BESTOpt: A Physics-Informed Neural Network Based Building Simulation, Control and Optimization Platform — A Case Study on Dynamic Model Evaluation

Anonymous Authors[1]

Abstract

This paper presents **BESTOpt**, a modular simulation platform for building energy systems modeling. Unlike traditional tools that treat building dynamics, HVAC systems, and grid interactions in isolation, BESTOpt provides an integrated framework for dynamic modeling and control co-optimization. At its core is a physics-informed modularized neural network (PI-ModNN) that incorporates state-space-informed structural priors and hard physical constraints, enabling accurate, interpretable, and generalizable predictions of space air temperature. We evaluate the dynamic modeling module of BESTOpt against a purely data-driven baseline (LSTM) using synthetic datasets generated from EnergyPlus. While LSTM achieves lower prediction errors under normal conditions, BESTOpt demonstrates superior generalization in abnormal scenarios such as HVAC shutdowns, highlighting its effectiveness in control-oriented tasks where response fidelity is critical. The platform enables integrated building-to-grid applications at a large scale.

1. Introduction

Buildings account for 30% of global final energy consumption and 27% of global energy-related emissions (IEA, 2022). Among various building energy consumers, Heating, Ventilation, and Air-Conditioning (HVAC) systems account for more than half of the used energy (File, 2015). However, 40% of this energy is wasted due to inappropriate HVAC control, mismatched operation schedules, and other inefficiencies (Meyers et al., 2010). Therefore, improving the energy efficiency of buildings has become a key strategy for mitigating climate change and achieving decarbonization targets.

In addition to energy efficiency, buildings are increasingly expected to support grid flexibility—the ability to shift or regulate energy demand to balance power supply and demand (Luo et al., 2022). With the growing penetration of renewable energy sources such as solar and wind, the mismatch between distributed energy resource (DER) generation and load demand can result in unstable photovoltaic (PV) power exports to the utility grid, leading to grid instability and voltage violations (Gandhi et al., 2020). Therefore, grid-interactive buildings that can dynamically adapt their energy usage patterns are essential for maintaining grid stability and reducing peak loads.

Simultaneously, the increasing frequency of extreme weather events and weather-related grid disturbances highlights the importance of thermal resilience—the capability of buildings to maintain safe indoor temperatures during power outages or equipment failures (Liu et al., 2023). Enhancing thermal resilience is critical for protecting occupant health and safety, especially in vulnerable populations.

Effectively addressing the above-mentioned challenges of energy efficiency, grid flexibility, and thermal resilience requires a holistic approach to modeling and controlling both building energy systems and DERs. However, current tools often treat these aspects in isolation, lacking an integrated platform to simulate system-level interactions and evaluate performance trade-offs under various scenarios.

1.1. Literature Review of Current Simulation Platforms

At the energy system level, a commonly used tool is **Modelica** (Mattsson et al., 1998), which is built upon a set of physics-based equations and provides high-fidelity models suitable for system-level energy modeling, control, and optimization.

At the building level, widely adopted tools include **EnergyPlus** (Crawley et al., 2001) and its derivatives such as **DesignBuilder** (Garg et al., 2020) and **OpenStudio** (Guglielmetti et al., 2011). These are also physics-based models

[1]Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute.

grounded in energy balance equations. However, such models often suffer from: 1) Extensive data requirements, 2) Significant modeling effort for individual buildings, and 3) Substantial computational burden.

On the other hand, data-driven building dynamic models typically face challenges such as: 1) sensitivity to data quality and quantity, 2) lack of interoperability, 3) limited generalization capabilities, and 4) absence of guarantees for physical consistency.

At the grid level, **CityLearn** (Vázquez-Canteli et al., 2019) is a well-developed tool that supports grid-level control and optimization. However, it relies on pre-calculated building dynamics from EnergyPlus, which are based on fixed schedules and do not dynamically interact with real-time building or HVAC system responses.

