E3S Web of Conferences **716**, 02038 (2026) https://doi.org/10.1051/e3sconf/202671602038
IAQVEC 2026

Physics-Embedded Transfer Learning Graph Neural Network for Building Energy Prediction

Yihui Li[1,2,3*], *Zhexuan Yu*[4], *Jun* Xiao[1,2], *Qiong* Wang[1], *Hao* Zhou[2,5] and *Borong* Lin[1,2**]

[1]School of Architecture, Tsinghua University, Beijing 100084, China
[2]Key Laboratory of Eco-planning & Green Building (Tsinghua University), Ministry of Education, Beijing 100084, China
[3]College of Environmental Design, University of California, Berkeley, CA 94720, USA
[4]Weiyang College, Tsinghua University, Beijing 100084, China
[5]Institute of Urban Governance and Sustainable Development, Tsinghua University, Beijing 100084, China

Abstract. Accurate and scalable building energy prediction in early design remains challenging due to heterogeneous spatial configurations, climate variability, and limited data. This study proposes a hybrid framework integrating transfer learning with physics-embedded graph neural networks (GNNs) to improve generalization and interpretability. Buildings are modeled as heterogeneous graphs, where spatial units are nodes with geometric and material attributes, and thermal interactions are edges. A transfer learning strategy adapts knowledge from data-rich source domains to data-scarce target cases, enhancing cross-building and cross-climate applicability. Fundamental heat transfer mechanisms—conduction, convection, and radiation—are embedded into the GNN message-passing process to enforce thermodynamic consistency and mitigate overfitting. Experiments on multi-climate datasets with diverse typologies show that the proposed method outperforms purely data-driven GNNs and conventional machine learning models, while requiring fewer target-domain samples. The framework also improves interpretability by linking learned representations with physical principles, supporting generalizable and data-efficient energy prediction for green building design.

Keywords. Graph neural network; Transfer learning; Physics-informed modeling; Building energy prediction; Heat transfer equations

1 Introduction

The building sector is a major contributor to global energy consumption and greenhouse gas emissions, accounting for approximately 34% of total energy use worldwide [1]. Accurate prediction of building energy consumption and thermal load is essential for optimizing energy management systems, reducing operational costs, and achieving carbon neutrality goals [2]. In the early stages of building design, internal disturbances such as occupant schedules and equipment usage are often highly uncertain. Consequently, the rapid and precise prediction of external disturbance loads—driven primarily by climatic factors—becomes critical for evaluating building performance and informing design decisions [3].

Traditionally, building energy prediction relies on physical simulation models (white-box models), such as EnergyPlus and TRNSYS. These models are grounded in fundamental thermodynamic and heat transfer principles, offering high interpretability and the ability to simulate complex physical processes. However, they require detailed building geometry and material properties, which are often unavailable in early design phases, and their high computational cost limits their application in real-time or large-scale scenarios [4]. Conversely, data-driven models (black-box models), particularly those based on machine learning and deep learning, have gained popularity due to their ability to learn complex patterns directly from historical data without explicit physical modeling [2]. Despite their success, pure data-driven methods often lack physical interpretability and suffer from poor generalization when faced with data-scarce environments or unseen building configurations [5].

To bridge the gap between interpretability and computational efficiency, hybrid approaches such as Physics-Informed Machine Learning have emerged. These methods integrate physical laws, typically in the form of partial differential equations, into the loss functions of neural networks to ensure that predictions remain physically consistent [4], [5], [6], [7]. Simultaneously, Graph Neural Networks (GNNs) have demonstrated significant potential in representing the complex spatial dependencies of buildings by abstracting zones and walls as nodes and their thermal interactions as

* Corresponding author: liyihui23@mails.tsinghua.edu.cn
** Corresponding author: linbr@tsinghua.edu.cn

edges [8], [9]. While some studies have explored GNNs for building energy modeling, there remains a lack of deep integration between GNN architectures and fundamental heat transfer equations, particularly for multi-scale buildings where both local heat exchange and global energy flow are significant [8].

Furthermore, transfer learning has been increasingly utilized to address the challenge of data scarcity in target buildings by leveraging knowledge from data-rich source domains [10], [11], [12]. Most existing transfer learning applications in this field are purely data-driven, focusing on feature alignment or parameter fine-tuning without considering the underlying physical rules that govern energy transfer across different climates [12]. This limitation often leads to "negative transfer" or reduced robustness when the source and target domains exhibit significantly different thermal behaviors [13].

To address these challenges, this study proposes a **Graph-based Physics-Embedded Transfer Learning (GPE-TL) framework** for predicting multi-zone building external disturbance loads. The framework introduces three key innovations:

- **Heterogeneous Graph Representation**: Buildings are represented as heterogeneous graphs where spatial units are encoded with geometric and material properties, and edges are defined by physical thermal interactions, allowing for flexible representation of complex multi-zone structures.
- **Physics-Embedded Message Passing**: Fundamental heat transfer equations (conduction, convection, and radiation) are integrated directly into the GNN's message-passing mechanism. This ensures that information propagation across the graph follows physical laws, enhancing both interpretability and data efficiency.
- **Cross-Climate Transfer Learning**: A physics-guided transfer learning scheme is developed to adapt knowledge learned from simulation-rich source domains to data-scarce target domains across different climate zones, significantly improving prediction robustness.

