
The Smart Buildings Challenge: Learning Foundation Models for Real-World Optimization at Scale

Judah Goldfeder
Columbia University
jag2396@columbia.edu

Ayush Rai
Harvard University
ayush_rai@seas.harvard.edu

Philippe M. Wyder
Distyl AI
philippe@distyl.ai

Na Li
Harvard University
nali@seas.harvard.edu

Victoria Dean
Princeton University
vdean@princeton.edu

Donna Vakalis
Mila
donna.vakalis@mila.quebec

Jiong Lin
Columbia University
jl6017@columbia.edu

Bing Dong
Syracuse University
bidong@syr.edu

Michael Rossetti
Google
mjrossetti@google.com

Nathan Kutz
Autodesk Research
kutz@uw.edu

Sohei Arisaka
Kajima
s.arisaka@kajima.com.sg

John Sipple
Google
sipple@google.com

Hod Lipson
Columbia University
hod.lipson@columbia.edu

Yoshua Bengio
Mila
yoshua.bengio@mila.quebec

Abstract

Buildings cause nearly 40% of carbon emissions in the US. Of total building emissions, HVAC systems are responsible for approximately 40-60% and account for 15% of global energy consumption. Despite advances in artificial intelligence, most buildings still operate as inefficiently as they did decades ago. Building optimization, perhaps more so than any other optimization problem, has enormous potential for Machine Learning (ML) driven impact on climate change, in both the short and long term. One of the biggest obstacles for building optimization is the lack of policy transfer. Most building control systems, whether simulation or learning based, don't generalize: they only work well for the building they were developed for, but not for other buildings. To solve this problem, we propose a challenge for the development of building foundation models that easily adapt to new buildings with minimal configuration and thereby enable building optimization to scale across the globe. Participants will build control policies on a set of simulated buildings, and have their policies evaluated on an unseen set of buildings. The buildings vary in size, layout, weather, material properties, and HVAC system. Successful outcomes from this initiative promise a massive leap forward in global sustainability efforts and demonstrate the great potential of ML to combat climate change.

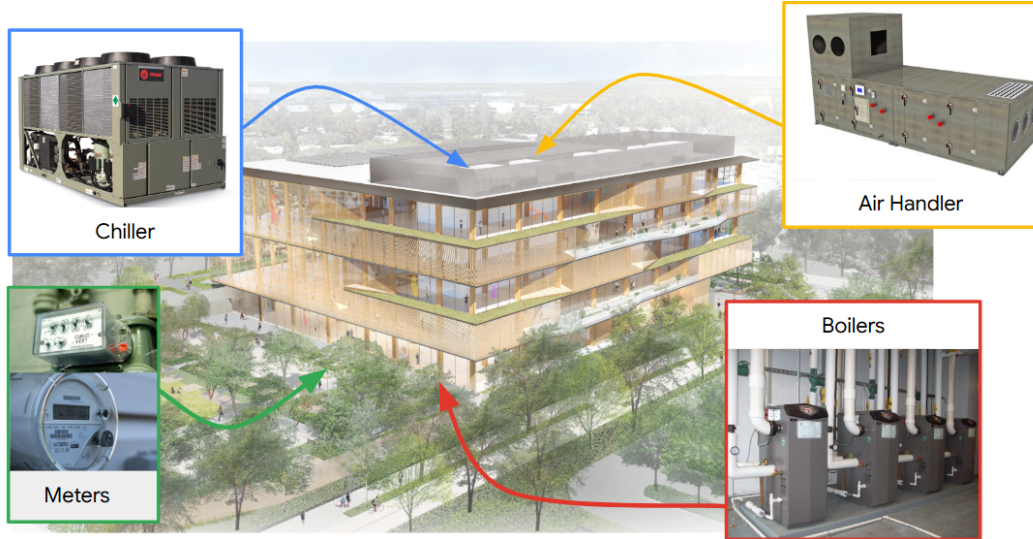


Figure 1: Illustration of an office building and its HVAC devices.

Keywords Foundation Models, Transfer Learning, Climate Change, Building Optimization.

1 Competition description

1.1 Background and impact

Commercial buildings account for 17% of total US carbon emissions [10]; within buildings, HVAC systems (see Fig. 1) are responsible for roughly 40–60% of energy use [32], or ~15% of global energy consumption [1]. Despite decades of active research [13, 39, 35, 31, 41, 23, 40, 7, 8], broad real-world adoption of optimized HVAC control remains elusive.

