack transparency and are limited in their interpretability. These models include machine learning algorithms like neural networks, which cannot explain how conclusions are drawn.

It is from these considerations that interpretable machine learning, also known as Scientific Machine Learning (SciML) [2], which falls under the grey-box category, was born with the aim of leveraging the physical information describing the system, ensuring interpretability of the model, without sacrificing the accuracy and

computational efficiency of traditional machine learning. This generalization and consistency with physical constraints problem particularly concerns neural network models, for which the quality and uniformity of the training and testing dataset are crucial. Recent research has focused on incorporating physical laws into neural networks, particularly through Physics-Informed Neural Networks (PINNs), which improve both model reliability and interpretability [14].

In an effort to advance energy-efficient building practices and tackle the challenges posed by the building sector's significant impact on global energy consumption, this work undertakes a comprehensive exploration of PiNNs applied to building energy consumption. The primary objective is to develop a neural network that integrates physical knowledge derived from the differential equations governing air temperature dynamics. This approach enables the prediction of air temperature based on a range of features, including external temperature, solar radiation, heating/cooling thermal power, and internal gains. The performance of the PiNN model is then compared to that of a Long Short-Term Memory (LSTM) model, which is recognized as one of the best architectures for time-series forecasting. This comparison aims to understand how the two models handle data scarcity. Together with this, the objective is also to evaluate the models' performances under varying climatic conditions, thus ensuring its adaptability across different environmental contexts. To achieve this, custom synthetic datasets are generated with varying climatic conditions and data percentages, including different durations (i.e, 2 weeks, 1 month, 3 months, 1 year, 2 years) and locations representing four distinct cities in different countries: Turin for Italy, Munich for Germany, Copenhagen for Denmark and Madrid for Spain. Although the datasets used in this study are synthetic and realistic, the underlying methodology is designed to be easily transferable to real-world data, making it applicable for broader implementation in actual building systems.

The rest of the paper is structured as follows. Section 2 provides a review of the literature on the application of PiNNs in the building sector. Section 3 outlines the developed methodology, while Section 4 presents and discusses the experimental results. Finally, Section 5 concludes with a summary of key findings and remarks.

2 Related works

PiNNs are widely used in complex physics fields like fluid and solid mechanics. Zhang et al. [34] developed a PiNN framework for learning relationships in incompressible fluid flow around a cylinder, applying physical regularization to enhance learning. In solid mechanics, Zhang et al. [33] created a PiNN model to identify soft tissue properties, using two neural networks to solve the forward problem and estimate material parameters. A typical way to model building thermal behavior was through grey-box model, such as RC networks. Hossain et al. [18] used Bayesian neural networks to identify these models from sensor data and provided pre-trained models to reduce training time, while in [20] the authors proposed an automated, scalable approach and a new indicator to assess model quality. Numerous examples of PiNN application in the building context are available in literature, each each addressing algorithmic challenges differently. In numerous studies, control-oriented models have been introduced. Chen et al. [1] proposed PhysCon, integrating domain knowledge via a physics-based loss term to predict lumped building thermal mass and plan a Demand Response strategy. Training is regulated by a time-lag and regularization term. Gokhale et al. [17] adopted the same loss function modification and introduced PhysNet, a two-module architecture encoding latent and observable states to predict outputs while embedding system physics. Later, it was enhanced for multi-time-step room temperature and thermal power predictions [27], enabling integration with a Monte Carlo Tree Search control algorithm.

PiNNs have been shown to be capable of handling both forward and inverse problems [7], sometimes jointly by exploiting the same optimization problem [4]. In [22], Luo et al. present a PiNN model designed to estimate the parameters of the Ordinary Differential Equation (ODE) governing the air conditioning load in a building. The training process is divided into supervised and unsupervised learning phases, each guided by a distinct term in the loss function: one related to data and the other to physics. [31] and [24] both demonstrate fully unsupervised learning approaches, as their PiNN models are trained without labeled data. In [31] the physics-informed model is trained offline to capture data hall dynamics and acts as a state estimator in a reinforcement learning control strategy. Likewise in [24] the authors develop HVAC system thermal model to be used in a Model Predictive Control. Notably, the physics-informed loss functions in both studies exclude any data term, consisting only of two components: one for the initial condition and another for the ODE residual, highlighting their purely unsupervised nature. Most of the PiNN models have a Multilayer Perceptron (MLP) model as an architecture, but it is not rare to find models with other structures. In [4] a Long Short-Term Memory (LSTM) neural network is trained in order to predict room temperature, while incorporating a physics-constrained loss function. This approach respects the governing ODEs of temperature variation and integrates the control problem optimization directly into the network's training process. Recurrent Neural Networks are also employed in [26], where the authors introduce a PiNN to model the thermal dynamics of an electric water heater, incorporating constraints on the direction of temperature gradients and permissible values. Nagarathinam et al. [25] combined a graph neural network with physics-informed learning, where the dynamics of the graph's interconnected nodes are governed by building behavioral equations, while the loss function incorporates both the residuals of the ODE and a term accounting for the initial conditions.

Physical knowledge can also be incorporated into the model architecture. Di Natale et al. [8] incorporated physical knowledge into model architecture by integrating a linear physical module within a Neural Network (PCNN), ensuring consistency without modifying the loss function. The model is thus defined physically consistent because it does not learn all the information from the data, but only those that are not easily characterized beforehand,which concerns the unforced dynamics of the building. PCNN has also been tested with multi-zone building [9], showing good results especially when considering a physical and black-box module for each thermal zone. Drgona et al. [10] developed a multi-block neural model linking heat transfer coefficients to eigenvalues, with penalties enforcing physical realism. The block structure applies separate constraints on inputs and disturbances, while slack variables in training penalize constraint violations to maintain realistic limits.

