054

000

HVAC-GRACE: Transferable Building Control via Graph Recurrent Neural Networks

Anonymous Authors¹

Abstract

Buildings consume 40% of global energy, with HVAC systems responsible for up to half of that demand (IEA, 2024). As energy use grows, optimizing HVAC efficiency is critical to meeting climate goals (Santamouris, 2016). While reinforcement learning (RL) offers a promising alternative to rule-based control, real-world adoption is limited by poor sample efficiency and generalisation (Nagy et al., 2023). We introduce HVAC-GRACE, a graph-based RL framework that models buildings as heterogeneous graphs and integrates spatial message passing directly into temporal GRU gates. This enables each zone to learn control actions informed by both its own history and its structural context. Our architecture supports zeroshot transfer by learning topology-agnostic functions-but initial experiments reveal this benefit depends on sufficient conditioned zone connectivity to maintain gradient flow. These findings highlight the promise and requirements of scalable, transferable building control.

1. Introduction and Related Work

Most buildings use rule-based thermostats with static setpoints, despite advanced strategies that could significantly reduce HVAC energy consumption (Drgoňa et al., 2020) and provide grid flexibility while reducing emissions (Zhou et al., 2023). While model-based, data-driven, and learningbased methods show promise, they lack generalisation and require extensive training, limiting deployment across buildings (Wang & Hong, 2020; Nagy et al., 2023).

Model Predictive Control requires costly building-specific models, while RL suffers from sample inefficiency, often needing years of training (Krishna et al., 2023) during which buildings experience suboptimal performance (Zhang et al., 2021). Pretraining in simulation is impractical due to simulator development costs. Furthermore, policies struggle to transfer between buildings due to varying characteristics (Zisman et al., 2023; Zhang et al., 2022), requiring retraining for each deployment. Recent RL and transfer-learning studies (Xu et al., 2020; Berkes, 2024) and morphology-aware methods in robotics (Wang et al., 2018; Huang et al., 2020; Gupta et al., 2022) show that structure helps, but buildings pose unique heterogeneity and coupling challenges.

Our work addresses a fundamental limitation: treating buildings as generic control problems using flat policies applied to concatenated state vectors that ignore inherent structural organisation. Buildings exhibit complex spatial and temporal relationships—zones have thermal adjacency determining heat transfer, weather affects zones differently based on orientation and insulation, and HVAC equipment has local effects propagating through structure. These relationships remain consistent across conditions, suggesting that structure-aware policies could transfer more effectively than those learning implicit representations from scratch.

Research in robotics demonstrates that structure-aware approaches improve RL performance. NerveNet (Wang et al., 2018) showed that encoding morphological structure as graphs enables better sample efficiency and generalisation by sharing parameters across similar components and explicitly modeling physical relationships. This trend extends to other morphology-aware methods: SMP (Huang et al., 2020) leverages graph representations for similar benefits, while Metamorph (Gupta et al., 2022) uses morphology-aware transformers to capture structural dependencies.

However, buildings present distinct challenges: heterogeneous node types with different thermal properties, complex multi-timescale dynamics, and variable control topologies requiring specialised architectural considerations.

Contributions: We introduce **HVAC-GRACE** (Graph Reinforcement Adaptive Control Engine), the first graph-based RL framework for building control.

We contribute (*i*) a heterogeneous graph representation for HVAC, (*ii*) a unified spatial-temporal Graph RNN with topology-agnostic functions, and (*iii*) empirical design insights showing that adequate conditioned-zone connectivity

¹Anonymous 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.

 $(\approx 50\%)$ is required for effective graph RL.

2. Methodology

057

058

059

074

075

090

109

2.1. Problem Formulation

060 We formulate HVAC control as a Markov Decision Process 061 $(\mathcal{S}, \mathcal{A}, \mathcal{P}, r, \gamma)$ where: \mathcal{S} includes zone temperatures, out-062 door weather (temperature, humidity), time features (hour, 063 day of year), and HVAC energy consumption (electricity, 064 gas); A represents heating and cooling temperature setpoints 065 for each controllable zone; \mathcal{P} defines physics-based tran-066 sitions through EnergyPlus (DOE, 2024) simulation; the 067 reward r(s, a) balances energy consumption and comfort 068 violations; and γ is the discount factor. 069

070 The agent learns a stochastic policy $\pi_{\theta}(a_t|s_t)$ that 071 maximizes expected discounted return $J(\theta) =$ 072 $\mathbb{E}_{\pi_{\theta}} \sum_{t=0}^{\infty} \gamma^t r(s_t, a_t).$