The remainder of this paper is organized as follows: Section 2 describes the proposed framework; Section 3 presents the experimental setup and data generation; Section 4 discusses the results and performance comparison; and Section 5 concludes the study.

2 Methodology

2.1 Graph structure definition and thermal process modeling

2.1.1 Definition of heterogeneous building graphs

Built on our previous work CUGER (Convex Unified Graph Encoding Representation)[9], each building can be represented as a heterogenous graph as:

$$G_{\text{building}} = (V_F \cup V_S, E_{FF} \cup E_{SF}, W_{SF}), \quad (1)$$

where surface nodes V_F : representing physical components such as floors, walls, windows, shading elements, and airwalls, modeled as polygonal surface elements; space nodes V_S : representing enclosed volumetric spaces bounded by the surfaces, such as rooms or thermal zones; face-face edges E_{FF}: undirected edges capturing physical connectivity between adjacent surfaces (e.g., wall-wall or floor-wall contact); space-face edges E_{SF}: directed edges from spatial nodes to their enclosing surface nodes, encoding boundary relationships. Each edge in E_{SF} is associated with a weight W_{SF}, which can represent surface area, orientation alignment, or heat transfer contribution.

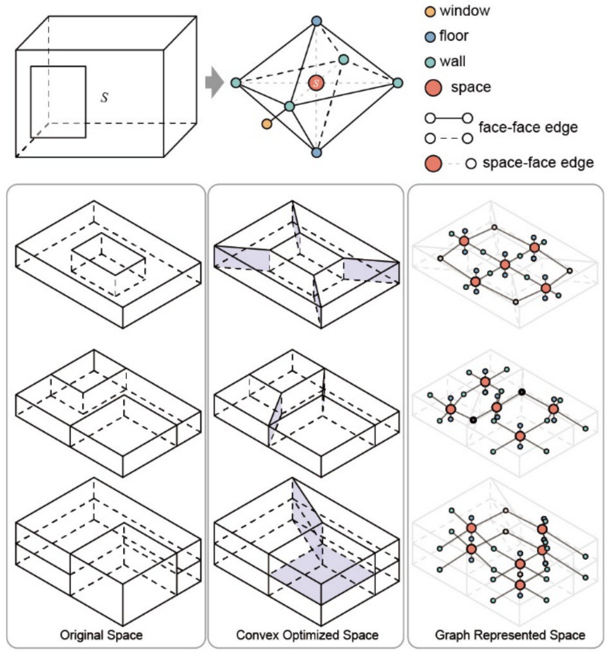

Fig. 1 CUGER graph representation of a building, illustrating simple room conversion, non-convex space subdivision via airwall nodes, and node–edge attribute support for graph-based simulation.

Airwall nodes are introduced as a special subclass of surface nodes to partition non-convex rooms into convex subspaces. This ensures that every space node corresponds to a convex region with no internal angle exceeding 180°, enabling consistent geometric reasoning and physically meaningful graph embeddings.

2.1.2 Physics embedding of external loads for graph modeling

We introduce an abstract environmental state function $\mathcal{S}(t)$ to represent the physical conditions affecting a space at time t, including air temperature, humidity, and radiative temperature. The indoor state of zone i is denoted as $\mathcal{S}_{\mathrm{int},i}(t)$, while the outdoor environment is represented by $\mathcal{S}_{\mathrm{ext}}(t)$. Energy exchange between environments is characterized by an operator $\mathcal{C}$, which abstracts thermal transmittance, ventilation effectiveness, and surface heat transfer properties. The major components of external disturbance heat gains are modeled as follows:

$$\dot{Q}_i(t) = \mathcal{C}\left[\mathcal{S}_{\mathrm{ext}}(t) - \mathcal{S}_{\mathrm{int},i}(t)\right] \tag{2}$$

In physical reality, energy transfer between indoor and outdoor environments occurs through envelope surfaces rather than directly at the space level. Therefore, we reformulate the above load model using a graph representation.