Current limitations. Most control policies—simulation or ML based—fail to generalize across buildings. The two primary reasons are system-specific dependencies (solutions are often tied to a particular building management system, weather pattern, HVAC system, or data ontology, which prevents them from adapting to other systems) and lack of transfer training (most controllers are trained on a single building because cross-building data is scarce and hard to come by).

Competition goal. A building foundation control model that could easily adapt to new buildings with minimal configuration would enable building optimization to scale to thousands of buildings with significantly less overhead—lowering the barrier of adoption. Recent evidence suggests that such a building model is within reach. Foundation models in other domains have demonstrated remarkable results in transfer learning, from vision[2] and language[30] to robotics[12] and control[47], and progress is especially bolstered when there is community support behind a well adopted benchmark[20, 46], something we are hoping to foster within the building science and ML community (a successful recent example of this in the related ML and sustainability community is the AI City Challenge[38]). This competition is trying to answer whether we can create a building foundation model that effectively transfers across many buildings. We will use a simulated environment for the contest, and will evaluate winning solutions on real buildings after the contest ends.

Impact. A scalable building model that unlocks even a 5% reduction in carbon emissions could reduce the total carbon emissions of the US by roughly 1%. This is especially true if steps are taken to ease public adoption[16]. Additionally, a building foundation model can lower costs of adopting optimal control mechanisms by removing the need for hard-to-calibrate simulations. Once adoption costs are lowered, widespread industry adoption can occur. We expect this competition to be of interest to the broader community, no domain knowledge will be needed to participate.

1.2 Related Work

HVAC control has garnered considerable interest[13] in recent years [26], and while alternatives like Model Predictive Control (MPC) exist[37], a growing portion of the literature has considered Reinforcement Learning [49, 29, 48, 14, 43, 42, 51, 11, 16]. Of course, this requires a simulation and many works have explored this topic[39, 35, 31, 41, 23, 40, 7, 25]. EnergyPlus [8], a high-fidelity simulator developed by the Department of Energy, is commonly used [45, 3, 53, 44, 5], but suffers from scalability and configuration challenges. To overcome the limitations of these methods, some work uses a hybrid approach [52, 4, 21, 50, 27, 9, 24], as does our simulator.

1.3 Novelty

To our knowledge, no previous competition has focused on foundation models for buildings, that is, on creating digital twin models that transfer from building to building. This is arguably the largest unsolved problem in this entire domain of research, and has been a major focus of recent work[6, 36, 33, 34, 28].

1.4 Data

The competition makes use of the HVAC simulation platform[21, 15] from the Smart Buildings Control Suite[17], along with some improvements from the open source community[22]. For the contest we created a random building generator that interfaces with the simulator. This allows for an infinite set of potential building simulators, each with different conditions. The generator varies the number of floors (2–4), width and length, perimeter shape, room layout, occupancy, envelope construction (six material properties), and the number of Air Handler Units (AHUs) and the zones they serve (at most 10 zones per AHU); the hot water system remains a single system per building. Each zone has its own VAV box. Each building is exposed as a Gym-compatible RL environment.