2.2. Heterogeneous Building Graph Construction

We represent buildings as heterogeneous directed graphs $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with three node types:

078
079Conditioned zones (\mathcal{V}_c) : Indoor spaces with thermostat
control that receive local observations and generate control
actions. Unconditioned zones (\mathcal{V}_u) : Indoor spaces with-
out active temperature control that provide thermal context.081
out active temperature control that provide thermal context.082
083
083084
084

085Edge relationships \mathcal{E} capture thermal connections: thermal086adjacency edges connect zones sharing surfaces, environ-087mental influence edges connect outdoor nodes to zones with088exterior surfaces, and self-loop edges enable temporal state089maintenance.

091 2.3. Temporal-Spatial Policy Architecture

 Our policy processes heterogeneous building states through two integrated stages, where spatial and temporal processing are unified within Graph RNN cells.

Stage 1: Input Processing and Heterogeneous Graph 096 Construction Raw observations from the building simu-097 098 lation must be transformed into a structured graph representation that captures the building's thermal topology. 099 100 Raw observations are first processed through type-specific input MLPs and parsed into heterogeneous graph format: $x^{t} = \text{InputMLP}(\text{parsed_obs}), \text{ where observations are struc-}$ tured as node dictionaries with zone temperatures, weather data, and temporal features mapped to their respective node 104 105 types. This preprocessing is essential because building zones have different thermal characteristics (conditioned 106 vs. unconditioned vs. outdoor) and require type-specific feature encoding to capture their distinct thermal behaviours

and control capabilities.

Stage 2: Integrated Spatial-Temporal Processing via Graph RNN Traditional approaches handle temporal and spatial dependencies separately: RNNs capture temporal patterns but ignore spatial relationships (Lipton et al., 2015), while GNNs capture spatial relationships but struggle with long-term temporal dependencies (Scarselli et al., 2009). However, in buildings these dependencies are tightly coupled—how a zone's temperature evolves depends critically on what neighbouring zones are doing.

Our Graph RNN unifies these relationships by replacing each GRU gate with a heterogeneous GNN. Standard GRU gates (reset, update, and new gates) are the core mechanism in RNNs for controlling how temporal memory is updated at each timestep. By implementing these gates as GNNs instead of simple linear transformations, we enable spatial context from neighbouring nodes to influence temporal memory updates.

Each heterogeneous GNN (HeteroGNN) performs typespecific message passing: (1) compute messages using functions $f_{\text{msg}}^{(\text{type}(u),\text{type}(v))}(h_u)$ that encode thermal physics relationships, (2) aggregate messages $\bar{m}_v = \text{AGG}(\{m_{u \to v} : u \in \mathcal{N}(v)\})$, and (3) update node representations via $f_{\text{update}}^{\text{type}(v)}(h_v, \bar{m}_v)$. Instead of zones updating memory in isolation, gates perform message passing across the building graph::

$$r_{\rm raw}^t = \text{HeteroGNN}_{\rm reset}(x^t, \mathcal{E}) \tag{1}$$

$$z_{\text{raw}}^{t} = \text{HeteroGNN}_{\text{update}}(x^{t}, \mathcal{E})$$
(2)

$$n_{\rm raw}^t = \text{HeteroGNN}_{\rm new}(x^t, \mathcal{E}) \tag{3}$$

GNN outputs combine with previous hidden states through type-specific transformations:

$$r_v^t = \sigma(r_{\text{raw},v}^t + W_r^{\text{type}(v)} h_v^{t-1})$$
(4)

$$z_v^t = \sigma(z_{\text{raw},v}^t + W_z^{\text{type}(v)} h_v^{t-1}) \tag{5}$$

$$\tilde{h}_{v}^{t} = \tanh(n_{\text{raw},v}^{t} + r_{v}^{t} \odot W_{h}^{\text{type}(v)} h_{v}^{t-1})$$
(6)

$$h_v^t = (1 - z_v^t) \odot \tilde{h}_v^t + z_v^t \odot h_v^{t-1}$$

$$\tag{7}$$

This enables spatial context to influence temporal memory updates. When Zone B computes its reset gate while adjacent Zone A is heating, the message function processes Zone A's state, affecting how Zone B updates its thermal memory—anticipating heat transfer and enabling coordinated control decisions.