Each indoor space $\mathcal{S}_i(t)$ is connected to a set of envelope surface nodes $\{\mathcal{F}_i\}$, including external surfaces $\mathcal{F}_{ext}$ and internal partition surfaces $\mathcal{F}_{int}$. External disturbance heat gains can then be expressed as a surface-wise aggregation:

$$\dot{Q}_{\mathrm{ext}} = \sum_{j \in \mathcal{F}_{\mathrm{ext}}} \mathcal{F}_j \left(\mathcal{S}_{\mathrm{ext}} - \mathcal{S}_{\mathrm{int}}\right) + \sum_{k \in \mathcal{F}_{\mathrm{ext}}} \mathcal{F}_k \left(\mathcal{S}_{\mathrm{int},\,k} - \mathcal{S}_{\mathrm{int}}\right) \tag{3}$$

In the graph, the interaction between a surface node $\mathcal{F}$ and a space node $\mathcal{S}$ is encoded by an edge weight $E_{\mathcal{F},\mathcal{S}}$, representing the effective thermal coupling. Under a linear approximation, the net external heat gain can be simplified as

$$\dot{Q}_{\mathrm{ext}} \propto \mathcal{S}_{\mathrm{int}} \sum_i \mathcal{F}_j E_{\mathcal{F}_i,\mathcal{S}_{\mathrm{int}}} \left(1 + \sigma_{\mathrm{ext},i}\right) \tag{4}$$

where $\sigma_{\mathrm{ext},i}$ is a binary or continuous mask indicating whether surface $\mathcal{F}_i$ is exposed to the outdoor environment:

$$\sigma_{\mathrm{ext},i} = \begin{cases} \sigma_i, & \mathcal{F}_i \in \mathcal{F}_{\mathrm{ext}} \\ 0, & \mathcal{F}_i \in \mathcal{F}_{\mathrm{int}} \end{cases} \tag{5}$$

This formulation unifies external and internal heat-transfer effects into a single graph aggregation process. Outdoor climatic influence is injected through masked surface–space edges, while internal surfaces preserve their original coupling weights. As a result, physical heat-transfer mechanisms are naturally embedded into the GNN message-passing scheme, providing a clear and interpretable bridge between thermal physics and data-driven graph learning. **Fig. 2** illustrates a case study of an architectural drawing to demonstrate the influence of outdoor nodes on spatial nodes.

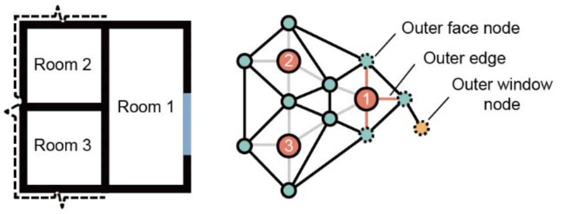

Fig. 2 A schematic diagram of the building plan and its corresponding graph structural slices. Part of Room 1's surface directly contacts the external environment, thus forming heterogeneous graph information such as outer face node and outer edge.

2.2 Graph neural network architecture

The proposed load prediction model embeds the physical graph structure of a building into a graph neural network (GNN). The building is represented as a multi-zone heterogeneous graph consisting of two types of nodes: space nodes $\mathcal{S}$ and envelope surface nodes $\mathcal{F}$. Each space node corresponds to a room or thermal zone, while each surface node represents an envelope element such as a wall, window, floor, or ceiling.

The learning objective is node-level prediction on space nodes, i.e., estimating the energy response $\dot{Q}_{\mathcal{S}}$ for each space. The GNN integrates geometric topology, thermal coupling, and external climatic disturbances through structured message passing, as described below.

2.2.1 Weather feature embedding and external injection

Outdoor meteorological conditions, including temperature, humidity, wind speed, and solar radiation, are first embedded into a latent feature vector using a multilayer perceptron (MLP):

$$x_{\mathrm{weather}} = \mathrm{MLP}(\mathcal{S}_{\mathrm{ext}}) \tag{6}$$

Rather than directly attaching weather features to space nodes, the external information is injected through surface–space edges, reflecting the physical fact that outdoor disturbances affect indoor spaces through envelope elements. For an edge from surface node $\mathcal{F}_j$ to space node $\mathcal{S}_i$, the edge feature is defined as

$$e_{\mathcal{F}_j \to \mathcal{S}_i} = \left[E_{\mathcal{F}_i,\mathcal{S}_i},\, x_{\mathrm{weather}}\right], \tag{7}$$

where $E_{\mathcal{F}_i,\mathcal{S}_i}$ represents the effective thermal coupling between the surface and the space.

To distinguish external and internal envelope elements, a binary or continuous mask $\sigma_{\mathrm{ext},j}$ is applied:

$$e_{\mathcal{F}_j \to \mathcal{S}_i} = \begin{cases} \left[E_{\mathcal{F}_i,\mathcal{S}_i},\, x_{\mathrm{weather}}\right], & \mathcal{F}_i \in \mathcal{F}_{\mathrm{ext}} \\ \left[E_{\mathcal{F}_i,\mathcal{S}_i},\, 0\right], & \mathcal{F}_i \in \mathcal{F}_{\mathrm{int}} \end{cases} \tag{8}$$

