SimBuild
2024

Eleventh National Conference of IBPSA-USA
Denver, Colorado
May 21 – 23, 2024

Physics-Informed Hybrid Modeling Approach for Room Temperature Prediction Using an RC Model and Siamese Neural Network

Chul-Hong Park[1], Seongkwon Cho[1], Tae Yong Song[1], Seon-Young Heo[1], Cheol-Soo Park[1]
[1]Department of Architecture and Architectural Engineering, College of Engineering,
Seoul National University, Seoul, Republic of Korea

Abstract

In this paper, a novel hybrid modeling approach is proposed to combine the advantages of physics-based and data-driven approaches for predicting the thermal behavior of the building. It incorporates an RC model and neural networks by designing custom layers and utilizing a neural network modeling technique known as "Siamese neural network". The neural network is used to predict various time-invariant and time-varying parameters in the RC model. The modeling technique allows flexibility in the model design, simultaneous training with both time-invariant and varying parameters present, warmup period and multiple-timestep forecasting per input during the training phase, and training the model with a limited number of measured states. To validate the proposed approach, it was applied to an existing building located in South Korea, using the measured data from a single air handling unit (AHU) serving an office area located on the 5th floor. The trained model was used to predict the room air temperature for the test period. It was found that a simple RC model combined with the Siamese neural network was good enough to predict room air temperature.

Introduction

Various modeling approaches for room temperature prediction and control optimization have been explored. The physics-based approach employs the 1st-principles equations to describe the thermal behavior of the building and various heating, ventilation, and air conditioning (HVAC)-related systems and its results are intuitive and easy to interpret. However, it requires demanding efforts, e.g., modeling of 3D building geometry, estimating unknown properties and operation characteristics, and so on. On the other hand, the data-driven approach derives the relationship between input and output using available data often through the use of artificial neural networks (ANN). It is convenient and easy to implement, but its results are often unintuitive to interpret and suffer from poor generalization. Its inability to distinguish between correlation and causation of the inputs and outputs when provided with limited data is also an issue (Manfren et al. 2022). As a result, one of the challenges lies in how to combine the two approaches and overcome their drawbacks. (Ahn et al. 2022, Gokhale et al. 2022).

A thermal-network model, or RC model is one of the most common approaches due to faster load calculation with major physical dynamics, and the model being able to be written in a state-space form to be solved by the state-space solver for online controls (Li et al. 2021). In an RC model-based approach, the building is represented as a network of thermal resistances (R) and capacitances (C), and a state-space formulation following the first law of thermodynamics is used to predict the future behavior of the room/building over a finite time horizon (Klanatsky et al. 2023).

The difficulty with this approach, however, is that while time-invariant parameters in the model can easily be fitted based on the measured data, time-varying parameters such as internal loads are difficult to incorporate due to the vast search space. For this reason, various approaches in incorporating time-varying parameters into the model were taken throughout the years. For example, Berthou et al. (2014) assumed a fixed occupancy profile, and the maximum heat gain due to occupancy is included as a time-invariant parameter to be multiplied with the aforementioned profile. Kim and Park (2017) estimated heat gains and losses caused by internal heat sources (people, lights, and equipment) and airflows (infiltration, ventilation, and air movement between zones) dynamically, using the Kalman filtering method. Wang et al. (2019) approximated effective solar heat gains and internal heat gains using global irradiation on the south façade and gross electricity demand, respectively. Klanatsky et al. (2023) used CO_2 concentration measurements for occupancy gain estimation and assumed heat generation rate for a resting person for internal gains.