Stage 3: Type-Specific Action Generation After processing through the Graph RNN, specialised policy heads generate control actions for conditioned zones only. For each conditioned zone $v \in \mathcal{V}_c$, we output Gaussian action distribution parameters:

$$\mu_v, \log \sigma_v = f_{\text{policy}}^{\text{conditioned}}(h_v^t)$$

Actions are sampled as $a_v \sim \mathcal{N}(\mu_v, \exp(\log \sigma_v))$. Figure 1 111 shows the framework for the successful implementation of 112 HVAC-GRACE.



Figure 1. The integration of heterogeneous message-passing GNNs within temporal GRU gates. This combined spatial-temporal model provides a principled method for learning dynamic thermal interactions across building zones.

2.4. Transferable Architecture

113

114

115

116

117

118

119

120

121

122

124

125

126

127

128

142

143

144

129 HVAC-GRACE is designed for zero-shot transfer by 130 learning type-specific functions that operate over node 131 types, rather than fixed input positions. Traditional RL 132 policies rely on fixed-dimensional state vectors a =133 $f_{\text{MLP}}([x_1, x_2, \dots, x_n])$, making them incompatible with 134 changing building topologies.

135 In contrast, our Graph RNN uses message functions $f_{msg}^{(type(u),type(v))}$, update functions $f_{update}^{type(v)}$, and policy heads 136 137 $f_{\text{policy}}^{\text{conditioned}}$ that generalise across graph structures. These 138 compact MLPs remain invariant to the number of zones 139 or layout, enabling policy transfer without architectural 140 changes. 141

3. Experimental Setup

Implementation: We demonstrate our approach using a 145 6-zone commercial building model (a Small Hotel layout with only 1 conditioned zone and 5 unconditioned zones) 147 simulated with EnergyPlus (DOE, 2024) through Minergym, 148 a Gym-compatible environment. The current preliminary 149 experiments focus on a scenario with Typical Meteorolog-150 ical Year (TMY) weather data for Pike County, Alabama 151 (hot-humid climate) for training, with generalisation evalua-152 tion on Montreal, Quebec weather data (continental climate) 153 to assess climate transferability. Comprehensive experi-154 ments involving multiple building topologies, diverse cli-155 mate zones, and systematic comparisons to baselines are 156 underway, aiming to thoroughly explore the transferability 157 and effectiveness of our proposed graph-based method. 158

159 Method: We implement the HVAC-GRACE approach, and 160 a comparison using PPO with identical hyperparameters: 161 64-dimensional hidden layers, PPO learning rate 3e-4, and 162 50K training timesteps (approximately 0.57 years). HVAC-163 GRACE uses K = 2 message passing layers with GRU 164

Algorithm 1 HVAC-GRACE Training Algorithm

- 1: Input: Building epJSON file, training episodes N
- 2: $\mathcal{G} = (\mathcal{V}, \mathcal{E}) \leftarrow \text{ConstructGraph}(epJSON)$
- 3: Initialise policy π_{θ} , critic V_{ϕ} , Graph RNN states $\{h_v^0\}_{v\in\mathcal{V}}$
- 4: for episode = 1 to N do
- Reset environment, initialise s_0 , $\{h_v^0\}_{v \in \mathcal{V}}$ 5:
- 6: for timestep t = 0 to T - 1 do
- 7: Stage 1: Input Processing
- parsed_obs \leftarrow ParseObservation(s_t) 8:
- 9: $x^t \leftarrow \text{InputMLP}(\text{parsed_obs})$
- Stage 2: Graph RNN Processing 10:
- $\{h_v^t\}_{v \in \mathcal{V}} \leftarrow \text{GraphRNNCell}(x^t, \mathcal{E}, \{h_v^{t-1}\}_{v \in \mathcal{V}})$ 11:
- Stage 3: Action Generation 12:
- for each conditioned zone $v \in \mathcal{V}_c$ do 13: 14:
- $\begin{aligned} \mu_v, \log \sigma_v &\leftarrow f_{\text{policy}}(h_v^t) \\ \text{Sample } a_v &\sim \mathcal{N}(\mu_v, \exp(\log \sigma_v)) \end{aligned}$ 15:
- end for 16:
- Execute actions $\{a_v\}_{v \in \mathcal{V}_c}$, observe s_{t+1}, r_t 17:
- 18: end for
- 19: Update policy π_{θ} and critic V_{ϕ} using PPO



cells for temporal encoding. Our reward balances energy efficiency with occupant comfort:

$$r_t = -\lambda_e \cdot \text{Energy}_t - \lambda_c \cdot \sum_{v \in \mathcal{V}_c} \text{ComfortViolation}_v^t$$

where the ComfortViolation term measures quadratic temperature deviations from the 20-24°C range. We set $\lambda_e = 1.0$ and $\lambda_c = 1.0$ with equal weighting to balance energy efficiency and occupant comfort.