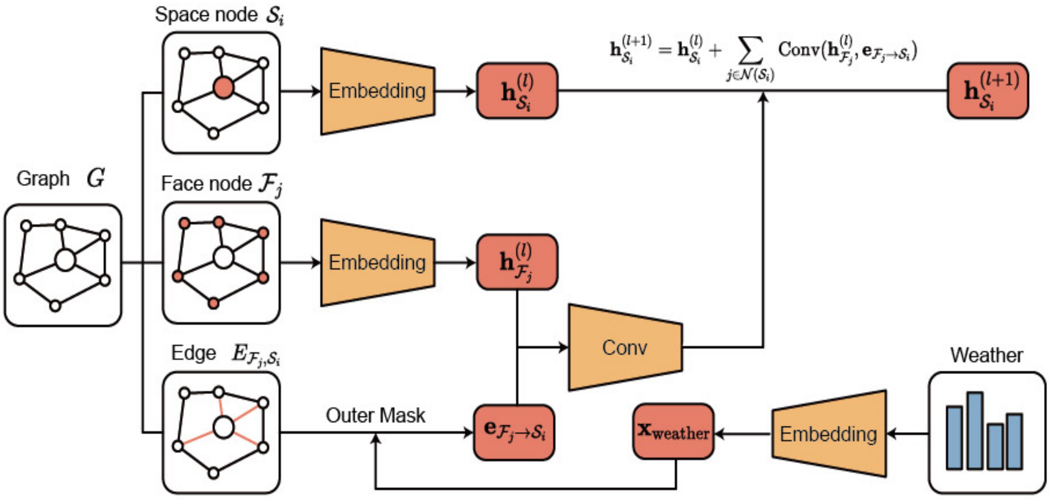

Fig. 3 GNN encoding architecture for heterogeneous data

This design ensures that outdoor climate information is only injected through physically exposed surfaces. **Fig. 3** shows the architecture for the embedding processing.

2.2.2 Space–surface message passing

Graph convolution is performed exclusively between space nodes and their adjacent surface nodes. Let $\mathcal{N}(\mathcal{S}_i)$ denote the set of envelope surfaces connected to space $\mathcal{S}_i$. At layer l, the space node representation is updated as

$$\boldsymbol{h}_{\mathcal{S}_i}^{(l+1)} = \boldsymbol{h}_{\mathcal{S}_i}^{(l)} + \sum_{j\in\mathcal{N}(\mathcal{S}_i)} Conv\left[\boldsymbol{h}_{\mathcal{F}_j}^{(l)}, \boldsymbol{e}_{\mathcal{F}_j\to\mathcal{S}_i}\right], \tag{9}$$

where $\boldsymbol{h}_{\mathcal{S}_i}^{(l)}$ and $\boldsymbol{h}_{\mathcal{F}_j}^{(l)}$ are the latent features of space and surface nodes, respectively.

The convolution operator $\mathrm{Conv}(\cdot)$ is implemented as a learnable weighted aggregation function (e.g., linear transformation), allowing the model to learn the relative contribution of different envelope elements to the space-level thermal response. The residual connection preserves stability and reflects energy accumulation over time.

2.2.3 Multi-layer aggregation and load prediction

By stacking multiple space–surface convolution layers, each space node progressively aggregates information from its surrounding envelope elements and injected outdoor disturbances. After L convolution layers, the final space node representation encodes both local physical properties and external climatic influence.

The space-level thermal load is then predicted using an output MLP:

$$\dot{Q}_{\mathcal{S}_i} = \mathrm{MLP}\left[\boldsymbol{h}_{\mathcal{S}_i}^{(L)}\right] \tag{10}$$

Through this architecture, the GNN jointly captures graph topology, thermal coupling, and external disturbance injection in a physically consistent manner. The resulting model provides a unified framework that bridges physics-based heat transfer modeling and data-driven graph learning for building energy load prediction.

2.3 Graph-based transfer learning framework

Transfer learning aims to improve model performance in data-scarce scenarios by reusing knowledge learned from related tasks or domains. In the context of building external loads prediction, we distinguish between a source domain $\mathcal{D}_S$, which contains multiple buildings with sufficient simulation data, and a target domain $\mathcal{D}_T$, corresponding to a new building with limited labeled samples. Although buildings in $\mathcal{D}_S$ and $\mathcal{D}_T$ may differ in geometry, topology, and climate, they share common physical heat-transfer mechanisms, which can be exploited for knowledge transfer.

Let $f(\cdot;\theta)$ denote the proposed physics-embedded GNN, parameterized by θ. The learning objective in the source domain is

$$\theta^* = \arg\min_{\theta} \mathbb{E}_{(G,\dot{Q})\sim\mathcal{D}_S} \mathcal{L}\left[f(G;\theta), \dot{Q}\right], \tag{11}$$

where G denotes the building graph and $\dot{Q}$ represents space-level external load response.

2.3.1 Domain-invariant graph feature transfer

The model parameters θ can be decomposed into two parts:

$$\theta = \{\theta_{\mathrm{GNN}}, \theta_{\mathrm{MLP}}\}, \tag{12}$$

where θ_{GNN} corresponds to the graph convolution layers and θ_{MLP} denotes the space-level prediction head.

The graph convolution layers encode physics-guided message passing between space and surface nodes, capturing generalizable thermal interaction patterns that are largely independent of specific building instances.