For urban-scale modeling, **CityBES** (Hong et al., 2016) is commonly used, but it is primarily a physics-based, design-oriented tool and is not designed for control optimization or integration with city-level energy systems.

In summary, there is a lack of simulation platforms that enable co-simulation and optimization across multiple components—namely buildings, HVAC systems, and grid components in an integrated and dynamic manner.

To fill this gap, this paper presents a simulation platform **BESTOpt** that integrates modular building dynamic models with DER components such as PV and batteries. This platform supports control optimization and forward emulation of energy performance, flexibility, and resilience across diverse building types and weather conditions. As a case study, we evaluate the building dynamic model performance using synthetic datasets generated by EnergyPlus software.

2. Methodology

2.1. Data Generation

We conduct our simulation in EnergyPlus based on a single-family prototype building (Mendon & Taylor, 2014) developed by Pacific Northwest National Laboratory. We use the EnergyPlus Runtime API to collect data from the space and HVAC system. This API allows a client to interface with EnergyPlus at runtime, enabling data sensing and actuation during a running simulation.

At each time step (15 minutes), we collect data on outdoor air temperature, solar radiation, occupancy level, HVAC power, and space air temperature via the EnergyPlus variable handle. We then send control signals via the actuator handle to adjust the supply air flow rate according to the control policy. Detailed simulation settings are summarized in Table 1.

Table 1. Simulation settings for EnergyPlus virtual testbed

Parameter	**Value**
Weather Condition	Denver, Climate 5, Cool Dry
Timestep	15 minutes
Cooling Setpoint (occupied)	U(22°C, 25°C)
Cooling Setpoint (unoccupied)	U(24°C, 31°C)
Heating Setpoint (occupied)	U(18°C, 21°C)
Heating Setpoint (unoccupied)	U(13°C, 16°C)
Baseline	On-Off control with deadband
Deadband	0.5°C
Depart Time	U(7:00, 10:00)
Arrive Time	U(16:00, 20:00)
Supply Air Temperature	13°C
Supply Air Flow Rate	0 to 0.16 m^3/s
Simulation Period	Whole year simulation
Peak Hour	15:00 to 18:00

2.2. Dataset and Baseline

The baseline used in this case study is the Long Short-Term Memory (LSTM) neural network. Four datasets are used to evaluate model performance. The first three datasets are generated using different sampled setpoints and occupancy schedules based on Table 1, to test the model under normal operating conditions. The fourth dataset simulates an HVAC shutdown after July 30th to evaluate the generalization ability of the proposed model.

2.3. Physics-Informed Modularized Neural Network

The thermal dynamic model in this study is based on a Physics-informed Modularized Neural Network (PI-ModNN) (Jiang & Dong, 2024; Jiang et al., 2025), which estimates space air temperature by modeling each heat transfer component through dedicated neural network modules as shown in Figure 1. This is a multi-step time stepper model structured around a discretized state-space formulation. The network takes in ambient temperature, solar radiation, time features, HVAC control inputs, and historical zone air temperature to predict future space air temperature over a specified horizon. Each module estimates a specific heat transfer component: the external module uses a recurrent neural network to capture envelope thermal inertia and solar gains; the internal gain module employs a multi-layer perceptron (MLP) driven by time-based features and occupancy data; the HVAC module maps control signals to