4. Results

4.1. Feasibility of Graph-Based RL and Topology Limitations

We first validate the feasibility of the HVAC-GRACE framework, which integrates GNNs with recurrent units (GRUs) for model-free reinforcement learning in HVAC control. The model successfully generates control actions from structured, heterogeneous graph inputs, confirming the technical viability of our novel spatial-temporal architecture. This result establishes the foundational capability to embed spatial message passing within temporal memory updates.

We tested HVAC-GRACE on a DOE Small Hotel model featuring six unconditioned zones and only one conditioned zone-a highly imbalanced topology with just 14% of nodes directly contributing to the policy gradient, as shown in Figure 2. This scenario revealed a fundamental limitation of graph-based control: with limited conditioned connectivity, the majority of the graph receives no learning signal.
This disrupts the gradient flow needed for effective message
passing, undermining the benefits of spatial reasoning.



Figure 2. Gradient propagation in HVAC-GRACE for (left) a wellconditioned building and (right) the imbalanced Small Hotel topology. Green nodes (conditioned zones) receive direct gradients from policy heads; orange nodes (unconditioned zones) receive gradients only through message passing connections. In sparse topologies, limited conditioned connectivity disrupts gradient flow to unconditioned zones, undermining graph-based learning.

This result illustrates that the success of our topologyagnostic method hinges on having both sufficient conditioned zones and connectivity to sustain gradient flow across the building graph. Without it, spatial dependencies are effectively unlearnable, and the model degrades to a singlezone controller. See Appendix A for experimental results on the Small Hotel topology and Appendix B for a detailed gradient-flow analysis.

4.2. Transferability by Design

Our architecture generalises across building topologies without retraining. Unlike flat policies, which are tightly coupled to a fixed number of input zones, HVAC-GRACE learns thermal interaction functions tied to node types. This enables zero-shot transfer to new buildings by:

(1) parsing a new building's epJSON file to extract node
and edge types, (2) loading pre-trained model weights unchanged, and (3) performing inference directly with the
same type-specific functions.

This capability reduces deployment cost and complexity. However, our initial experiments highlight an important caveat: successful transfer requires sufficient connectivity among conditioned zones. In topologies like the Small Hotel, where only 14% of zones receive learning signals, the graph-based reasoning mechanism fails to activate meaningfully, undermining generalisation. These findings suggest that while HVAC-GRACE is structurally transferable, its effectiveness depends on topological conditions that support stable gradient propagation.

5. Discussion and Next Steps

While the current results did not yield immediate numerical advantages, this exploration strongly emphasises the critical role of building topology in determining the utility of structured RL representations. The demonstrated pipeline and identification of topology-related limitations represent essential steps toward achieving robust, scalable, and transferable RL-based HVAC control solutions in the future.

This preliminary result clearly illustrates that building topology strongly determines the suitability and effectiveness of graph-based reinforcement learning for HVAC control. We hypothesise that successful graph-based learning necessitates: A) Balanced Ratio of Conditioned Zones: A minimum threshold (50–60%) of conditioned zones to ensure stable gradient flow and meaningful message passing across the building's graph representation. B) Zones should exhibit strong thermal interactions and sufficient spatial-temporal coupling to generate rich and informative graph representations.

The Small Hotel's 14% conditioned zone ratio is significantly below this threshold, which explains the ineffectiveness of the graph-based method in this initial demonstration. Recognising these findings, our immediate ongoing research includes:

Evaluating Alternative Topologies: We have already begun testing the approach on other building configurations, notably the DOE SmallOffice building, which exhibits a more balanced ratio of conditioned zones and diverse thermal interactions.

Defining the Applicability Threshold: Systematically identifying which buildings and topologies benefit most from graph-based RL, establishing clear guidelines on when graph-based models provide substantial advantages over simpler, non-structured RL approaches.

Exploring Offline Training and Transferability: Leveraging the topology-agnostic nature of our method, we plan extensive experiments training offline across diverse building topologies, which could substantially improve policy robustness and transferability in practical, real-world deployments.