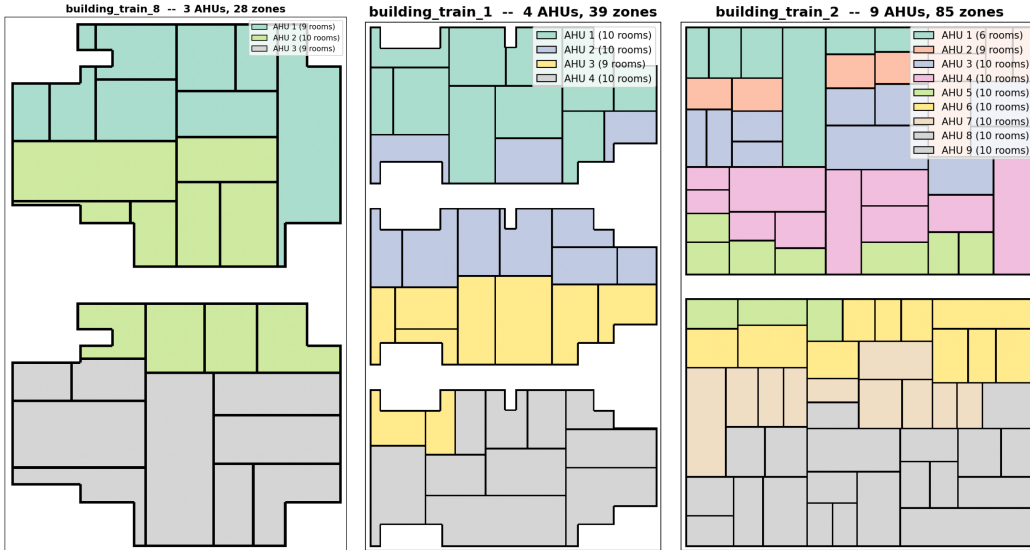


Figure 2: Three training buildings. Floor plan, zone count, and AHU-to-zone assignment differ per building (colors show the AHU serving each zone).

Building split. Participants will be given 6 training buildings, each paired with one recorded weather strip from each of five climate stations (30 training scenarios). The agent will be evaluated on 6 unseen buildings from the same generator and parameter ranges, at 5 climate stations that appear nowhere in the training set, so the transfer is across both buildings and climates.

Availability and licensing. The simulated buildings and corresponding weather scenarios are generated procedurally by an open-source generator built on the Smart Buildings Control Suite [17].

They contain no personally identifiable information, no real building telemetry, and no human subjects data, and therefore comply fully with the NeurIPS Code of Ethics on privacy, consent, and fair use. The generator, the 6 development buildings, and the training weather strips will be released under an open-source license alongside the starter kit at competition launch. The 6 held-out evaluation buildings and the 5 held-out stations are kept confidential by the Organizers and released publicly only after the competition concludes.

Scale and statistical power. Because the environment is procedural, the effective dataset is unbounded: participants may sample arbitrarily many training buildings, and Organizers can sample fresh evaluation buildings to defeat memorization. The default training pool of 6 buildings $\times$ 5 stations yields 30 development scenarios; the evaluation set is 6 unseen buildings $\times$ 5 unseen stations, giving 30 scored rollouts per submission. The simulator is deterministic, so a single rollout per pair carries no sampling noise and repeated rollouts would add nothing (1.6); breadth across buildings and climates, rather than repetition, is what supports conclusions about transfer.

Ground-truth confidentiality. The 6 held-out building configurations are generated from random seeds known only to the Organizers and stored in a private repository. No artifact derived from these buildings — floorplans, AHU divisions, baseline returns — is released until after the competition closes.

Feature leakage. Three mechanisms address leakage: (i) the train/evaluation split is across building configurations rather than across time within a single building, so participants cannot memorize per-building physics; (ii) evaluation buildings are generated from seeds withheld from participants and never appear in any released artifact prior to the competition close; (iii) the policy receives the action spec, floorplan, and AHU divisions of each evaluation building only at runtime inside the evaluation sandbox, so they cannot be inspected, indexed, or used to pre-compile per-building heuristics during training.

1.5 Tasks and application scenarios

Real-world scenario. This task mirrors managing a diverse real-estate portfolio where per-building tuning is too costly and risky. A successful policy acts as a universal controller that adapts to unseen buildings using only floorplans and HVAC inventory, enabling scalable, portfolio-wide building optimization.

Scientific challenge. Varying building dynamics prevent overfitting, yet shared physics and established baselines (1.7) make the problem tractable. The core hurdle is transferability: participants must learn a single controller that successfully conditions its decisions on the specific building it is operating: a foundation model for smart buildings. An open question that we also hope to explore is, whether domain specific models will win out[19], or whether generic foundation models will generalize best.

Ethics. The competition uses purely simulated environments with no human subjects or personal data. Post-contest live trials follow strict institutional safety protocols, ensuring participants never directly control physical equipment. The project aims solely to reduce carbon emissions and poses no foreseeable societal harm.

Action space

Each environment has a hybrid action with discrete and continuous parts. Every AHU and the Hot Water System (HWS) has one discrete run command (on/off). Each AHU has three continuous setpoints: supply-air temperature, static pressure (duct airflow), and recirculation fraction. The HWS has two: supply-water temperature and pump differential pressure. Each VAV box (one per zone) has two: damper position and reheat valve position. Since an AHU sends one supply-air temperature to all its zones, the dampers and reheat valves are the only per-zone levers. Discrete commands are in $\{0, 1\}$; continuous setpoints are normalized to $[-1, 1]$. For a building with K AHUs and V zones, the action is a dict with `discrete_action` of shape $(K+1,)$ and `continuous_action` of shape $(3K+2+2V,)$. The simulator enforces ventilation floors (minimum static pressure for a running

AHU, at most 70% recirculation, dampers at least 10% open), so shutting all air off is not an energy strategy. Per-index mappings are given in the building metadata.