The literature primarily focuses on specific use cases [27] or architectural innovations aimed at improving the performance of PiNN models [8], [17], highlighting their versatility in building applications, while the scientific novelty of this work is the prioritization of the evaluation of PiNN models' behavior under multiple varying scenarios of training data availability. This study specifically investigates how physics-informed models compare to conventional deep learning approaches when faced with one of the most pressing challenges in building management and control: the data scarcity. By addressing this critical limitation, the study positions PiNNs as a promising solution to improve model reliability and effectiveness in data-constrained scenarios. Moreover, while most prior research has focused exclusively on HVAC power data, either for heating or cooling, this work takes a more comprehensive approach by undertaking datasets that encompass the entire year, capturing both heating and cooling scenarios. Besides that, while other studies have applied the same methodology to different building scenarios, such as [1], our work extends the analysis to various climatic conditions. By utilizing datasets from four cities with distinct weather patterns (Turin, Munich, Copenhagen, Madrid), this study assesses the model's adaptability to diverse environmental contexts, providing a broader evaluation of its generalization capabilities and making it reproducible and location independent.

3 Methodology

This section outlines the development and evaluation process of PiNN model for predicting indoor air temperature dynamics in buildings, as illustrated in Fig. 1.

The methodology begins with the generation of datasets representing the same building in different cities. This process involves the customization of *weather files* and *building model*, *EnergyPlus* simulations, and *data normalization*. The second step involves problem formulation, including the definition of the governing *Ordinary Differential Equation (ODE)* that captures the physical behavior of air temperature variations within a building. This equation, along with its associated *parameters* that are properly estimated, is embedded into the *PiNN* framework to ensure the model adheres to the underlying physical dynamics.

During the *Training* phase, the datasets are divided into several configurations with increasing amounts of data to evaluate the integration of physics-based constraints into the neural network. An LSTM model is trained on the same datasets, enabling a comparative analysis to benchmark the PiNN's performance against that of a conventional deep learning model. In the *Testing* phase, this evaluation highlights the benefits of the PiNN approach for temperature prediction, while systematically assessing its adaptability and robustness across diverse environmental and training scenarios. The rest of this section will describe in detail the proposed methodology.

3.1 Datasets

The datasets comprise simulated data generated using the output from EnergyPlus [5] simulations. EnergyPlus is a building simulation program capable of modeling both conditioned and unconditioned spaces, offering customizable timesteps and detailed

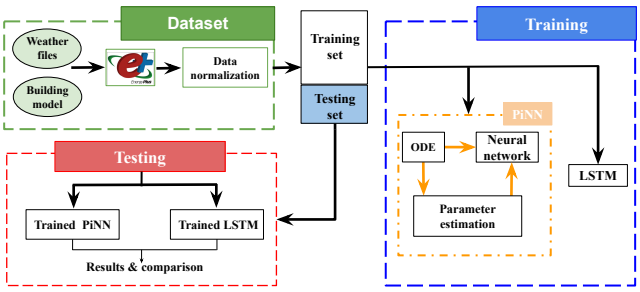


Figure 1: Methodology of the presented work

representations of dynamic interactions. In this study, the simulated building is a simplified model consisting of a single thermal zone, commonly referred to as a shoebox model, and an ideal HVAC system.

To generate the data required for the model, as shown in the top-left block in Fig. 1, EnergyPlus simulations rely on weather files in the *.epw* format. EnergyPlus offers example weather files for nearly every country worldwide [12], which are generally based on "typical meteorological year" data rather than specific years. These files represent averaged conditions derived from historical weather data to ensure broad applicability. However, for this study, it was essential to use weather files corresponding to three specific years, 2021, 2022, and 2023, for each city to capture year-specific climatic variations. To generate realistic datasets, additional weather data had to be integrated into the example *.epw* files and adapted to meet the study's specific requirements. The Open Meteo Historical Weather API [36] offers a collection of reanalyzed data, combining weather station, aircraft, buoy, radar, and satellite observations to generate a comprehensive record of past weather conditions. To accurately modify the *.epw* files, eleven variables were necessary for each city to adjust the data: i) air temperature, ii) air dewpoint temperature, iii) air relative humidity, iv) rainfall, v) snow depth, vi) surface pressure, vii) wind speed, viii) wind direction, ix) Global Horizontal Irradiation (GHI), x) Diffuse Horizontal Radiation (DHI), and xi) Direct Normal Irradiance (DNI). The customized epw files accept data at an hourly resolution, which is then interpolated by the software to achieve the desired timestamp.

The selected cities, Turin (Italy), Munich (Germany), Copenhagen (Denmark), and Madrid (Spain), represent a diverse range of climatic conditions. Turin captures a Mediterranean climate, with hot summers and cold winters, Munich shows cold winters, but warm summers, Copenhagen represents cooler northern conditions, and Madrid exemplifies a warmer southern European climate, with very hot summers, mild winters, and less rainfall. The finalized epw files were used in the simulation within the EnergyPlus software, along with the building model input file, to extract the datasets directly from the output files. Each dataset is structured with labeled data points consistent with the governing ODE employed in the model. The input variables include HVAC thermal power, past indoor temperature, solar gains, represented by the global horizontal radiation and denoted as G, internal gains, specifically the sum of occupants' and lighting equipment thermal gains, denoted as I, and outdoor temperature, while the output variable is the indoor air temperature.

The resulting complete datasets span three years of continuous data for each city, simulated with a 10-minute temporal resolution.

It was decided to work with synthetic data in order to conduct a deeper and more detailed analysis, both in terms of the cities analyzed and training data and cope with the data scarcity issue. This approach allowed for the evaluation of results for the same building, but located different cities. Anyway, this method can be replicated with real-world data, if available.

3.2 Problem Formulation

To incorporate physical knowledge into the proposed framework, an appropriate Resistance-Capacitance (RC) model has been adopted. This approach leverages the analogy between heat transfer and electrical circuits, where thermal resistance and capacitance are represented by electrical resistors and capacitors [32]. Consequently, the system's behavior can be mathematically expressed using one or more Ordinary Differential Equations.

