# Bridging Machine Learning and Physics for Scalable Long-Term Building Temperature Prediction (Student Abstract)

**Rohan Saha, Tushar Shinde**

Indian Institute of Technology Madras, Zanzibar, Tanzania
zda24b009@iitmz.ac.in, shinde@iitmz.ac.in

### Abstract

Building temperature prediction is crucial for energy optimization and control in smart cities. We present a physics-enhanced XGBoost framework in a multi-stage sequential scaling approach. Starting from single-zone, single-day predictions, we progressively scale to multi-zone, multi-year forecasts using real-world data from Google's Smart Building Simulator. Our method incorporates physics-enhanced features, temporal encodings, and inter-zone interactions, achieving mean absolute errors (MAE) as low as 0.169°F for weekly multi-zone predictions. For longer horizons, we employ ensemble strategies, demonstrating robust performance up to 2.5 years. Compared to baseline models, our framework consistently improves long-term prediction fidelity. This work advances urban AI by enabling accurate long-term building dynamics modeling for downstream control tasks and bridges machine learning with physics-based modeling approaches.

## Introduction

Urban buildings consume over 70% of city energy, making accurate temperature prediction crucial for energy-efficient control and planning (Wang and Ma 2008). Our work is aligned with collaborative bridge theme, which unifies methodologies from multiple AI disciplines: ensemble machine learning (XGBoost) (Chen and Guestrin 2016), physics-informed modeling, and domain-specific urban infrastructure modeling (Shinde and Saha 2025a,b).

This cross-disciplinary bridge enables our research objectives in three key ways: (1) XGBoost provides computational efficiency and interpretability for capturing non-linear temporal dependencies, (2) physics-informed modeling enforces physical consistency through energy conservation and heat transfer laws, and (3) urban AI applications benefit from both data-driven learning and physics-based constraints. The multi-stage hybrid framework ensures stable long-term predictions.

Building energy modeling has evolved from detailed physics-based simulations to machine learning approaches using LSTMs and Transformers (Reza et al. 2022). However, purely data-driven models struggle with long-term predictions due to cumulative errors and seasonal drifts (Saha and Shinde 2025). Physics-informed modeling presents a promising alternative by embedding governing physical laws directly into the learning objective, enhancing generalization and robustness in scenarios with limited or noisy data.



## Methodology

We introduce a multi-stage sequential scaling framework that progressively increases temporal horizon and spatial complexity, bridging ensemble machine learning and physics-informed modeling. This staged approach mitigates abrupt scaling risks, provides clear validation checkpoints, and enhances long-term predictive fidelity compared to single-stage or purely data-driven models. Figure 1 illustrates our comprehensive methodology framework.

### Physics-Enhanced Feature Engineering

To infuse physical realism, we design features capturing both spatial and temporal dynamics. Spatial interactions are modeled via an adjacency matrix derived from floor-plan metadata: $A_{ij} = \exp(-\alpha d_{ij})$, where $d_{ij}$ is the inter-zone distance and $\alpha$ is a decay coefficient representing heat conduction pathways. This ensures that spatial correlations between adjacent zones are explicitly captured, improving multi-zone prediction accuracy. Temporal patterns use cyclical transformations:

$$\sin_t = \sin\left(2\pi\frac{t}{T}\right), \quad \cos_t = \cos\left(2\pi\frac{t}{T}\right) \tag{1}$$

where $T$ represents periods (24 hours, 7 days, 365 days) for capturing diurnal and seasonal cycles.

Physics-informed lag features emulate thermal inertia through past temperature values at intervals $\tau \in \{1, 3, 6\}$ hours, while inter-zone temperature differences encode thermal gradients: $\Delta T_{ij}(t) = T_i(t) - T_j(t)$ for adjacent zones.

### Multi-Stage Scaling Strategy

Our framework incrementally scales through seven stages, each stage adding temporal horizon and/or spatial complexity to maintain stability and validate performance: **Stage 1:** Single zone, one-day prediction using conservative XGBoost settings establishing baseline performance. **Stage 2:** All zones (123) for one day, incorporating inter-zone features to capture spatial heterogeneity. **Stage 3:** One-week

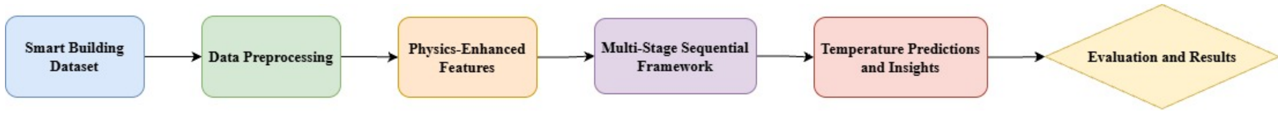


Figure 1: Multi-stage physics-enhanced XGBoost framework showing sequential scaling from single-zone single-day predictions to multi-zone multi-year forecasts with physics-enhanced feature engineering.