Using the ANN-based data-driven approach for these predictions combined with the physics-based nature of the RC network can be beneficial. However, this approach has not been thoroughly explored in previous studies. With this in mind, the authors intend to develop a hybrid model that can predict room temperature using an RC model with ANNs to predict time-varying parameters in the matrix. The main objective of this study is to develop a temperature prediction model that can be constructed from minimal building properties and data collected from the air handling unit, and then possibly be applied to model predictive control (MPC). With this, the following challenges arise:

1. *Simultaneous estimation of time-invariant and time-varying parameters*: some parameters are assumed time-invariant (e.g., internal mass), while others are considered as time-varying parameters (e.g., internal loads) that are predicted using ANNs. All of these parameters are required for the matrix calculation, and therefore simultaneous estimation of these parameters is required. In other words, the time-invariant values have to be trained alongside the weights inside the ANNs.
2. *Hidden states*: there are state/states that are part of the matrix, but are not being measured, thus referred to as hidden states. In the case of this study, the temperature of the envelope is a hidden state. To circumvent the absence of these measurements, often a warmup period is introduced, where the hidden states are given initial values and matrix calculations are carried out for a certain number of timesteps. After this warmup period, it is assumed that the hidden states have reasonable values, ready for the main calculations. The hybrid model must utilize this warmup period before forecasting, meaning that for every training sample, the matrix calculation needs to be carried out multiple times.
3. *Forecasting multiple timesteps*: for reasonable parameter and state prediction, multiple-timestep forecasting is required. Similar to the warmup period, this requires multiple matrix calculations per input, or training sample.
4. *Cannot directly train the model using the prediction outputs*: while the model predicts the parameters (time-invariant and time-varying), there is no way to directly train the model based on these predictions, because the ground truths of these parameters are unknown. In other words, the model needs to be trained based on the calculated states using the predicted parameters.

To address the aforementioned challenges, the authors present a novel physics-informed hybrid modeling approach that combines an RC model, matrix calculation layer, and a neural network modeling technique known as "Siamese neural network". The structure as well as the key components of this complex model (hereinafter referred to as "hybrid model") will be further explained in the next section.

Methods

Hybrid model structure

The physics-informed hybrid model is a complex network that combines conventional neural networks, and an RC model-based matrix calculation layer using a technique called "Siamese neural network" (Figure 1). A Siamese neural network refers to sections inside a neural network that share the weight matrices (Koch et al. 2015). In other words, a part of a neural network that is called multiple times within a single input, which is crucial for the repeated calculations needed for the aforementioned warmup and multiple-timestep forecasting.

The key components of this hybrid model are as follows:

1. Multiple matrix calculation per input array: as discussed previously, due to the warmup period and multi-timestep forecast, multiple matrix calculation is required ($n_{total} = n_{warmup} + n_{forecast}$ timesteps). As a result, a single input and output array is designed to be concatenated arrays containing multiple inputs and outputs (states) of n_{total} timesteps. When inputted into the hybrid model, the input layer inside the model is responsible for splicing the input array into separate timestep-wise pairs and appropriately connecting these pairs to corresponding layers.
2. Time-varying parameter prediction network ("subnetwork"): a lightweight neural network responsible for predicting time-varying parameter values based on relevant input states. This subnetwork is called n_{total} times per input array and is thus referred to as a Siamese neural network.
3. RC model-based matrix calculation layer ("matrix layer"): a custom layer designed to hold time-invariant parameters as well as perform a state-space matrix calculation using given states, time-invariant, and time-varying parameters. This layer is also called n_{total} times per input array.
4. Custom loss function: for $n_{forecast}$ predictions made per input, both the predicted states (x_{pred}) and the change in states (Δx_{pred}) are important. While the former can be considered as the most important factor, the latter is also significant better ensure that the causality between inputs and outputs is reflected. As a result, a weighted sum of the loss functions (L_{pred}, L_{diff}) is used (1), where L_{pred} is the predicted state error (2), while L_{diff} represents the error in

predicted state change (3), and w_1 and w_2 are the weights for L_{pred} and L_{diff}, respectively.

$$L = w_1 L_{pred} + w_2 L_{diff} \quad (1)$$

$$L_{pred} = \sum_{i=1}^{n_{forecast}} \left| x_{pred,i} - x_{truth,i} \right| \quad (2)$$

$$L_{diff} = \sum_{i=1}^{n_{forecast}} \left| \Delta x_{pred,i} - \Delta x_{truth,i} \right| \quad (3)$$

The model was constructed using *Tensorflow* for Python (Abadi et al. 2015) and was fitted using the *Adam* optimizer in default settings. It is worth noting that this complex network is exclusively used for fitting and training the time-invariant parameters and subnetworks simultaneously. For actual prediction purposes (e.g., MPC), the fitted parameters and trained subnetworks can be extracted and used separately (Figure 2).