In this study, the ODE corresponds to a simple single-zone model comprising only a capacitance value and a resistance (1R1C), which represent the thermal interaction between indoor and outdoor environments. Within this analogy, thermal input flows are modeled as current sources, temperature differences as voltage differences, and the outdoor temperature as a voltage generator, as shown in Fig. 2. This representation effectively captures the temperature gradient between the building's interior and exterior, treating it as an energy source driving thermal flow through the building envelope. The indoor temperature, the main variable to be determined, corresponds to a node in the circuit, where Kirchhoff's law is applied. The resulting lumped-parameter differential equation (1) simplifies the system by neglecting spatial dimensions and includes key parameters: the thermal resistance *Rv* between indoor and outdoor environments, the internal thermal capacitance *Ci*, and an α parameter. The latter adjusts the solar radiation flux by accounting for factors such as the surface area of solar incidence and the complex dynamics of heat exchange.

$$\frac{dT_i}{dt} = \frac{1}{Rv \cdot Ci} \cdot (T_{ext} - T_i) + \frac{G \cdot \alpha}{Ci} + \frac{\dot{Q}}{Ci} + \frac{I}{Ci} \tag{1}$$

Here, T_i and T_{ext} are the room temperature and the outdoor air temperature, G and I are the solar gains and the internal heat gains respectively and $\dot{Q}$ is the HVAC power supplied to the building.

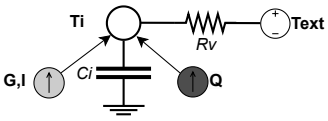


Figure 2: 1 Resistance - 1 Capacitance model

3.3 Physics-informed Neural Network

The solution to any equation can be approximated by a neural network. Automatic differentiation techniques can be used to physics-inform the models [28] by calculating their derivatives with respect to input coordinates, such as space and/or time. The residual of the neural network approximation is denoted as *f(t,x)*, and as it approaches zero, it indicates that the network adheres to the physical law. The neural network generates the solution elements of the differential equation as outputs and the automatic differentiation takes place in the subsequent layer. Domain knowledge is incorporated outside the network architecture by restructuring the loss function into a two-terms formulation, see Eq. (3), thus the model is penalized whenever it does not respect the physical principles.

$$Loss = L_{nn} + L_{phys} \tag{2}$$

The first term, L_{nn}, represents the conventional prediction error over the training data and is typically defined as the Mean Squared Error (MSE) between the predicted and the actual value. On the other hand, L_{phys} incorporates the physical-knowledge and expresses the residuals of the governing equation. The loss function is thus expressed as follows:

$$Loss = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{u}_{i+1} - \hat{\mathbf{u}}_{i+1})^2 + \frac{1}{N}\sum_{i=1}^{N}|f(t_i, x_i)|^2 \tag{3}$$

where u_{i+1} is the observed output, $\hat{u}_{i+1}$ is the predicted one, N indicates the number of samples, ti and xi the spatio-temporal coordinates.

In the proposed work, physical law is explicitly incorporated into a fully connected neural network architecture. The model is trained to predict the next value of indoor air temperature, utilizing external weather data, such as outdoor air temperature and solar radiation, as well as inputs from the HVAC system's thermal power, previous indoor air temperature and internal heat gains, all obtained through EnergyPlus simulations. The residual *f(t,x)*, expressed in Eq. (4), is thus specifically formulated to reflect the governing equation of indoor air temperature dynamics in Eq. (1), but it reduces to *f(t)* as it only depends on time:

$$f(t) = \frac{\partial u}{\partial t} - \frac{1}{Rv}\frac{T_{ext}}{Ci} + \frac{1}{Rv}\frac{u}{Ci} - \frac{1}{Ci}(\alpha \cdot G + I + \dot{Q}) \tag{4}$$

where *u=u(t)* is the neural network approximation. The conventional term into the loss function can be written as:

$$L_{nn} = \frac{1}{N}\sum_{i=1}^{N}\left(T_i - \hat{T}_i\right)^2 \tag{5}$$

where T_i is the synthetic room air temperature and $\hat{T}_i$ is the predicted one. This explicit formulation of the optimization problem ensures that the physical laws governing the indoor air temperature dynamics are seamlessly embedded into the learning process, enhancing the model's ability to predict temperature variations in response to environmental and system-related factors.

To optimize the PiNN model's architecture for temperature prediction, various configurations were explored. They included different numbers of hidden layers, ranging from 2 to 8 layers, in alignment with the approaches discussed in [7]. Additionally, batch sizes (64, 100, 128, and 256) were tested to identify the optimal configuration for model training. The model is a single-step prediction model, where the output is predicted based solely on the input at the previous time step. A grid search was performed to tune the hyperparameters, and the best performing configuration was selected based on validation accuracy and loss minimization. The optimal architecture found through this grid search consisted of 8 layers MLP, with 8 neurons in each layer. Batch size was adjusted depending on the specific dataset, allowing for flexibility in handling varying data volumes and characteristics. The optimization

algorithm used for training is the Adaptive Moment Estimation (ADAM Optimizer) with a learning rate of $1e-4$.

3.4 Long Short-Term Memory Neural Network

The LSTM architecture belongs to the RNN category and is designed to solve the "vanishing-gradient" problem, allowing the model to retain information over long distances. This makes LSTM particularly suitable for time-series handling and forecasting. Therefore, LSTM-based architectures are considered a state-of-the-art method in this field [15], achieving satisfying performance in indoor temperature prediction [21], [13], [23]. This is why it has been chosen as the benchmark model in this work (Fig. 1). The architecture that led to the best results was determined through a grid search of the hyperparameters. The optimal configuration consists of two layers, each with 64 units, and a batch size of 256. Similar to the previous model, the ADAM optimizer was used with a learning rate of $1e-3$.

3.5 Testing

As outlined in Fig. 1, in the lower block, the final part of the implemented methodology involves the testing and validation phase of the trained PiNN and LSTM architectures. Their performance was evaluated using the Mean Absolute Error (MAE), a widely used metric to assess prediction accuracy. The mathematical expression for MAE is:

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{6}$$

where y_i are the predicted values, $\hat{y}_i$ are the simulated values, and N is the number of predictions. A lower MAE indicates a better model performance, as it reflects a smaller average absolute error between the predicted and observed values. When discussing indoor air temperature in a building, it is essential to consider its impact on the thermal comfort of the occupants. Thermal comfort is commonly evaluated using the Predicted Mean Vote (PMV), based on Fanger's thermal comfort theory, which assigns a numerical value reflecting the occupants' thermal sensation. A variation in air temperature of +2 or -2°C only slightly affects the PMV index. Furthermore, studies like [6] demonstrate that a temperature variation of 2°C does not significantly impact the occupants' performance. Therefore, an acceptable range for temperature predictions is typically considered to be between -2 and +2 °C.