Observation Space

Each environment returns a flat observation vector whose length depends on the building. It contains, per zone, air temperature, heating/cooling setpoints, damper and reheat positions, and airflow setpoint; per AHU, supply-air temperature, static pressure, airflow, recirculation, and fan status; the HWS supply-water temperature; building occupancy and comfort-schedule flags; outside air temperature; and time of day and day of week. The policy also receives building metadata: action indices, observation field names, the AHU-to-zone assignment, floor plan, schedule, and normalization constants.

For each unseen building in the evaluation set, the floorplan, AHU divisions, and action spec are delivered to the policy at runtime as part of the observation (see 1.4).

Policy API

Contestants submit a Python class with two methods, subclassing the policy base class supplied by the evaluation gateway, and hand that class to the inference server:

```
from kaggle_evaluation import smart_buildings_inference_server

class MyPolicy(smart_buildings_inference_server.Policy):
    def __init__(self, building_meta_data, action_spec):
        super().__init__(building_meta_data, action_spec)
        ...
    def get_action(self, obs) -> dict: ...

smart_buildings_inference_server.SmartBuildingsInferenceServer(MyPolicy).run()
```

The constructor is called once per building with its metadata (floor plan, AHU-to-zone mapping, action and observation indices) and `action_spec` (keys, shapes, and ranges). `get_action` maps the current observation to an action dict matching `action_spec` exactly. The class name is the participant's own choice; only the two signatures are fixed. Submissions take the form of a Kaggle notebook; the full submission instructions and the current template notebook are given on the Kaggle competition submission page.

1.6 Metrics

Evaluation uses the regret-based reward, at each 5-minute step t ,

$$r_t = w_p p_t - w_c c_t - w_e e_t, \quad (w_p, w_c, w_e) = (0.6, 0.2, 0.2).$$

The reward is a weighted sum of three terms: p_t is occupant comfort (productivity), c_t is the energy cost (electricity and gas), and e_t is the carbon emission. The comfort term drops as zone temperatures leave the comfort band; the drop is linear up to 5 K outside the band and quadratic and unbounded beyond, so abandoning comfort is never free. It is summed over zones weighted by each zone's own occupancy rather than its floor area, so a single badly conditioned zone can dominate an episode. Weights are fixed by the Organizers and identical for all submissions. Each scenario runs for 30 days at a 5-minute timestep and every step is scored, so the episodic return is $R^{(s)}(\pi) = \sum_{t=1}^T r_t^{(s)}$ with $T = 8640$.

The reward is bounded above by zero. All three terms are penalties. The comfort term p_t is a regret: it is 0 when every occupied zone sits inside its comfort band and falls negative otherwise, without bound. Energy cost c_t and carbon emission e_t are non-negative quantities entering with a negative sign, since the building cannot export energy. Hence $r_t \leq 0$ always, with equality only when comfort is perfect and the plant draws nothing. Measured over a full 8640-step episode, the largest per-step reward observed is exactly 0, attained at 95 steps, all of them with the building unoccupied and the plant idle.

Leaderboard score. Let $\bar{R}^{(s)}(\pi) = R^{(s)}(\pi)/T$ be the mean per-step reward on scenario s . The leaderboard metric is the mean absolute error of these values against a target of zero,

$$\text{MAE}(\pi) = \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} |\bar{R}^{(s)}(\pi) - 0| = -\frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} \bar{R}^{(s)}(\pi),$$

where $\mathcal{S}$ is the set of evaluation scenarios. **Lower is better.** The second equality holds because $r_t \leq 0$, so the metric is simply the mean magnitude of the per-step loss. The target of zero is not an arbitrary reference point: it denotes perfect comfort delivered at no energy and no carbon cost, which is the unattainable ideal of the control problem. Because the transformation is monotone, minimizing MAE is identical to maximizing the return on every individual scenario.