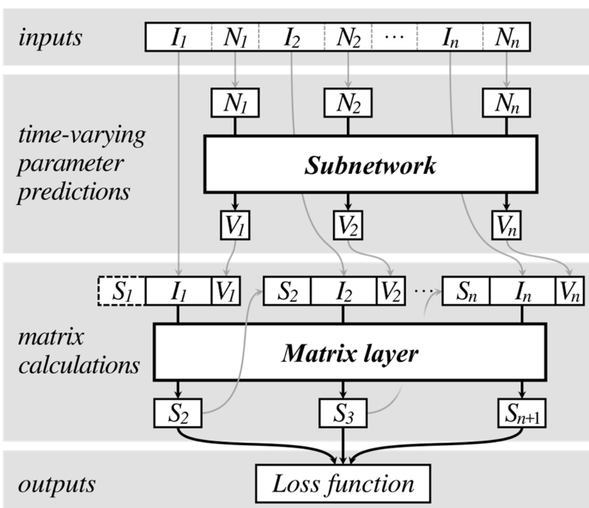

Figure 1 Hybrid model structure for training

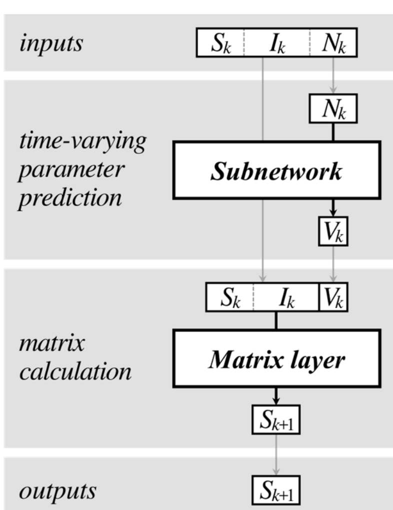

Figure 2 Hybrid model structure for predictions

- S_k: predicted states at timestep k
- I_k: measured states at timestep k
- N_k: subnetwork inputs for timestep k
- V_k: subnetwork outputs for timestep k
- n: total number of timesteps (= n_{total})

Target building

In order to evaluate the proposed modeling technique, a case study was carried out on an existing building. The target building is an industrial building located in Pyeongtaek, South Korea, and the study was mainly focused on developing a simulation model that can predict the room air temperature of a large office area located on the 5th floor served by a single AHU (Table 1). For this purpose, operational data from July 1st to 31st, 2022, were obtained at a sampling time of five minutes including outdoor air temperature, global solar irradiance, indoor air temperature, air flow rate through supply and return fans, etc. The hybrid model was trained using the data from July 1st to 24th and was validated for the testing period from July 25th to 30th (weekdays).

Table 1 Target building and zone properties

Property	Value
Building location	Pyeongtaek, South Korea
Zone location	5th floor
Zone type	office
Floor area	2,280 m^2
Aspect ratio	1.58
Building azimuth	–4°
Envelope	metal panel finish, Low-E glazing
Win-to-wall ratio	0.54(S), 0.66(N), 0.00(E,W)
AHU max. airflow	54,000 CMH

RC model

An RC network with a relatively small number of nodes was selected due to the lack of measured data in the target building. Apart from data collected from the weather station, the only available temperature data from inside the building was room temperature (T_i), and therefore a model with minimal states was considered appropriate. As a result, an RC network with two resistances and two capacitances (i.e., 2R2C), where the envelope and the internal thermal mass were lumped into a single capacitance (C_m) to reduce hidden (unknown) states, was selected. The boundaries between the target floor and the floors above and below the target floor were assumed adiabatic because the upper and lower floors were air-conditioned in a similar manner to the target floor. The structure and parameters of this model are shown in Figure 3 and Table 2, respectively.