4 Experimental results

This section evaluates the results obtained with the proposed methodology. First, the estimated ODE parameters and the details of the adopted training and testing datasets are described. Afterwards, the performance of the PiNN model for indoor air temperature prediction is evaluated and compared to that of the LSTM model.

4.1 Ordinary Differential Equation's parameters

For each considered city, the ODE parameters, such as thermal resistance (Rv), capacitance (Ci), and the adjustment factor for solar radiation (α), are optimized, through a non-linear least-squares fitting procedure, using the given data to accurately capture the building's thermal dynamics. The results are listed in Table 1.

Table 1: Ordinary Differential Equation parameters

	Turin	Munich	Copenhagen	Madrid
$R_v\ [K/W]$	$4.5e{-3}$	$4.2e{-3}$	$3.7e{-3}$	$4.8e{-3}$
$C_i\ [kJ/kg]$	27.9	24.3	39.2	41.7
$\alpha\ [m^2]$	3.1	4.1	5.6	2.3

Table 2: Dataset Periods for Training and Testing

	Training period	Testing period	Number of samples
1.	01.01 - 13.01 2021	14.01 - 31.01 2021	1344
2.	01.01 - 31.01 2021	01.01 - 28.02 2021	2688
3.	01.01 - 31.03 2021	01.10 - 31.12 2021	8064
4.	2021	2022	52560
5.	2021 - 2022	2023	105120

Considering that the analyzed building, described by the EnergyPlus model, remains the same, any differences in parameter estimation are likely due to variations in weather conditions. Turin and Munich experience similar climates, although Munich is colder in winter. This is probably why the values for these two cities are comparable, with α differing by only 1 m^2, Rv by $3e-4$ K/W, and Ci by 3.6 kJ/kg.

Copenhagen and Madrid, on the other hand, show distinct values, likely due to challenging conditions that need to be considered. For instance, Copenhagen experiences extreme winds throughout the year, while Madrid faces significant temperature differences between winter and summer, with very hot days recorded during the summer months. Indeed, the Ci value for Madrid (41.7 kJ/kg) and Copenhagen (39.2 kJ/kg) are quite different from those of the other two cities. The same can be said for the α values, respectively of 5.6 and 2.3 m^2.

4.2 Datasets divisions

As mentioned in Section 3, the models were trained on different subsets of the data, which included 2 weeks, 1 month, 3 months, 1 year, and 2 years, in order to evaluate their performance with varying data volumes. Table 2 outlines the selected periods. In total, five different cases were considered.

The first case involved training the model on the first two weeks of January 2021, with the last two weeks used for testing. The second case trained the model on the entire month of January 2021 and tested it on the whole February dataset of the same year.

The third case trained the model on the first 3 months of 2021 (January, February, March) of the year and tested it on the last 3 months of 2021 (October, November, December), providing comparisons between different seasons (winter for training and autumn for testing). The last two configurations incorporated both cooling and heating data: the fourth case trained the model on the entire year of 2021 and tested it on 2022, while the final case used 2021 and 2022 for training and 2023 for testing. The same configurations were adopted for each city.

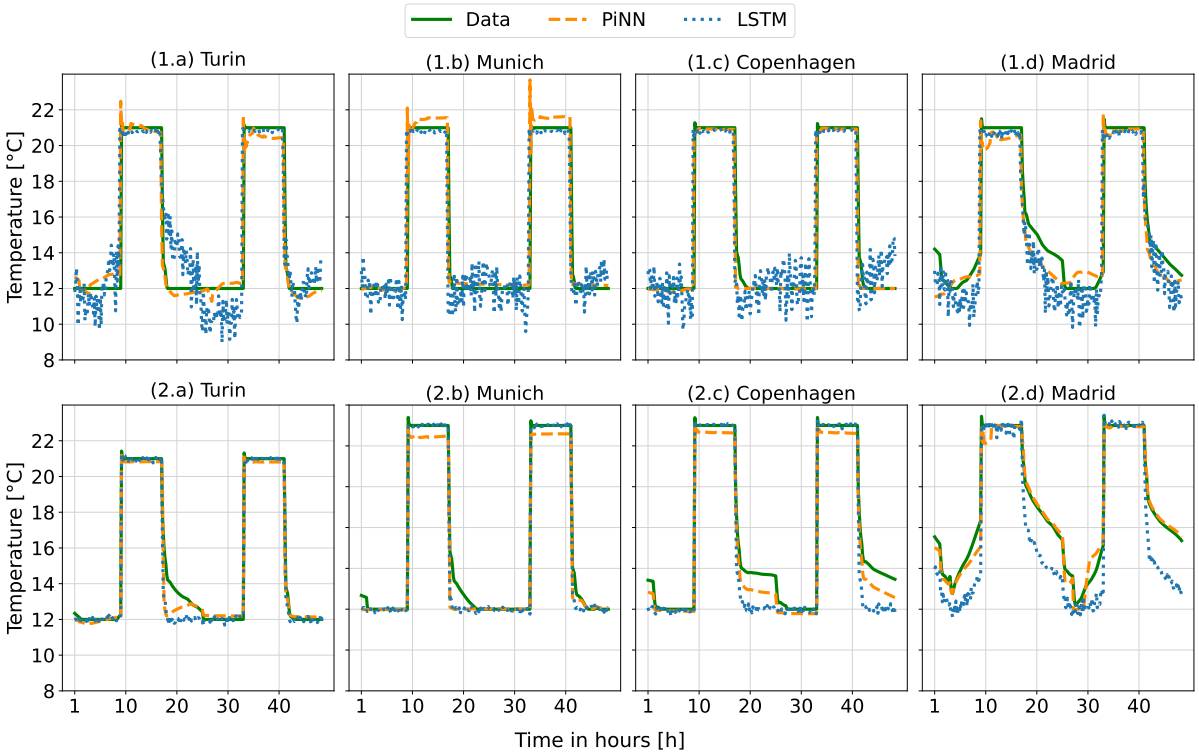


Figure 3: Prediction comparison between PiNN and LSTM for bi-weekly (Fig.1) and yearly (Fig. 2) results