A submitted policy is evaluated on each of the 6 held-out buildings at each of 5 weather stations, giving 30 scenarios. The simulator is deterministic: the same policy on the same scenario reproduces the same return bit for bit, across processes and machines, so each (building, weather) pair is run exactly once and no averaging over repeated rollouts is required. The headline leaderboard score is the mean over all 30 scenarios, with no grouping by building or weather.

The Organizers re-evaluate the top 7 submitted policies end-to-end at a fixed simulator commit before announcing final standings; because the simulator is deterministic, this re-run verifies rule compliance rather than estimating variance.

1.7 Baselines

Baselines. Building upon earlier work[18], we ship four hand-written reference controllers, all of which run the plant: **Bang-Bang** (`bang_bang`) switches each AHU on or off around the comfort band; **Fixed Setpoint** (`fixed_setpoint`) holds one comfort band around the clock with no night setback; **PID** (`pid`) uses proportional control with fan pressure scaled by control effort; and **PID + Water** (`pid_water`) adds modulating reheat and a demand-driven hot-water reset. Each is a controller a practitioner might plausibly write. An **All Off** policy is shipped as an additional reference point but is not counted among the four, because it does not run the plant at all. The baselines do not enter the leaderboard metric, which needs no reference policies; they give participants a published performance range to develop against, and their returns on the public training scenarios are shipped with the starter kit. On those scenarios `bang_bang` sets the lower end of the range, while the upper end is set by `pid_water` or, in hot and humid weather, by `fixed_setpoint`.

Each scenario runs for 30 days, all of which are scored. No single baseline is best everywhere, which confirms the transfer challenge. For each policy on each scenario we produce the overall return and a plot charting zone temperature, return, occupancy, and reward-component contributions over time (Figure 3); a video rendering of the building over the scenario can be produced on request.

2 Organizational aspects

2.1 Protocol

The competition will be hosted on Kaggle, with the prize pool jointly sponsored by Kajima, Distyl, and KODE Labs. These will be collectively referred to as Organizers from here on. A Participant or group of Participants (a “Team”, up to four members) submits an Entry consisting of (i) a Python module implementing the `TransferPolicy` interface, (ii) the model weights or other artifacts needed to instantiate it, and (iii) a reproducibility bundle (code archive, seed list, dependency manifest). The latest valid Submission per Team is the official Entry. Submissions are sandboxed and executed by the Organizers against the held-out 6-building evaluation suite via the `evaluate.py` interface.

Two-phase structure. The competition runs in two phases. During the **development phase**, the Kaggle public leaderboard scores submissions against a rotating, publicly-visible subset of evaluation (building, weather) pairs; this gives participants signal without exposing the full held-out set. During the **final phase**, all submitted policies are re-evaluated against the complete held-out set (1.6), and final standings are determined by the leaderboard metric of 1.6, mean absolute error against zero, on the private leaderboard, where lower is better. The public/private split mitigates overfitting to the visible evaluation signal.