From this model, heat balance equations (4, 5) are formulated, and then, a state-space representation of the building (6) is obtained.

$$C_m \frac{dT_m}{dt} = \frac{T_o - T_m}{R_1} + \frac{T_i - T_m}{R_2} + F_{s1} Q_s \quad (4)$$

$$C_i \frac{dT_i}{dt} = \frac{T_m - T_i}{R_2} + F_{s2} Q_s - Q_c + Q_{int} \quad (5)$$

$$\begin{bmatrix} \dot{T}_i \\ \dot{T}_m \end{bmatrix} = \begin{bmatrix} -\frac{1}{C_i R_2} & \frac{1}{C_i R_2} \\ \frac{1}{C_m R_2} & -\frac{1}{C_m R_1} - \frac{1}{C_m R_2} \end{bmatrix} \begin{bmatrix} T_i \\ T_m \end{bmatrix} + \begin{bmatrix} \frac{F_{s2} Q_s + Q_c + Q_{int}}{C_i} \\ \frac{T_o}{C_m R_1} + \frac{F_{s1} Q_s}{C_m} \end{bmatrix} \quad (6)$$

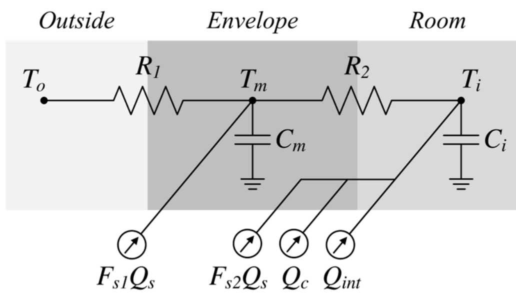

Figure 3 RC model

Table 2 RC model parameters

Symbol	Unit	Description
T_o	K	outdoor temperature
T_m	K	envelope/thermal mass temperature
T_i	K	room temperature
C_m	kJ/K	envelope/thermal mass capacitance
C_i	kJ/K	room capacitance
R_1	K/kW	outdoor–envelope thermal resistance
R_2	K/kW	envelope–room thermal resistance
Q_s	kW/m^2	global horizontal irradiance
Q_c	kW	coil load
Q_{int}	kW	internal loads
F_{s1}	m^2	envelope solar radiation factor
F_{s2}	m^2	transmitted solar radiation factor

It is worth stressing that the listed parameters should be considered as “fictitious” values. In other words, the parameter values themselves are neither direct nor explicit representations of actual properties. Thus, these values are not to be interpreted as a face value, but to be analyzed or compared based on scale and magnitude to better represent the thermal behavior of the building.

Solar radiation

To reduce complications and free-hanging parameters that can lead to failed learning processes and unstable results, the solar radiation was computed using solar geometry instead of relying on a subnetwork. As a result, the absorbed solar radiation factor (F_{s1}) and transmitted factor (F_{s2}) were obtained through the following steps:

1. For solar radiation calculation, approximation of direct normal irradiance (DNI) and diffuse horizontal irradiance (DHI) was required. As a result, Direct Insolation Simulation Code (DISC) (Maxwell 1987) was used to convert the measured global horizontal irradiance (GHI) to DNI and DHI for its simplicity of implementation.
2. The building was assumed rectangular, and from the floor area, aspect ratio, and building azimuth angle, surface incident solar radiation for each side of the walls was calculated (ASHRAE 2021).
3. Calculated incident solar radiation was proportioned to walls and windows based on the window-to-wall ratio of each side.
4. Using solar absorptance of the walls (α) and solar transmittance of windows (τ), solar absorption to walls and solar transmission through windows were calculated.