Table 3: PiNN and LSTM validation loss comparison

		Validation Loss - MAE [°C]				
		2 weeks	1 month	3 months	1 year	2 years
Turin	PiNN	0.40	0.90	0.59	0.55	0.49
	LSTM	0.56	1.57	0.68	0.17	0.59
Munich	PiNN	0.43	0.15	0.56	0.30	0.34
	LSTM	0.55	0.49	0.73	0.13	0.18
Copenhagen	PiNN	0.10	0.70	0.79	0.67	0.29
	LSTM	0.51	0.78	0.94	0.21	0.16
Madrid	PiNN	0.33	1.42	1.84	0.34	0.12
	LSTM	0.62	1.81	1.35	0.26	0.28

4.3 Comparative analysis of PiNN and LSTM models

Table 3 presents the prediction performance of the PiNN model compared to the fully data-driven LSTM model in terms of MAE. It is evident that when the models are trained with smaller datasets, such as the 2 weeks, 1 month, and 3 month cases, the physics-informed model outperforms the conventional LSTM. This is attributed to the physical principles incorporated into the PiNN model's training, enabling it to achieve better temperature forecasting results. For the last two cases, involving 1 and 2 years of data, the LSTM model demonstrates lower errors, though its performance remains comparable to that of the PiNN, and this may result from its ability to learn complex temporal patterns, given sufficient data.

When examining the data for Turin, the PiNN model achieves its lowest error of 0.40 °C when trained with just 2 weeks of data. The highest error occurs in the 1-month dataset case, where the PiNN error is 0.90 °C. However, even in this scenario, the PiNN outperforms the LSTM, which has a significantly higher error of 1.57 °C, reinforcing the PiNN's superior performance when trained on smaller datasets. The dataset where the PiNN shows the most consistent prediction performance across all scenarios is Munich. In this case, the lowest error occurs with the 1 month dataset (0.15 °C), while the highest error is 0.56 °C with the 3 months dataset, when the LSTM stands with 0.73 °C. This suggests that the physical model and the parameters incorporated into the PiNN's architecture are particularly well-suited to the data from Munich, contributing to better stability and lower errors over a range of time periods. Madrid, on the other hand, is the dataset where both models perform worst, with errors surpassing 1 °C for both the 1 month and 3 months cases, reaching up to 1.8 °C. It is important to note that the simpler 1R1C model, which underpins the PiNN's predictions, may limit its performance in more complex building systems or under certain conditions. The 1R1C model's assumptions about thermal behavior might not capture all the challenges of real-world systems, which could help explain why, in some cases, the PiNN's performance is less accurate. However, in all cases, the errors remain below the 2 °C threshold, which is still considered acceptable based on Fanger's perception model, meaning the predictions are still within reasonable comfort limits.

Between all cases, the best result for the PiNN corresponds to a 0.10 °C MAE achieved with the Copenhagen dataset, when the model was trained on the first two weeks of January 2021 and tested on the last two weeks of the same month, compared to the 0.51 °C MAE assessed by the LSTM model. However, for the Madrid dataset with 3 months of training data, the LSTM (1.35 °C) edges

out the PiNN (1.84 °C), likely due to the increased complexity and variability of the dataset. This result is particularly significant as it represents the PiNN's most challenging scenario among the first three cases, providing valuable insights into its limitations. The substantial differences in external conditions across these months may require a more sophisticated thermal model to fully capture the interactions. In this case, the LSTM's data-driven approach might be better equipped to adapt to the changing dynamics present in the dataset. Generally speaking, the 3 months training case is where both models exhibit the worst performance, which could be attributed to the comparison of data from two different seasons. For example, the October dataset may resemble a summer climate, with no heating power supplied, while the training data from January, February, and March reflects winter conditions. These seasonal discrepancies introduce variability that both models struggle to capture effectively, impacting their predictive accuracy.

Physical adherence is also evident when examining the prediction outcomes. Figure 3 shows the differences between the PINN and LSTM models in predicting two days in January. The top part (Fig. 3.1) displays the results obtained after training on data from the first two weeks of January 2021, while the bottom part (3.2) presents the results based on training with the full dataset from 2021. For nearly every city, the PINN model accurately captures the temperature trend, maintaining its prediction close to the temperature setpoint during both day and night hours, when no thermal power is supplied to the building. While the PINN model sometimes exhibits slight underpredictions, as seen in Munich and Copenhagen in Figure 3.2b and 3.2c, or overpredictions, as for Munich in Figure 3.1b, the error remains within the acceptable threshold of ±2 °C. The LSTM model, on the other hand, produces a less stable and highly fluctuating output, which is inconsistent with the underlying physical phenomenon. This issue improves when the LSTM model is trained on a full year's worth of data, allowing it to follow the trend more consistently. However, in the case of Madrid (Fig. 3.2c), the model's performance deteriorates, resulting in a less accurate prediction.

To gain deeper insights into the error tendencies and the differences between the PiNN and LSTM outcomes when trained with limited data, Fig. 4 illustrates the residual autocorrelation factors for the two weeks (Fig. 4.1) and for one year (Fig. 2.2), up to a lag of 20. By default, a lag of 1 is used, meaning that predictions are based on the current timestep to forecast the next value. Given that the time-series timestep is 10 minutes, a lag of 20 corresponds to 200 minutes. In general, for Figure 4.1, the residual autocorrelation for the PINN model is lower, indicating that the model is better at capturing the underlying temporal structure of the data. In particular, for Copenhagen (Fig. 4.1c), the PINN model exhibits its best performance, with the autocorrelation factor approaching zero. For the one-year dataset, the residual autocorrelation values are higher, but they decrease with higher lags, suggesting that long-term dependencies are better handled by both models. For Munich (Fig. 4.2b), both models show very similar outcomes, while for Copenhagen (Fig. 4.2c), the LSTM model performs better.