Impact Statement

Graph-based policy architectures could enable the rapid, cost-effective deployment of intelligent HVAC systems at scale— providing a critical pathway toward the 50% reduction in building energy consumption needed to meet global climate targets.

References

- Berkes, A. Hvac-dpt: A decision pretrained transformer for hvac control. *arXiv preprint arXiv:2411.19746*, 2024.
- DOE. Energyplus, 2024. URL https://energyplus.net/.
- Drgoňa, J., Arroyo, J., Figueroa, I. C., Blum, D., Arendt, K., Kim, D., Ollé, E. P., Oravec, J., Wetter, M., Vrabie, D. L., et al. All you need to know about model predictive control for buildings. *Annual Reviews in Control*, 50: 190–232, 2020.
- Gupta, A., Fan, L., Ganguli, S., and Fei-Fei, L. Metamorph: Learning universal controllers with transformers. *arXiv preprint arXiv:2203.11931*, 2022.
- Huang, W., Mordatch, I., and Pathak, D. One policy to control them all: Shared modular policies for agent-agnostic control. In *International Conference on Machine Learning*, pp. 4455–4464. PMLR, 2020.
- IEA. Energy efficiency 2024, 2024. URL https://www.iea.org/reports/ energy-efficiency-2024. Licence: CC BY 4.0.
- Krishna, G. S. A., Zhang, T., Ardakanian, O., and Taylor, M. E. Mitigating an adoption barrier of reinforcement learning-based control strategies in buildings. *Energy and Buildings*, 285:112878, 2023.
- Lipton, Z. C., Berkowitz, J., and Elkan, C. A critical review of recurrent neural networks for sequence learning. *arXiv preprint arXiv:1506.00019*, 2015.
- Nagy, Z., Henze, G., Dey, S., Arroyo, J., Helsen, L., Zhang,
 X., Chen, B., Amasyali, K., Kurte, K., Zamzam, A.,
 Zandi, H., Drgoňa, J., Quintana, M., McCullogh, S., Park,
 J. Y., Li, H., Hong, T., Brandi, S., Pinto, G., Capozzoli,
 A., Vrabie, D., Bergés, M., Nweye, K., Marzullo, T., and
 Bernstein, A. Ten questions concerning reinforcement
 learning for building energy management. *Building and Environment*, 241:110435, 2023. ISSN 0360-1323.
- Santamouris, M. Cooling the buildings–past, present and future. *Energy and Buildings*, 128:617–638, 2016.
- Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M., and Monfardini, G. The graph neural network model. *IEEE Transactions on Neural Networks*, 20(1):61–80, 2009. doi: 10.1109/TNN.2008.2005605.
- Wang, T., Liao, R., Ba, J., and Fidler, S. Nervenet: Learning structured policy with graph neural networks. In *International conference on learning representations*, 2018.

- Wang, Z. and Hong, T. Reinforcement learning for building controls: The opportunities and challenges. *Applied Energy*, 269:115036, 2020.
- Xu, S., Wang, Y., Wang, Y., O'Neill, Z., and Zhu, Q. One for many: Transfer learning for building hvac control. In Proceedings of the 7th ACM international conference on systems for energy-efficient buildings, cities, and transportation, pp. 230–239, 2020.
- Zhang, T., Baasch, G., Ardakanian, O., and Evins, R. On the joint control of multiple building systems with reinforcement learning. In *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*, pp. 60–72, 2021.
- Zhang, T., Afshari, M., Musilek, P., Taylor, M. E., and Ardakanian, O. Diversity for transfer in learning-based control of buildings. In *Proceedings of the Thirteenth ACM International Conference on Future Energy Systems*, pp. 556–564, 2022.
- Zhou, X., Xue, S., Du, H., and Ma, Z. Optimization of building demand flexibility using reinforcement learning and rule-based expert systems. *Applied Energy*, 350: 121792, 2023. ISSN 0306-2619. doi: https://doi.org/10. 1016/j.apenergy.2023.121792.
- Zisman, I., Kurenkov, V., Nikulin, A., Sinii, V., and Kolesnikov, S. Emergence of in-context reinforcement learning from noise distillation. *arXiv preprint arXiv:2312.12275*, 2023.