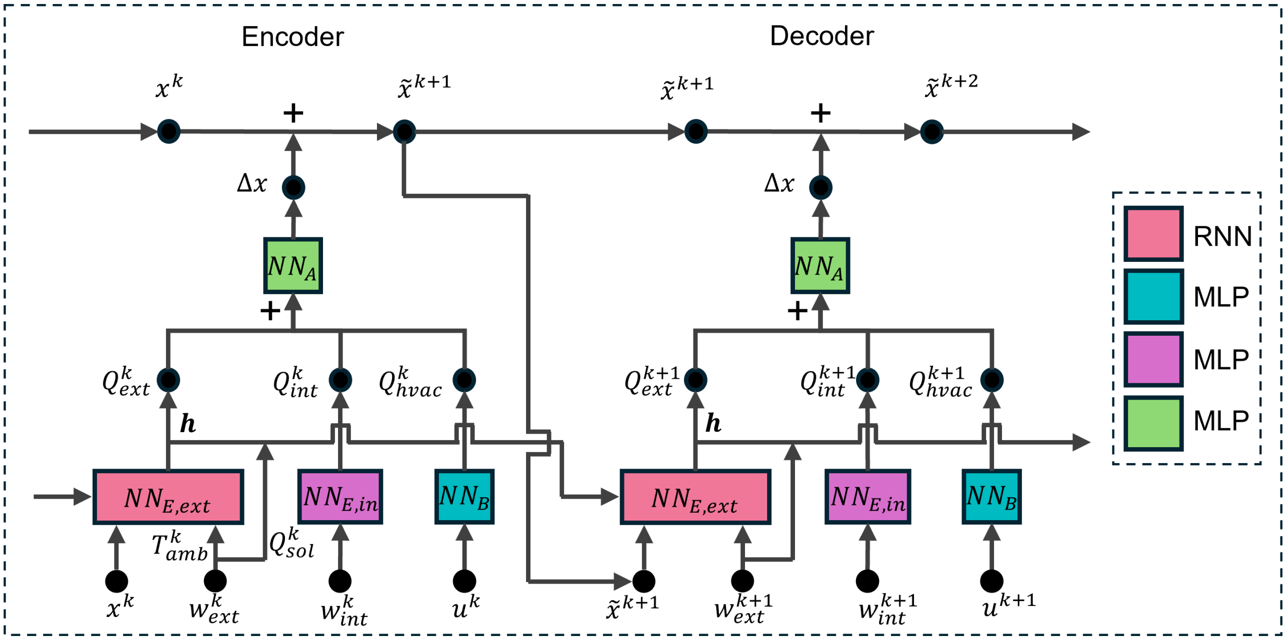

Figure 1. Physics-Informed Building Dynamic Model used in BESTOpt

heat input using either MLP or RNN-based architectures; and a residual module learns the unknown thermal capacity and integrates temperature dynamics over time. These components together support both accuracy and physical interpretability.

To improve the model's responsiveness to HVAC-driven fluctuations and mitigate the common over-smoothing issue in neural networks, we adopt a physics-informed loss function that penalizes discrepancies in both absolute temperature and per-step fluctuation. This approach encourages the model to better capture rapid changes without sacrificing long-term trend accuracy. The proposed loss combines a traditional mean-squared error term with a fluctuation-based penalty weighted by a tunable coefficient.

In addition, we impose physics-inspired constraints to ensure that model responses follow the physical principle that increased cooling or heating should lead to monotonic decreases or increases in temperature, respectively. This is achieved by enforcing positivity on the gradients of zone temperature with respect to heat inputs, which can be satisfied by constraining network weights to be positive and using activation functions like ReLU. These constraints are especially applied to the HVAC and envelope modules to maintain consistency with the heat balance equation.

2.4. Model Training and Validation

The model training process begins with data cleaning, if the space air temperature remains unchanged for four consecutive hours, that period is considered missing. To preserve temporal continuity, training data loaders are created within each clean segment. To enhance generalization across different prediction intervals, multi-horizon forecasting (e.g., 2, 4, 8 to 24 hours) is incorporated into training batches. Two techniques are used to improve training: early stopping, which stops training when validation loss plateaus to prevent overfitting, and a mixed-data training strategy, where the encoder is trained using both ground truth and predicted temperatures to better learn the dynamic module. Hyperparameter tuning is performed using Optuna, with learning rates, hidden dimensions, and sequence lengths optimized. The final configuration for each case is selected based on validation performance and used for evaluation.