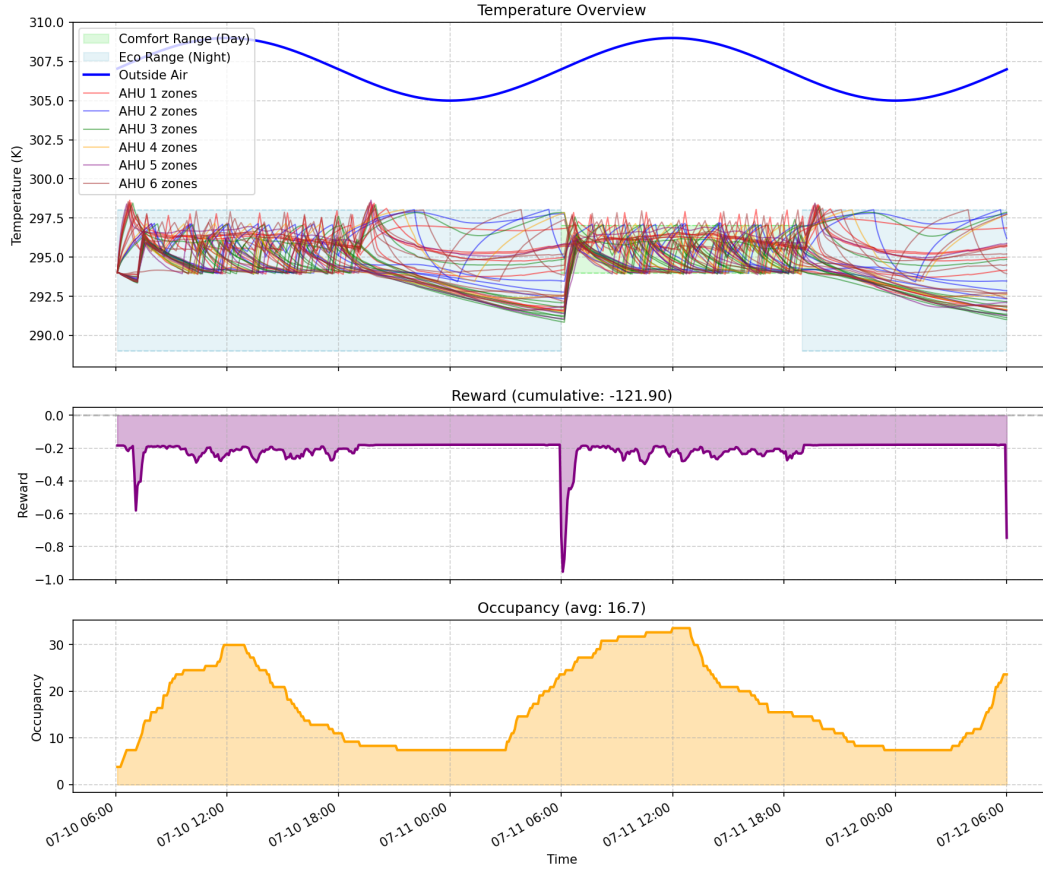


Figure 3: One rollout: zone temperatures against the comfort band, per-step reward, and occupancy.

Anti-cheating and fairness.

1. The held-out 30 scenarios are never released to participants, in any form, before the competition closes.
2. Policies are executed in an isolated container at evaluation time with no network access, blocking external lookup tables or remote inference.
3. Per-rollout wall-clock is bounded at $10\times$ the bang-bang baseline runtime so solutions cannot succeed by unbounded online planning.
4. Submitted code archives are scanned for hard-coded building IDs or sensor signatures before scoring; matches are grounds for disqualification.

Evaluation harness. Participants develop against `evaluate.py`, the harness shipped in the starter kit: a set of (building, weather) configs is specified and each is rolled out once for $T = 8640$ steps, after which `score_leaderboard.py` reports the leaderboard metric of 1.6 together with the shipped baseline returns for comparison, so the four baselines need not be re-run. Scoring on the leaderboard is performed instead by an inference-server gateway, which owns the simulator, iterates the held-out scenarios and queries the submitted policy once per step; a submission therefore contains no scenario list and no rollout loop. Reference policies (the four baselines from 1.7), the public training scenarios with their baseline returns, and example invocations are shipped in the starter kit (<https://github.com/smart-buildings-2026/starter-kit>).

2.2 Rules and Engagement

The Challenge will be run in accordance with the rules and regulations posted on the Kaggle page.

2.3 Schedule and readiness

Our competition preparation and execution timeline is outlined below.

Date	Task / Event
Nov 13, 2026	Round One deadline (NeurIPS poster; \$2K in prizes)
Dec 13, 2026	Paris Workshop
Jan 15, 2027	Final deadline (\$38K in prizes)

2.4 Competition promotion and incentives

Incentives. The \$40,000 total prize pool will be split across podium positions on the private leaderboard. Beyond monetary awards, top teams will receive: (i) an invited podium talk at the NeurIPS 2026 Competition Track workshop; (ii) co-authorship on the consortium write-up of the competition results (named-author status for teams contributing a methods description that meets the standard threshold; consortium-author status otherwise); (iii) invitation to the post-contest live-trial program at the four physical partner sites.