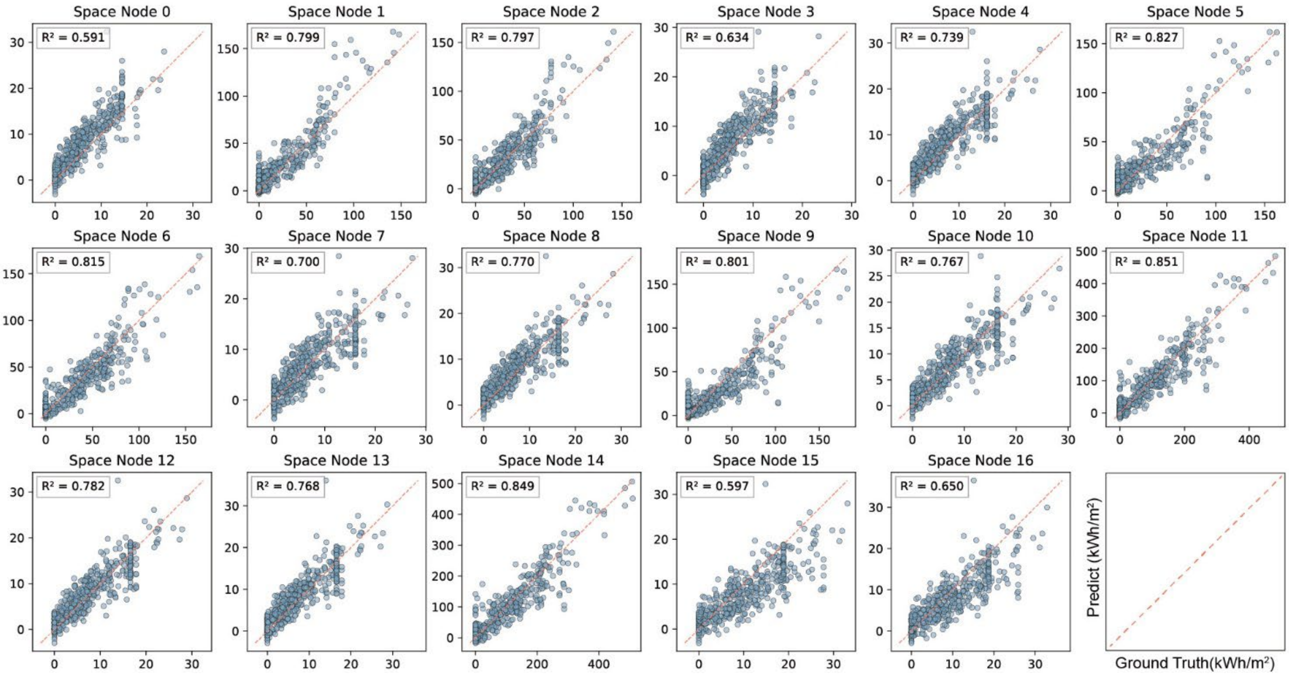

Fig. 4 Scatter plots of predicted vs. ground-truth values for each space node.

Therefore, during transfer from $\mathcal{D}_S$ to $\mathcal{D}_T$, the parameters θ_{GNN} are frozen and directly reused:

$$\theta_{GNN}^{T} = \theta_{GNN}^{S} \tag{13}$$

This strategy preserves physically consistent feature extraction and prevents overfitting when only limited target-domain data are available.

2.3.2 Target-domain adaptation via prediction head fine-tuning

The output MLP maps the learned space-level graph representations to external disturbance load predictions, which may vary across buildings due to differences in envelope properties, shading conditions, and climate exposure. To account for these variations, only the prediction head is fine-tuned using target-domain samples:

$$\theta_{MLP}^{T} = \arg\min_{\theta_{\text{MLP}}} \mathbb{E}_{(G,\dot{Q})\sim\mathcal{D}_T} \mathcal{L}\left[f(G;\theta_{GNN}^{S},\theta_{\text{MLP}}),\dot{Q}\right] \tag{14}$$

By restricting adaptation to the MLP layer, the proposed framework efficiently transfers physics-guided knowledge of envelope-driven heat transfer while adapting to building-specific external load characteristics. This strategy enables robust external load prediction across buildings and climates with minimal target-domain data, as demonstrated in the experimental evaluation.

3 Experimental settings

This section introduces the experimental design and datasets used to evaluate the proposed physics-informed graph neural network (GNN) and the graph-based transfer learning framework. Two experiments are conducted. The first examines single-building external load prediction under a limited data regime. The second evaluates cross-building transfer learning for external load prediction.

3.1 Experiment I: Single-building external load prediction

The first experiment focuses on hourly external load prediction for a single building using physically structured graph representations. The prediction target is the space-level external thermal load, driven by hourly weather conditions. No internal gains or occupancy-related loads are considered, ensuring that the task isolates envelope–weather interactions.

The dataset consists of one full year of simulation data, with hourly resolution (8760 samples). Each sample includes:

- Hourly weather variables (e.g., outdoor temperature, solar radiation, wind-related features),
- A fixed building graph encoding spatial nodes (spaces) and envelope component nodes (walls, windows, roofs),
- Edge attributes coupling weather inputs with envelope component features.