Parameters

Model parameters addressed in the previous sections can be divided into time-invariant and time-varying parameters. As mentioned previously, time-invariant parameters are fitted similarly to conventional RC model-based approaches, while time-varying parameters are predicted using subnetworks (ANNs).

Time-invariant parameters and their respective ranges used in this study are listed in Table 3. The ranges were selected based on rough estimations of the physical properties. Due to the lack of detailed information regarding the properties of the target building, the ranges of the parameters were purposefully set widely. For example, the range of C_i was selected based on the air volume of the target zone (12,540 m^3), and physical properties (e.g., density and specific heat) of air. The range of C_m was set based on the exterior area (1,078 m^2) and physical properties of typical metal panel-finished envelopes. The ranges of R_1 and R_2 were set based on the thermal resistance of the identical envelope assumption as well as the average convective heat transfer coefficient for exterior (0.043 m^2·K/W) and interior surfaces (0.11 m^2·K/W), respectively (MOLIT 2022). The resistances were assumed time-invariant, as the changes in heat transfer coefficients were negligible when combined with the resistance of the envelope.

Table 3 Time-invariant parameters

Symbol	Unit	Range
C_i	kJ/K	772 – 77,192
C_m	kJ/K	6,254 – 625,399
R_1	K/kW	0.095 – 9.476
R_2	K/kW	0.098 – 9.787
α	–	0 – 1
τ	–	0 – 1

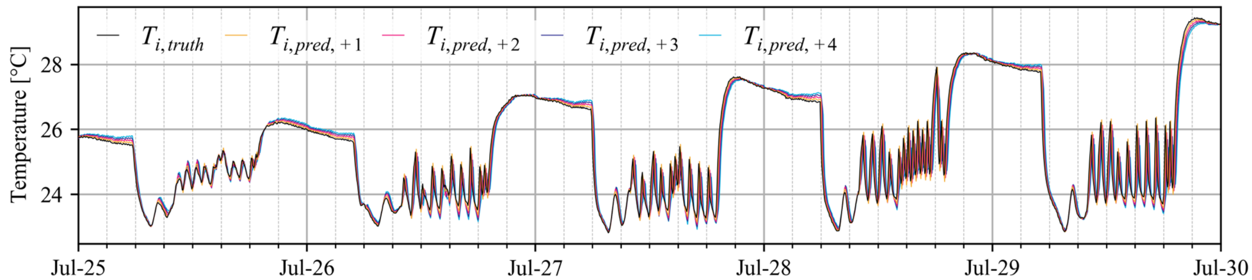

Figure 5 Room air temperature predictions

For this study, internal load (Q_{int}) was selected as the time-varying parameter to be predicted. The properties of the subnetwork responsible for this parameter estimation are listed in Table 4. The normalized output of the subnetwork is multiplied by the maximum internal load of 50 kW.

Table 4 Subnetwork properties

Property	Value
Input	*h* (time of day)
Output	Q_{int} (internal loads)
# of hidden layers	2
# of neurons in each hidden layer	20, 20
Activation function in hidden layers	SELU
Activation function in output layer	sigmoid

Results

Model training and parameter fitting

The fitted values of time-invariant parameters are listed in Table 5. In addition, the time-varying parameter Q_{int} relative to the time of day estimated by the subnetwork is visualized in Figure 4 (converted from kW to W/m^2).