A. HVAC-GRACE Performance Analysis

275

289 290 291

292

293 294

295

296

297

299

300

301

302 303

304

305

306

307

308

309

310

Table 1 compares PPO and HVAC-GRACE performance on our generalisation test (Alabama training \rightarrow Montreal evaluation). HVAC-GRACE's substantially worse performance stems from the building topology creating sparse graph connectivity that leads to zero gradient flows during training, preventing effective policy optimisation. We find that using a flat MLP value function does not improve the performance of HVAC-GRACE either, suggesting that the graph-structured policy is fundamentally ill-suited for this building topology with predominantly unconditioned zones and sparse inter-zone connectivity.

Table 1. Performance Comparison: Alabama Training \rightarrow Montreal Generalisation			tion
Method	Alabama	Montreal	Gap
PPO (MLP Policy + MLP Value)	-118.89	-158.46	39.57
HVAC-GRACE (Graph Policy + Graph Value)	-353.30	-368.27	14.97

B. Gradient-Flow Pathology in Sparse Topologies

Why insufficient conditioned connectivity breaks graph learning.

Gradient Flow Mechanics: Policy gradients originate exclusively from conditioned zones that generate actions through policy heads. For graph-based learning to succeed, these gradients must propagate through message passing connections to create meaningful representations across the entire building graph.

- Normal case (balanced topology). Gradients originate at multiple conditioned zones and propagate to neighbors via bidirectional message passing. This creates rich learning signals across the graph: conditioned zones receive direct policy gradients, while unconditioned zones receive gradients through bidirectional Conditioned-Unconditioned message functions during backpropagation.
- **Sparse case (low conditioned connectivity).** When buildings have few conditioned zones relative to total nodes, the majority of the graph receives severely limited gradient signals. This leads to weak learning in unconditioned zones that depend entirely on sparse policy gradient sources, where message functions may not develop strong computational dependencies, and where unconditioned zones' influence on policy decisions may be minimal in the reward structure. In extreme cases like the Small Hotel topology mentioned in prior work (14% conditioned connectivity), unconditioned zones may lack direct connections to conditioned zones entirely, resulting in complete gradient isolation and rendering graph-based learning ineffective for these nodes.
- Outdoor node gradient isolation. Outdoor nodes receive zero gradients because edges are unidirectional (Outdoor→Conditioned/Unconditioned only) by physical design, as buildings do not influence outdoor weather conditions.
 While the Outdoor→Conditioned message functions still receive gradients and can learn optimal weather processing, the outdoor node representations themselves cannot adapt. This represents a fundamental architectural choice where physical realism constrains gradient flow.
- Empirical symptoms. High return variance (unstable learning), low explained variance in the graph-based value network (indicating poor state representation), negligible attention weights on message passing edges, and worse performance than a flat MLP baseline despite the sophisticated graph architecture.
- Critical threshold hypothesis: Based on these findings, we hypothesise that effective graph-based HVAC control
 requires approximately 50-60% conditioned zone connectivity to maintain sufficient gradient density for stable learning
 across the building graph.
- **Building topology recommendations:** Our method is best suited for buildings with dense thermal coupling between conditioned zones, such as open-plan offices, educational facilities, and residential buildings where zones share significant thermal boundaries. Buildings with isolated conditioned zones (e.g., hotel rooms with individual HVAC units, data centers with isolated server zones) may not benefit from graph-based approaches due to limited inter-zone message passing opportunities. The ideal topology features bidirectional thermal connections between most conditioned zones and meaningful thermal influence from unconditioned spaces like atriums or shared mechanical areas.

HVAC-GRACE: Transferable Building Control via Heterogeneous Graph Neural Network Policies

330	Future research directions: The current architecture processes outdoor conditions through dedicated outdoor nodes
331	that cannot receive gradients due to physical constraints on edge directionality. Future research will investigate whether
332	removing outdoor nodes and directly incorporating weather features into each zone's representation improves learning
333	efficiency by providing gradient-accessible weather processing while maintaining physical interpretability of the
334	building graph structure. Additionally, we will conduct systematic investigations of graph topology effects on learning
335	performance, including quantitative analysis of the relationship between conditioned zone connectivity ratios and
336	training stability, comparative studies across diverse building archetypes with varying thermal coupling patterns, and
337	development of graph topology metrics that can predict the suitability of buildings for graph-based HVAC control
338	methods.
339	
340	
341	
342	
343	
344	
345	
346	
347	
348	
349	
350	
351	
352	
353	
354	
355	
356	
357	
359	
350	
260	
261	
262	
302	
203	
364	
305	
366	
367	
368	
369	
370	
371	
372	
373	
374	
375	
376	
377	
378	
379	
380	
381	
382	
383	
384	