3. Results and Discussion

3.1. Model Performance under Normal Condition

Table 2. Model performance (MAE in °C) over one-month evaluation across different datasets.

Model	Dataset 1	Dataset 2	Dataset 3
LSTM	0.19	0.11	0.14
BESTOpt	0.33	0.26	0.28

Model performance under normal operating conditions is compared in Table 2. The LSTM model outperforms the physics-informed BESTOpt model across all three datasets, achieving mean absolute errors (MAEs) of 0.19, 0.11, and

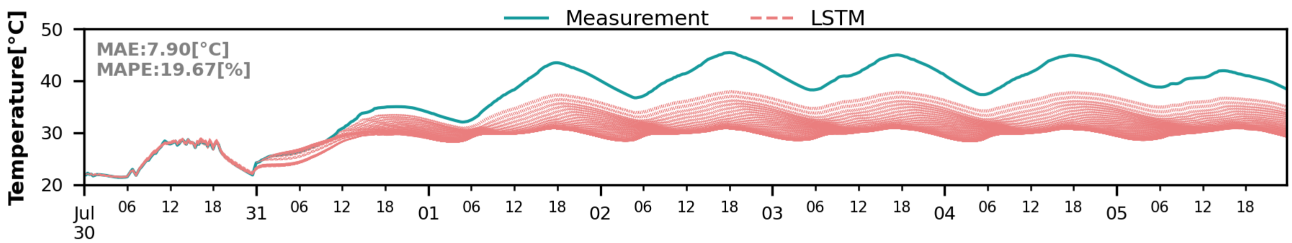

Figure 2. LSTM under unseen condition.

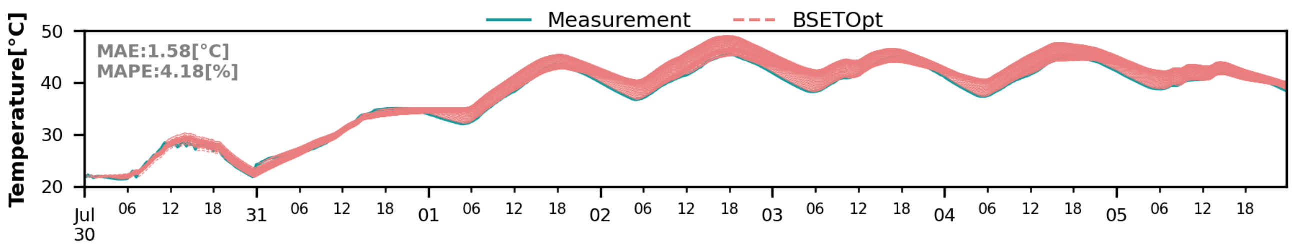

Figure 3. BESTOpt under unseen condition.

0.14 °C compared to 0.33, 0.26, and 0.28 °C for BESTOpt. One possible explanation is that the datasets were generated using EnergyPlus, where certain heat gains—such as internal heat gains—are driven by the occupancy schedule, which is a predefined constant value and follows repetitive temporal patterns. As a result, the LSTM model can effectively utilize time-related features (e.g., hour of day) to learn these patterns and produce accurate forecasts. In contrast, BESTOpt integrates physics-based constraints, which can enhance generalizability and interpretability under abnormal conditions, but may also restrict the model's flexibility during training.

3.2. Model Performance under Unseen Conditions

To evaluate generalization ability of proposed model, we tested its performance under unseen conditions, as illustrated in Figure 2 and Figure 3. In these figures, the green line represents the ground truth, while the red lines show the predicted space air temperature over a 24-hour horizon with 15-minute intervals. Both models perform well on July 30, when the system operates under normal conditions. However, after the air conditioning system is shut off on July 31, the indoor temperature continues to rise due to summer heat. The LSTM baseline fails to capture this behavior, whereas the physics-informed dynamic model in BESTopt successfully reflects the thermal dynamics and yields significantly lower prediction errors. This result highlights the advantage of embedding physical priors into data-driven models, particularly for safety-critical scenarios and abnormal operations where training data is limited or unavailable.