Table 5 Fitted time-invariant parameters

Parameter	Unit	Fitted value
C_i	kJ/K	77,176
C_m	kJ/K	153,294
R_1	K/kW	5.087
R_2	K/kW	0.098
α	–	0.008
τ	–	0.007

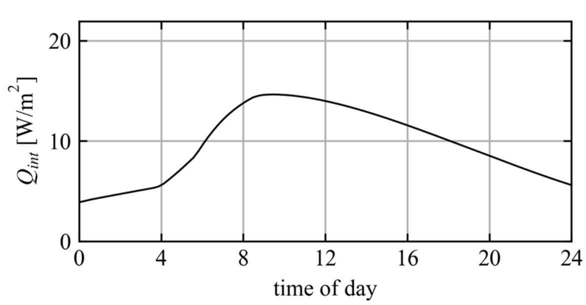

Figure 4 Fitted time-varying parameter (Q_{int})

Room air temperature prediction

In order to evaluate the prediction performance of the trained hybrid model, the model was used to forecast the room temperature for the test period (July 25th – 30th). For each timestep, four consecutive matrix calculations were carried out to obtain the future room temperatures with the prediction horizon of four timesteps (20 minutes) (Figure 5). R^2 score, root mean squared error (RMSE), and mean absolute error (MAE) were used as the accuracy metric for the predictions. The comparisons between measured ($T_{i,truth}$) and predicted room air temperatures are shown in Figure 6, and accuracies for each horizon are listed in Table 6. It was found that the model showed acceptable prediction accuracies for all horizons with all R^2 scores being above 0.9, and RMSE and MAE values under 1°C.

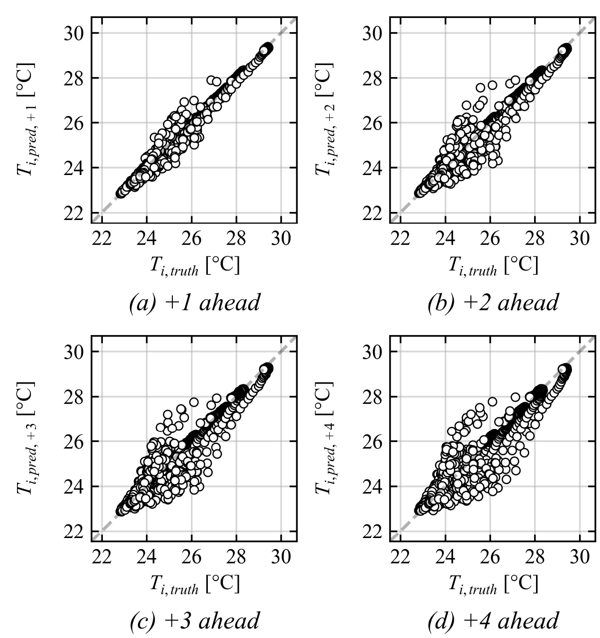

Figure 6 Measured vs. predicted temperatures

Table 6 Room air temperature prediction accuracies

Metric	+1	+2	+3	+4
R^2	0.98	0.95	0.92	0.90
RMSE [°C]	0.22	0.38	0.48	0.54
MAE [°C]	0.12	0.21	0.29	0.34

Conclusion

In this paper, a novel hybrid modeling approach was proposed to combine an RC model and neural networks. The building is represented as a minimalistic RC network, and the neural network is used to predict various time-invariant and time-varying parameters required for the state-space matrix calculation obtained from the RC model. To enable simultaneous fitting of these parameters, as well as incorporate warmup periods and multi-timestep forecast, Siamese neural networks were utilized as part of the model structure. The advantages and significance of the proposed hybrid modeling approach can be summarized as:

1. A simple RC model combined with the Siamese neural network is found to be good enough to predict room air temperature because it can combine the advantages of physics-based and data-driven approaches.
2. The model can be trained with both time-invariant and varying parameters present.
3. The model structure is flexible and easy to modify based on the desired RC model form, subnetwork structures, n_{total}, and so on.
4. The combined structure of the Siamese neural network and the matrix calculation layer (Figure 1) allows for the training of the model with a limited number of measured states.
5. The Siamese neural network and the matrix calculation layer can later be extracted and used separately for T_i prediction independent of n_{total}.

The authors observed that the available ranges of parameters had a significant impact on the resulting accuracy. In the absence of a definitive way of selecting these ranges, in this study, the authors found that updating the ranges heuristically based on the fitted results was enough for acceptable accuracies. A more robust, objective method for range selection should be further explored.