To further investigate the PiNN's capabilities, Fig. 5 highlights the error distributions for both the PiNN and LSTM models when trained on the Munich full 2021 dataset and tested on the 2022 one. Munich was selected as it represents a scenario where both models exhibit competitive performance, with the LSTM achieving its lowest error of 0.13 °C compared to the PiNN's 0.30 °C. This balanced comparison allows for a more comprehensive analysis of the strengths of the PiNN with respect to the LSTM.

The boxplots break down the errors across quarterly intervals, roughly corresponding to the four seasons: winter (January to March), spring (April to June), summer (July to September), and autumn (October to December). The y-axis represents prediction deviations in degrees Celsius (°C), where positive and negative values indicate over- and under-predictions, respectively.
A tighter interquartile range and shorter whiskers are ideal for achieving more consistent predictions, as they indicate less variability and greater stability in the results. This trend is reflected in the performance of the PiNN and LSTM models across the seasonal intervals. Overall, the PiNN consistently shows lower mean errors than the LSTM, except in November, where the errors are comparable. However, during January and Febraury, the PiNN's upper whisker is higher, suggesting it tends to over-predict air temperature in these months. In contrast, the LSTM exhibits a wider error spread, particularly during the summer months (July-September), when temperature dynamics likely involve more complex interactions. The PiNN, however, maintains a more compact distribution, reflecting its stable performance across seasons.
For the December data, shown in Fig. 5d, although the PiNN's distribution shows less symmetry, its mean error is still lower than the LSTM's, approximately 0.03 °C. This further supports the PiNN's superior accuracy and stability, even when the distribution is less symmetric, reinforcing its ability to maintain robust performance across varying seasonal conditions.

All in all, the analysis highlights the PiNN model's strength in delivering stable and accurate predictions across various training scenarios, particularly under conditions of limited data availability. While the LSTM model shows promise in data-rich settings, the PiNN consistently demonstrates a better balance between accuracy and physical consistency, making it particularly well-suited for practical applications in building energy management. This is evident across the four selected cities, Turin, Munich, Copenhagen, and Madrid, which represent a diverse range of climatic conditions and validate the PiNN's adaptability to varying environmental contexts.

5 Conclusion

In this work, a PiNN was developed to predict the indoor air temperature of a building. The model was trained using four different datasets representing various climatic conditions. These datasets were further divided into subsets with varying amounts of training data to evaluate the model's performance under different levels of data availability and conditions. The PiNN's performance was compared with that of a conventional deep learning model, specifically an LSTM neural network. The physics-informed model demonstrated superior performance when trained with limited data, achieving the lowest MAE of 0.10 °C, underscoring its ability to generalize effectively through the incorporation of physical constraints. When trained with the largest dataset, it achieved results comparable to the LSTM model. Additionally, the PiNN's performance remained consistent across datasets from different cities, suggesting

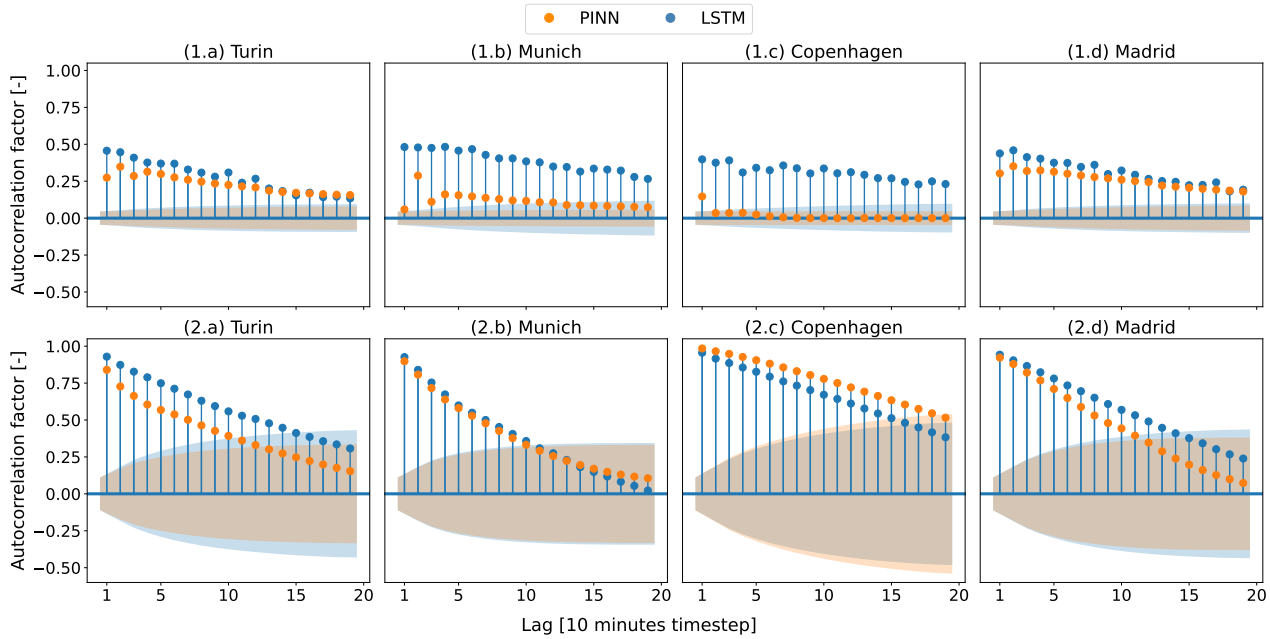


Figure 4: Residual autocorrelation for the model's bi-weekly (Fig. 1) and yearly (Fig.2) results