4. Conclusion

This study proposes a building dynamic simulation platform, BESTOpt, which leverages physics-informed machine learning to enhance model generalization compared to traditional purely data-driven approaches. BESTOpt can serve as a virtual testbed for advanced controller design and enables integration across building, HVAC, and grid-connected DERs, paving the way for future multi-component, complex building energy system control and optimization.

Software and Data

Acknowledgments

Impact Statement

This study highlights the transformative potential of embedding physical knowledge into data-driven models to enhance their generalizability, interpretability, and robustness. By enabling seamless integration with modular system components, the proposed approach offers a scalable, hierarchical framework that supports coordinated interaction, control, and optimization across diverse layers of future building-grid energy systems. This work paves the way for more intelligent, resilient, and efficient energy management strategies in complex cyber-physical infrastructures.

References

Crawley, D. B., Lawrie, L. K., Winkelmann, F. C., Buhl, W. F., Huang, Y. J., Pedersen, C. O., et al. Energyplus:

creating a new-generation building energy simulation program. *Energy and Buildings*, 33(4):319–331, 2001.

File, M. Commercial buildings energy consumption survey (cbecs). Technical report, U.S. Department of Energy, Washington, DC, USA, 2015.

Gandhi, O., Kumar, D. S., Rodríguez-Gallegos, C. D., and Srinivasan, D. Review of power system impacts at high pv penetration part i: Factors limiting pv penetration. *Solar Energy*, 210:181–201, 2020.

Garg, V., Mathur, J., and Bhatia, A. *Building energy simulation: A workbook using DesignBuilder™*. CRC Press, 2020.

Guglielmetti, R., Macumber, D., and Long, N. Openstudio: an open source integrated analysis platform. Technical report, National Renewable Energy Laboratory (NREL), Golden, CO, 2011.

Hong, T., Chen, Y., Lee, S. H., and Piette, M. A. Citybes: A web-based platform to support city-scale building energy efficiency. *Urban Computing*, 14:2016, 2016.

IEA. Iea tracking report for buildings. `https://globalabc.org/index.php/resources/publications/iea-tracking-report-buildings`, 2022. Accessed: 2025-06-10.

Jiang, Z. and Dong, B. Modularized neural network incorporating physical priors for future building energy modeling. *Patterns*, 5(8), 2024.

Jiang, Z., Wang, X., and Dong, B. Physics-informed modularized neural network for advanced building control by deep reinforcement learning. *arXiv preprint arXiv:2504.05397*, 2025.

Liu, Z., Chen, Y., Yang, X., and Yan, J. Power to heat: Opportunity of flexibility services provided by building energy systems. *Advances in Applied Energy*, 11:100149, 2023.

Luo, Z., Peng, J., Cao, J., Yin, R., Zou, B., Tan, Y., and Yan, J. Demand flexibility of residential buildings: definitions, flexible loads, and quantification methods. *Engineering*, 16:123–140, 2022.

Mattsson, S. E., Elmqvist, H., and Otter, M. Physical system modeling with modelica. *Control Engineering Practice*, 6(4):501–510, 1998.

Mendon, V. V. and Taylor, Z. T. Development of residential prototype building models and analysis system for large-scale energy efficiency studies using energyplus. Technical report, Pacific Northwest National Laboratory (PNNL), Richland, WA, United States, 2014.

Meyers, R. J., Williams, E. D., and Matthews, H. S. Scoping the potential of monitoring and control technologies to reduce energy use in homes. *Energy and Buildings*, 42 (5):563–569, 2010.

Vázquez-Canteli, J. R., Kämpf, J., Henze, G., and Nagy, Z. Citylearn v1.0: An openai gym environment for demand response with deep reinforcement learning. In *Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*, pp. 356–357, 2019.