Two models are compared: a multilayer perceptron (MLP) baseline and the proposed physics-informed graph neural network (GNN). The MLP directly maps weather

https://doi.org/10.1051/e3sconf/202671602038

features to external load, whereas the GNN leverages a graph representation of the building envelope, where nodes correspond to envelope components and edges encode their physical adjacency and heat transfer relationships. Physical priors, such as surface orientation and thermal properties, are embedded into node and edge features.

3.2 Experiment II: Cross-building transfer learning for external load prediction

To investigate the effectiveness of cross-building transfer learning under limited data availability, a two-stage training strategy was designed for external load prediction. The core motivation is to leverage structural and physical representations learned from a source building dataset to improve prediction performance on a target building with scarce labeled data.

The transfer learning framework consists of a pretraining stage and a fine-tuning stage. During pretraining, the GNN model is trained on a source-domain dataset to learn generic building-level representations. In the transfer stage, the pretrained model is adapted to the target-domain dataset by freezing all graph convolution layers and updating only the fully connected (FC) layers. This design aims to preserve transferable spatial–topological features while allowing lightweight adaptation to the target building.

Two transfer scenarios were considered.

- Task 1 uses a building dataset (with **681 different buildings** and their external load in **the same weather condition**) split into source and target domains with a 6:4 ratio, simulating a partial data availability scenario within similar building contexts.
- Task 2 uses an additional independent dataset (with **220 different buildings** and their external load in **the different weather condition**) as the target domain, representing a more challenging cross-building transfer setting.

The source and target sample sizes for Task 1 are 544 and 137, respectively, while Task 2 contains 681 source samples and 220 target samples.

For both tasks, pretraining was conducted for 50 epochs with a batch size of 8 and a learning rate of 1×10^{-3}. Transfer learning was then performed for 50 epochs using a reduced learning rate of 1×10^{-4}, following standard fine-tuning practice. All evaluations were performed on denormalized (real-scale) energy values to reflect practical prediction accuracy.

4 Results and discussion

4.1 GNN prediction performance evaluation results in a single building

GNN model was trained and evaluated on a single-building dataset using the optimal hyperparameter configuration identified in the previous section, consisting of two heterogenous layers, one MLP layer, a hidden dimension of 16, a batch size of 72, and a learning rate of 5×10^{-3}. The dataset was randomly split into training and testing subsets with an 80:20 ratio, and the model was trained for 50 epochs using ReLU activations.

Under this configuration, the training process converged within approximately 40 epochs, with stable training and testing losses. To assess robustness, the experiment was repeated five times with different random initializations. This model achieved an average MAE of 7.96 ± 0.49, an RMSE of 16.41 ± 0.40, and an R^2 of 0.88 ± 0.01, indicating stable and accurate zone-level energy predictions.

Fig. 4 shows predicted versus ground-truth hourly energy consumption for individual zones. Zones with higher energy intensity and larger dynamic ranges exhibit stronger correlations, while low-load zones show larger relative dispersion. Overall, GNN maintains consistent predictive performance across heterogeneous zones within the building.

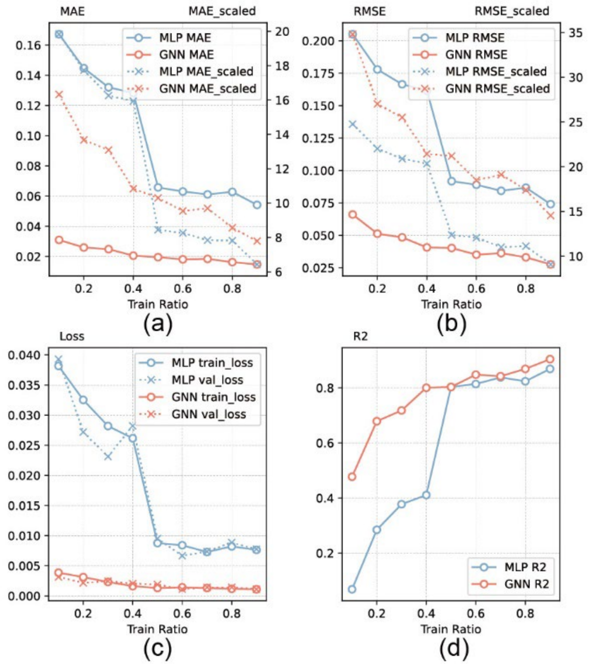

Fig. 5 (a) MAE and scaled MAE as a function of training data ratio. (b) RMSE and scaled RMSE across different training ratios. (c) Training and validation loss comparison under varying data availability. (d) R^2 versus training ratio.