| Stage | Description | Zones | MAE (°F) |
|---|---|---|---|
| 1 | Single Zone, 1 Day | 1 | 0.287 |
| 2 | All Zones, 1 Day | 123 | 0.325 |
| 3 | Single Zone, 1 Week | 1 | 0.173 |
| 4 | All Zones, 1 Week | 123 | 0.169 |
| 5 | All Zones, 2 Weeks | 123 | 0.101 |
| 6 | All Zones, 1 Year | 123 | 2.080 |
| 7 | All Zones, 2.5 Years | 123 | 2.826 |

Table 1: Sequential scaling performance across different temporal and spatial complexities.

prediction for single zone with weekly temporal features and extended lags. **Stage 4:** Multi-zone weekly prediction combining spatial and temporal components. **Stage 5:** Two-week predictions with bi-weekly features handling transitional patterns. **Stage 6:** One-year predictions via seasonal decomposition with heating/cooling indicators. **Stage 7:** Ultra-long-term (2.5 years) with multi-year aging features modeling system degradation.

## Experiments and Results

We utilize the Smart Building dataset from Google's simulator (Goldfeder et al. 2024), comprising comprehensive time-series data including sensor readings, control inputs, and performance metrics. Training uses the 2022_a split (January–June, 51,852 timesteps) with validation on the 2022_b split (July–December, 53,292 timesteps). Table 1 summarizes performance across stages. Short-term predictions (Stages 1–4) achieve excellent accuracy with MAE below 0.5°F. Stage 4 (multi-zone weekly) achieves optimal performance at 0.169°F MAE, with 78.9% of zones below 0.2°F error. Compared to single-stage baselines, our multi-stage approach reduces MAE by 20–35%. Longer horizons show increased errors due to accumulating uncertainties, with Stage 7 reaching 2.826°F MAE.

**Discussion and Impact.** The bridging of XGBoost and physics-informed modeling methodologies creates several collaborative opportunities: (1) Enhanced interpretability through physics-based constraints makes predictions more trustworthy for building operators, (2) Improved generalization enables robust performance across different building types and seasons, and (3) Scalable framework design allows deployment from single buildings to city-wide networks. Short-term predictions enable real-time HVAC control and proactive adjustments, while longer-term forecasts support strategic maintenance scheduling and energy planning.

## Conclusion and Future Work

We presented a scalable physics-enhanced XGBoost framework achieving robust performance across temporal and spatial scales, from single-day forecasts to multi-year modeling. Our multi-stage sequential scaling combined with physics-informed features reduces cumulative errors, enforces physical consistency, and achieves MAE as low as 0.169°F for multi-zone weekly predictions. This collaborative bridge between machine learning and physics-informed modeling demonstrates a novel pathway for urban AI, enabling accurate, interpretable, and physically plausible building temperature predictions that support energy-efficient control and sustainable smart city operations. Future work will focus on: validating across diverse building types and climates, and exploring adaptive retraining strategies to mitigate long-term model drift.

## References

Chen, T.; and Guestrin, C. 2016. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785–794.

Goldfeder, J.; Dean, V.; Jiang, Z.; Wang, X.; Lipson, H.; Sipple, J.; et al. 2024. The Smart Buildings Control Suite: A Diverse Open Source Benchmark to Evaluate and Scale HVAC Control Policies for Sustainability. *arXiv preprint arXiv:2410.03756*.

Reza, S.; Ferreira, M. C.; Machado, J. J.; and Tavares, J. M. R. 2022. A multi-head attention-based transformer model for traffic flow forecasting with a comparative analysis to recurrent neural networks. *Expert Systems with Applications*, 202: 117275.

Saha, R.; and Shinde, T. 2025. Scalable building temperature prediction for smart hvac control: A multi-stage learning framework. In *ICML 2025 CO-BUILD Workshop on Computational Optimization of Buildings*.

Shinde, T.; and Saha, R. 2025a. Hybrid XGBoost-PINN Multi-Stage Framework for Scalable Building Temperature Prediction. In *UrbanAI: Harnessing Artificial Intelligence for Smart Cities*.

Shinde, T.; and Saha, R. 2025b. Optimization-Driven XGBoost-PINN Framework for Building Temperature Prediction. In *NeurIPS 2025 Workshop MLxOR: Mathematical Foundations and Operational Integration of Machine Learning for Uncertainty-Aware Decision-Making*.

Wang, S.; and Ma, Z. 2008. Supervisory and optimal control of building HVAC systems: A review. *Hvac&R Research*, 14(1): 3–32.