In addition, loss function design, specifically weight selection for the two losses was also crucial for not only prediction accuracy but also how well the causalities between inputs and outputs were reflected. As a result, poor weight selection often resulted in unstable, failed training of the model, as illustrated in Figure 7.

Another noteworthy characteristic of this methodology is that we *purposefully* introduce several of the *fictitious* state variables and their relevant fictitious properties. In other words, some of the estimated variables as well as parameters might not have physical meaning.

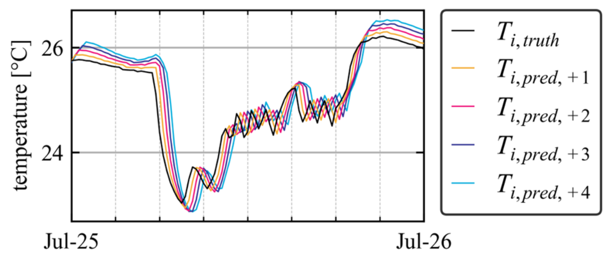

Figure 7 Example of a failed training

In future studies, the presented methodology will be further developed to be more robust and generalizable, and the aforementioned causality-based validation will be addressed. The modeling approach will be applied to other buildings with different scales, purposes, operations, etc. In addition, various forms of RC models and subnetworks will be explored. Also, more reliable methods for selecting parameter ranges as well as the loss weights will be addressed. Finally, the proposed modeling approach will be utilized in applications such as MPC and provide reliable predictions for optimal and rational decision-making.

Acknowledgment

This work was supported by the Korea Institute of Energy Technology Evaluation and Planning(KETEP) and the Ministry of Trade, Industry & Energy(MOTIE) of the Republic of Korea (No. 20202020800360).

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., et al. 2015. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems (Version 2.4.0), https://www.tensorflow.org/.

Ahn, K.U., Park, C.S., Kim, K.J., Kim, D.W., and Chae, C.U. 2022. Hybrid model using Bayesian neural network for variable refrigerant flow system, Journal of Building Performance Simulation 15(1), 1–20.

ASHRAE. 2021. ASHRAE Handbook – Fundamentals 2021, American Society of Heating, Refrigerating and Air-Conditioning Engineers.

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

Gokhale, G., Claessens, B., and Develder, C. 2022. Physics informed neural networks for control oriented thermal modeling of buildings, Applied Energy 314, 118852.

Kim, D.W., and Park, C.S. 2017. Application of Kalman Filter for Estimating a Process Disturbance in a Building Space, Sustainability 9, 1868.

Klanatsky, P., Veynandt, F., and Heschl, C. 2023. Grey-box model for model predictive control of buildings, Energy and Buildings 300, 113624.

Koch, G., Zemel, R., and Salakhutdinov, R. 2015. Siamese neural networks for one-shot image recognition, Proceedings of the 32nd International Conference on Machine Learning, July 6–11, Lille, France.

Li, Y., O'Neill, Z., Zhang, L., Chen, J., Im, P., and DeGraw, J. 2021. Grey-box modeling and application for building energy simulations – A critical review, Renewable and Sustainable Energy Reviews 146, 111174.

Manfren, M., James, P.A., and Tronchin, L. 2022. Data-driven building energy modelling – An analysis of the potential for generalisation through interpretable machine learning, Renewable and Sustainable Energy Reviews 167, 112686.

Maxwell, E.L. 1987. A Quasi-Physical Model for Converting Hourly Global Horizontal to Direct Normal Insolation, Technical Report No. SERI/TR-215-3087, Golden, CO: Solar Energy Research Institute.

MOLIT. 2022. Green Buildings Construction Support Act. Ministry of Land, Infrastructure and Transport.

Wang, Z., Chen, Y., and Li, Y. 2019. Development of RC model for thermal dynamic analysis of buildings through model structure simplification, Energy and Buildings 195, 51–67.