GNN was further compared with a baseline MLP model under varying training data ratios. While the MLP directly maps weather inputs to zone energy demand, GNN incorporates building topology through message passing between space and face nodes. As shown in **Fig. 5**, GNN converges faster and achieves lower training and validation losses than MLP. Across different training ratios, GNN consistently outperforms MLP in terms of MAE, RMSE, and R^2, particularly in low-data regimes, demonstrating superior data efficiency. With increased training data, the performance gap between the two models narrows, and MLP shows improved absolute accuracy. These results indicate that GNN is better suited for structure-aware energy prediction tasks under limited data availability, while MLP benefits more strongly from large datasets.

4.2 Transfer learning results and effectiveness analysis

The quantitative results of the transfer learning experiments are summarized in Tab. 1 and **Fig. 6**. Overall, transfer learning leads to consistent but task-dependent improvements compared with training from scratch on the target-domain data.

For Task 1, where the source and target domains are derived from the same dataset, transfer learning yields a modest improvement of approximately 3% in R^2 compared with direct training. This indicates that when domain similarity is high and data distribution differences are limited, the benefit of transfer learning is constrained but still observable.

For Task 2, which involves an independent target dataset, the performance gain becomes more pronounced. Transfer learning improves the R^2 score by approximately 12%, accompanied by noticeable reductions in MAE and RMSE. This suggests that pretrained representations learned from a larger and structurally similar source building can effectively compensate for data scarcity in the target domain.

Tab. 1 Model performance metrics on different tasks

	TrainLoss	Val_MAE	Val_RMSE	Val_R2
Original	0.010356	49.3019	79.5901	0.7495
Task 1 pretrained	0.008992	33.1911	56.2691	0.7903
Task 1 transfer	0.009029	49.7019	79.6080	0.7726
Task 2 pretrained	0.010199	43.2707	74.1928	0.7755
Task 2 transfer	0.011612	41.2325	65.7275	0.7784

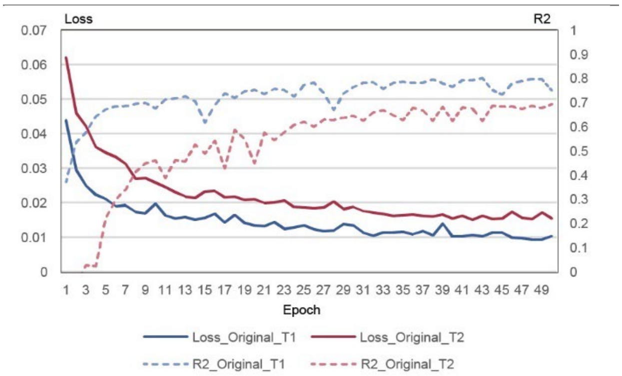

Fig. 6 Training dynamics of the original (non-transfer) models on Task 1 (T1) and Task 2 (T2), showing the evolution of loss (dashed lines) and R² (solid lines) over training epochs.

From a computational perspective, transfer learning also demonstrates clear efficiency advantages. Training from scratch on the small target dataset (220 samples) for 50 epochs requires approximately one minute. In contrast, the transfer strategy involves a one-time pretraining cost (about two minutes) and a very short fine-tuning phase on new data (around four seconds), while enabling reuse of the pretrained model for future buildings. This trade-off is favorable in multi-building deployment scenarios.

Overall, while transfer learning does not dramatically alter prediction accuracy in all cases, it provides measurable performance gains under small-sample conditions, particularly for cross-building generalization, and significantly reduces the marginal cost of adapting the model to new buildings. These results support the practicality of transfer learning as an effective strategy for data-limited building energy prediction tasks.

4.3 Discussion

The results demonstrate that incorporating building topology through a heterogeneous GNN yields clear advantages over purely data-driven baselines, particularly under limited data availability. The consistently higher accuracy and faster convergence observed in low-data regimes confirm that explicit structural priors help constrain the hypothesis space and improve generalization at the zone level. The transfer learning experiments further support this interpretation: when the target building shares structural similarity with the source, pretrained representations can effectively reduce data dependency and adaptation cost, making the approach practical for multi-building deployment scenarios.

However, several limitations should be acknowledged. First, the experiments are restricted to a small number of buildings and climates, and the observed transfer gains may not directly generalize to buildings with substantially different typologies, operational schedules, or HVAC control strategies. Second, the current model focuses primarily on envelope-driven and weather-driven loads, while internal gains and occupant behavior are treated implicitly, which may limit performance in zones dominated by stochastic usage patterns. Third, although the GNN improves interpretability at a structural level, the learned message-passing features remain only partially transparent in physical terms. Finally, the study does not include a dedicated ablation analysis to rigorously isolate and quantify the contribution of the proposed physical embedding mechanism; therefore, the specific performance gains attributable to this design choice cannot yet be conclusively demonstrated.

Future work will focus on extending validation to more diverse building stocks and climates, incorporating richer operational and occupancy-related features, and exploring physics-informed regularization to further align learned representations with physical energy balance constraints. In addition, adaptive transfer strategies that explicitly account for cross-building heterogeneity may further improve robustness and scalability.