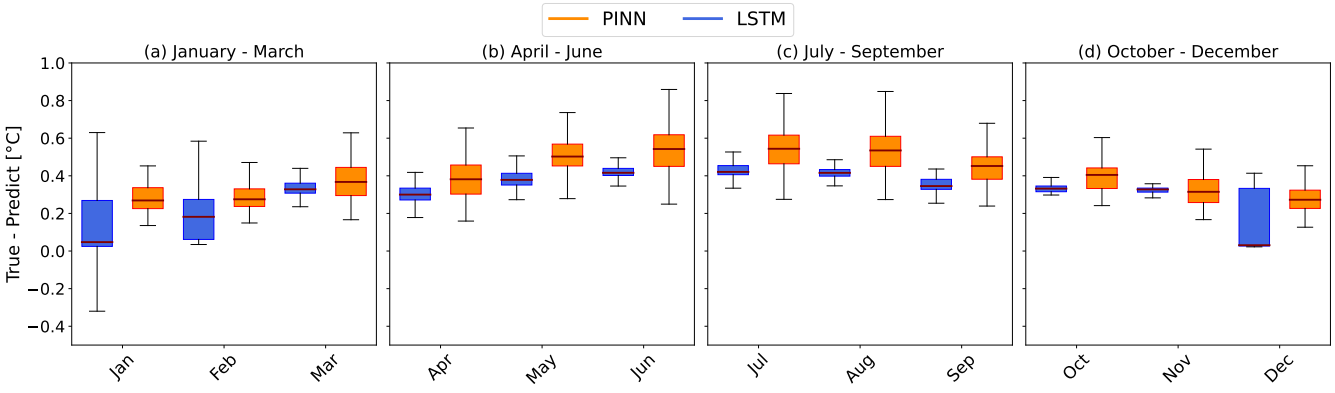


Figure 5: Munich boxplot seasonal comparison

that the methodology is unaffected by the building's location and is therefore robust and repeatable without modification.

Considering the model's structure, the PiNN's performance was limited by the building's physical representation, particularly the 1R1C equation. Upcoming investigations will assess the impact of using more detailed RC models, such as 2R2C or 3R2C, to better describe thermal dynamics by incorporating additional temperature nodes, both measurable and non-measurable, as well as multiple thermal zones. In this study, the ODE parameters were estimated outside the training process, which limited their accuracy. Future work could address this limitation by integrating these parameters directly into the neural network, enabling the model to optimize them as part of its training process. While the current study focused on assessing the model performance across different climatic conditions, future work will also explore the influence of occupancy and internal gains, such as electrical appliances and heating/cooling systems, on model behavior. These factors will be investigated to improve model accuracy and provide a more comprehensive understanding of their effect on building energy performance. Further research directions could, furthermore, include exploring alternative architectures for physics-informed learning, such as Recurrent Neural Networks (RNNs), which are particularly well-suited for time series data and may enhance model accuracy. Additionally, comparing the PiNN architecture with other well-established models from the literature, including simpler, traditional approaches, such as Euler Explicit Scheme, would help highlight the trade-offs between complexity and performance.

Another promising avenue could involve integrating the PiNN into a reinforcement learning control strategy to enable real-time temperature optimization and adaptive control. Future work will consider the computational challenges associated with such integrations, as well as their practical application in enhancing energy-efficient building management.

References

[1] Yongbao Chen, Qiguo Yang, Zhe Chen, Chengchu Yan, Shu Zeng, and Mingkun Dai. 2023. Physics-informed neural networks for building thermal modeling and demand response control. *Building and Environment* 234 (2023), 110149.

[2] Zhe Chen, Fu Xiao, Fangzhou Guo, and Jinyue Yan. 2023. Interpretable machine learning for building energy management: A state-of-the-art review. *Advances in Applied Energy* 9 (2023), 100123.

[3] European Commission. 2020. Stepping up Europe's 2030 climate ambition. (2020). https://eur-lex.europa.eu/legal-content/EN/TXT/uri=CELEX:52020DC0562

[4] Nikolina Čović, Hrvoje Pandžić, and Yury Dvorkin. 2022. Learning indoor temperature predictions for optimal load ensemble control. *Electric power systems research* 211 (2022), 108384.

[5] Drury B Crawley, Linda K Lawrie, Curtis O Pedersen, and Frederick C Winkelmann. 2000. Energy plus: energy simulation program. *ASHRAE journal* 42, 4 (2000), 49–56.

[6] Weilin Cui, Guoguang Cao, Jung Ho Park, Qin Ouyang, and Yingxin Zhu. 2013. Influence of indoor air temperature on human thermal comfort, motivation and performance. *Building and Environment* 68 (2013), 114–122. doi:10.1016/j.buildenv.2013.06.012

[7] Salvatore Cuomo, Vincenzo Schiano Di Cola, Fabio Giampaolo, Gianluigi Rozza, Maziar Raissi, and Francesco Piccialli. 2022. Scientific machine learning through physics–informed neural networks: Where we are and what's next. *Journal of Scientific Computing* 92, 3 (2022), 88.

[8] L. Di Natale, B. Svetozarevic, P. Heer, and C.N. Jones. 2022. Physically Consistent Neural Networks for building thermal modeling: Theory and analysis. *Applied Energy* 325 (2022), 119806. doi:10.1016/j.apenergy.2022.119806

[9] Loris Di Natale, Bratislav Svetozarevic, Philipp Heer, and Colin Neil Jones. 2023. Towards scalable physically consistent neural networks: An application to data-driven multi-zone thermal building models. *Applied Energy* 340 (2023), 121071.

[10] Ján Drgoňa, Aaron R Tuor, Vikas Chandan, and Draguna L Vrabie. 2021. Physics-constrained deep learning of multi-zone building thermal dynamics. *Energy and Buildings* 243 (2021), 110992.

[11] EnergyPlus. 2021. EnergyPlus Documentation. https://energyplus.net/documentation

[12] EnergyPlus. 2021. EnergyPlus Weather Data. https://energyplus.net/weather

[13] Zhen Fang, Nicolas Crimier, Lisa Scanu, Alphanie Midelet, Amr Alyafi, and Benoit Delinchant. 2021. Multi-zone indoor temperature prediction with LSTM-based sequence to sequence model. *Energy and Buildings* 245 (2021), 111053.

[14] Salah A Faroughi, Nikhil Pawar, Celio Fernandes, Maziar Raissi, Subasish Das, Nima K. Kalantari, and Seyed Kourosh Mahjour. 2023. Physics-Guided, Physics-Informed, and Physics-Encoded Neural Networks in Scientific Computing. arXiv:2211.07377 [cs.LG] https://arxiv.org/abs/2211.07377

[15] Rafael Natalio Fontana Crespo, Alessandro Aliberti, Lorenzo Bottaccioli, Edoardo Pasta, Sergej Antonello Sirigu, Enrico Macii, Giuliana Mattiazzo, and Edoardo Patti. 2024. A comparative analysis of Machine Learning Techniques for short-term grid power forecasting and uncertainty analysis of Wave Energy Converters. *Engineering Applications of Artificial Intelligence* 138 (2024), 109352. doi:10.1016/j.engappai.2024.109352

[16] Maryam Gholamzadehmir, Claudio Del Pero, Simone Buffa, Roberto Fedrizzi, and Niccolo' Aste. 2020. Adaptive-predictive control strategy for HVAC systems in smart buildings – A review. *Sustainable Cities and Society* (2020).

[17] Gargya Gokhale, Bert Claessens, and Chris Develder. 2022. Physics informed neural networks for control oriented thermal modeling of buildings. *Applied Energy* 314 (2022), 118852.

[18] Md Monir Hossain, Tianyu Zhang, and Omid Ardakanian. 2021. Identifying grey-box thermal models with Bayesian neural networks. *Energy and Buildings* 238 (2021), 110836.

[19] IEA. 2022. *World Energy Outlook 2022, IEA, Paris.* Technical Report. https://www.iea.org/reports/world-energy-outlook-2022

[20] Julien Leprince, Henrik Madsen, Clayton Miller, Jaume Palmer Real, Rik van der Vlist, Kaustav Basu, and Wim Zeiler. 2022. Fifty shades of grey: Automated stochastic model identification of building heat dynamics. *Energy and Buildings* 266 (2022), 112095.

[21] Ming Li, Yijun Li, and Xinli Min. 2020. Practice and Application of LSTM in Temperature Prediction of HVAC System. In *2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC).* 1000–1004. doi:10.1109/ITOEC49072.2020.9141910

[22] Xiao Luo, Yifei Wang, Qing Zhu, Hanyang Liu, Shuhong Wang, and Minghe Wu. 2024. Physics-Informed Neural Network for Parameter identification of Air Conditioning Load Models. In *2024 7th International Conference on Energy, Electrical and Power Engineering (CEEPE).* IEEE, 948–953.

[23] Fatma Mtibaa, Kim-Khoa Nguyen, Muhammad Azam, Anastasios Papachristou, Jean-Simon Venne, and Mohamed Cheriet. 2020. LSTM-based indoor air temperature prediction framework for HVAC systems in smart buildings. *Neural Computing and Applications* 32 (2020), 17569–17585.

[24] Srinarayana Nagarathinam, Yashovardhan S Chati, Malini Pooni Venkat, and Arunchandar Vasan. 2022. PACMAN: physics-aware control MANager for HVAC. In *Proceedings of the 9th ACM international conference on systems for energy-efficient buildings, cities, and transportation.* 11–20.

[25] Srinarayana Nagarathinam and Arunchandar Vasan. 2024. PhyGICS–A Physics-informed Graph Neural Network-based Intelligent HVAC Controller for Open-plan Spaces. In *Proceedings of the 15th ACM International Conference on Future and Sustainable Energy Systems.* 203–214.

[26] Surya Venkatesh Pandiyan and Jayaprakash Rajasekharan. 2022. Physics-Informed Neural Network Model for Flexibility Modeling of Electric Water Heaters. In *2022 International Conference on Smart Energy Systems and Technologies (SEST).* IEEE, 1–6.

[27] Fabio Pavirani, Gargya Gokhale, Bert Claessens, and Chris Develder. 2024. Demand response for residential building heating: Effective Monte Carlo Tree Search control based on physics-informed neural networks. *Energy and Buildings* 311 (2024), 114161.

[28] M. Raissi, P. Perdikaris, and G.E. Karniadakis. 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* 378 (2019), 686–707. doi:10.1016/j.jcp.2018.10.045

[29] Hannah Ritchie, Pablo Rosado, and Max Roser. 2020. Emissions by sector. *Our World in Data* (2020). https://ourworldindata.org/emissions-by-sector.

[30] Amir Shahcheraghian, Hatef Madani, and Adrian Ilinca. 2024. From white to black-box models: A review of simulation tools for building energy management and their application in consulting practices. *Energies* 17, 2 (2024), 376.

[31] Ruihang Wang, Zhiwei Cao, Xin Zhou, Yonggang Wen, and Rui Tan. 2023. Phyllis: Physics-Informed Lifelong Reinforcement Learning for Data Center Cooling Control. In *Proceedings of the 14th ACM International Conference on Future Energy Systems.* 114–126.

[32] Junlu Yang, Hanning Wang, Linmiao Cheng, Zhi Gao, and Fusuo Xu. 2024. A review of resistance–capacitance thermal network model in urban building energy simulations. *Energy and Buildings* (2024), 114765.

[33] Enrui Zhang, Minglang Yin, and George Em Karniadakis. 2020. Physics-informed neural networks for nonhomogeneous material identification in elasticity imaging. *arXiv preprint arXiv:2009.04525* (2020).

[34] Tongtao Zhang, Biswadip Dey, Pratik Kakkar, Arindam Dasgupta, and Amit Chakraborty. 2020. Frequency-compensated pinns for fluid-dynamic design problems. *arXiv preprint arXiv:2011.01456* (2020).

[35] Fangzhou Guo a Jinyue Yan a Zhe Chen a, Fu Xiao a b. 2023. nterpretable machine learning for building energy management: A state-of-the-art review. (2023).

[36] Patrick Zippenfenig. 2023. *Open-Meteo.com Weather API.* doi:10.5281/zenodo.7970649