5 Conclusion

This study proposed a Graph-based Physics-Embedded Transfer Learning framework for multi-zone building external disturbance load prediction, aiming to bridge the gap between physical interpretability and data efficiency. By representing buildings as heterogeneous graphs composed of space and envelope surface nodes, and by embedding fundamental heat transfer mechanisms directly into the message-passing process, the proposed approach provides a physically grounded yet flexible modeling paradigm suitable for early-stage design and data-scarce scenarios.

The experimental results demonstrate that the physics-informed GNN consistently outperforms a purely data-driven MLP baseline in single-building prediction tasks, particularly under limited training data. The advantage is most evident in low-data regimes, confirming that explicit structural and physical priors help constrain learning and improve generalization. Moreover, the transfer learning experiments show that freezing physics-guided graph convolution layers and fine-tuning only the prediction head enables effective knowledge reuse across buildings and climates. This strategy yields measurable performance gains and substantially reduces the marginal cost of adapting models to new buildings, supporting its practicality for scalable deployment.

Overall, the framework demonstrates a practical and scalable approach for structure-aware and transferable building energy modeling, with clear potential for extension to more comprehensive building performance prediction tasks.

The study was supported by the National Natural Science Foundation of China (Grant No.52425801, 52394223 and 52130803), and the Tsinghua University Initiative Scientific Research Program (Grant No.20257020014).

References

[1] Y. C. for E. + A. United Nations Environment Programme, "Building Materials and the Climate: Constructing a New Future," Sep. 2023. [Online]. Available: https://wedocs.unep.org/20.500.11822/43293

[2] L. Zhang *et al.*, "A review of machine learning in building load prediction," *Appl. Energy*, vol. 285, p. 116452, Mar. 2021, doi: 10.1016/j.apenergy.2021.116452.

[3] D. Lin, X. Xu, K. Liu, T. Wu, X. Wang, and R. Zhang, "Interpretable data-driven urban building energy modeling considering inter-building effect," *Build. Environ.*, vol. 274, p. 112688, Apr. 2025, doi: 10.1016/j.buildenv.2025.112688.

[4] D. Chen, C.-K. Chui, and P. S. Lee, "Physics informed machine learning based predictive control for intelligent operation of edge datacenters," *Appl. Energy*, vol. 402, p. 126975, Jan. 2026, doi: 10.1016/j.apenergy.2025.126975.

[5] V. R. Chifu, T. Cioara, C. B. Pop, I. Anghel, and A. Pelle, "Physics-informed neural networks for heat pump load prediction," *Energies*, vol. 18, no. 1, Art. no. 8, 2025, doi: 10.3390/en18010008.

[6] Z. Jiang *et al.*, "Physics-informed machine learning for building performance simulation-A review of a nascent field," *Adv. Appl. Energy*, vol. 18, p. 100223, Jun. 2025, doi: 10.1016/j.adapen.2025.100223.

[7] Y. Chen, Q. Yang, Z. Chen, C. Yan, S. Zeng, and M. Dai, "Physics-informed neural networks for building thermal modeling and demand response control," *Build. Environ.*, vol. 234, p. 110149, Apr. 2023, doi: 10.1016/j.buildenv.2023.110149.

[8] Y. Hu, X. Cheng, S. Wang, J. Chen, T. Zhao, and E. Dai, "Times series forecasting for urban building energy consumption based on graph convolutional network," *Appl. Energy*, vol. 307, p. 118231, Feb. 2022, doi: 10.1016/j.apenergy.2021.118231.

[9] Z. Yu, Y. Li, J. Xiao, H. Zhou, and B. Lin, "Physical embedding on building surface spatial relationships shows better performance in graph-based daylight prediction," *Build. Environ.*, vol. 290, p. 114141, Feb. 2026, doi: 10.1016/j.buildenv.2025.114141.

[10] X. Yan, X. Wu, H. Ta, and Y. Cao, "A transfer learning strategy with multi-source data for building cooling load prediction under data scarcity across multiple architectural scenarios," *Appl. Therm. Eng.*, vol. 288, p. 129589, Mar. 2026, doi: 10.1016/j.applthermaleng.2025.129589.

[11] M. Ribeiro, K. Grolinger, H. F. ElYamany, W. A. Higashino, and M. A. M. Capretz, "Transfer learning with seasonal and trend adjustment for cross-building energy forecasting," *Energy Build.*, vol. 165, pp. 352–363, Apr. 2018, doi: 10.1016/j.enbuild.2018.01.034.

[12] C. Wu, X. Zhang, Y. Li, Z. Chen, and Z. Zhang, "Cross-building energy consumption forecasting with deep transfer learning," *J. Build. Eng.*, vol. 117, p. 114695, Jan. 2026, doi: 10.1016/j.jobe.2025.114695.

[13] F. Peng, T. Su, Q. Zeng, and X. Han, "Climate-adaptive energy forecasting in green buildings via attention-enhanced Seq2Seq transfer learning," *Sci. Rep.*, vol. 15, no. 1, p. 31829, Aug. 2025, doi: 10.1038/s41598-025-16953-y.