References

- [1] Nilofar Asim, Marzieh Badiei, Masita Mohammad, Halim Razali, Armin Rajabi, Lim Chin Haw, and Mariyam Jameelah Ghazali. Sustainability of heating, ventilation and air-conditioning (hvac) systems in buildings—an overview. International journal of environmental research and public health, 19(2):1016, 2022.
- [2] Muhammad Awais, Muzammal Naseer, Salman Khan, Rao Muhammad Anwer, Hisham Cholakkal, Mubarak Shah, Ming-Hsuan Yang, and Fahad Shahbaz Khan. Foundation models defining a new era in vision: a survey and outlook. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2025.
- [3] Donald Azuatalam, Wee-Lih Lee, Frits de Nijs, and Ariel Liebman. Reinforcement learning for whole-building hvac control and demand response. Energy and AI, 2:100020, 2020.
- [4] Yasaman Balali, Adrian Chong, Andrew Busch, and Steven O’Keefe. Energy modelling and control of building heating and cooling systems with data-driven and hybrid models—a review. Renewable and Sustainable Energy Reviews, 183:113496, 2023.
- [5] Mangesh Basarkar. Modeling and simulation of hvac faults in energyplus. 2011.
- [6] Anaïs Berkes, Yoshua Bengio, David Rolnick, and Donna Vakalis. A hot dataset: 150,000 buildings for hvac operations transfer research. In Proceedings of the 12th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, pages 171–180, 2025.
- [7] Michael Blonsky, Jeff Maguire, Killian McKenna, Dylan Cutler, Sivasathya Pradha Balamurugan, and Xin Jin. Ochre: The object-oriented, controllable, high-resolution residential energy model for dynamic integration studies. Applied Energy, 290:116732, 2021.
- [8] Drury B Crawley, Linda K Lawrie, Frederick C Winkelmann, Walter F Buhl, Y Joe Huang, Curtis O Pedersen, Richard K Strand, Richard J Liesen, Daniel E Fisher, Michael J Witte, et al. Energyplus: creating a new-generation building energy simulation program. Energy and buildings, 33(4):319–331, 2001.
- [9] Ján Drgoňa, Aaron R Tuor, Vikas Chandan, and Draguna L Vrabie. Physics-constrained deep learning of multi-zone building thermal dynamics. Energy and Buildings, 243:110992, 2021.
- [10] EIA. Frequently Asked Questions (FAQs) - U.S. Energy Information Administration (EIA) — eia.gov. <https://www.eia.gov/tools/faqs/faq.php?id=86&t=1>, 2024. [Accessed 04-06-2024].

- [11] Xi Fang, Guangcai Gong, Guannan Li, Liang Chun, Pei Peng, Wenqiang Li, Xing Shi, and Xiang Chen. Deep reinforcement learning optimal control strategy for temperature setpoint real-time reset in multi-zone building hvac system. Applied Thermal Engineering, 212:118552, 2022.
- [12] Roya Firoozi, Johnathan Tucker, Stephen Tian, Anirudha Majumdar, Jiankai Sun, Weiyu Liu, Yuke Zhu, Shuran Song, Ashish Kapoor, Karol Hausman, et al. Foundation models in robotics: Applications, challenges, and the future. The International Journal of Robotics Research, page 02783649241281508, 2023.
- [13] Kwong Fai Fong, Victor Ian Hanby, and Tin-Tai Chow. Hvac system optimization for energy management by evolutionary programming. Energy and buildings, 38(3):220–231, 2006.
- [14] Cheng Gao and Dan Wang. Comparative study of model-based and model-free reinforcement learning control performance in hvac systems. Journal of Building Engineering, 74:106852, 2023.
- [15] Judah Goldfeder and John Sipple. Real-world data and calibrated simulation suite for offline training of reinforcement learning agents to optimize energy and emission in buildings for environmental sustainability. 2024.
- [16] Judah Goldfeder and John Sipple. Reducing carbon emissions at scale: Interpretable and efficient to implement reinforcement learning via policy extraction. In Proceedings of the 11th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, pages 403–407, 2024.
- [17] Judah Goldfeder, Victoria Dean, Zixin Jiang, Xuezheng Wang, Bing dong, Hod Lipson, and John Sipple. The smart buildings control suite: A diverse open source benchmark to evaluate and scale hvac control policies for sustainability, 2025. URL <https://arxiv.org/abs/2410.03756>.
- [18] Judah Goldfeder, Gabriel Guerra Trigo, Philippe Martin Wyder, Neil Kachappilly, and Hod Lipson. Evaluating performance of reinforcement learning agents to control buildings efficiently. In 2025 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), pages 2903–2909. IEEE, 2025.
- [19] Judah Goldfeder, Philippe Wyder, Yann LeCun, and Ravid Shwartz Ziv. Ai must embrace specialization via superhuman adaptable intelligence. arXiv preprint arXiv:2602.23643, 2026.
- [20] Judah Goldfeder, Philippe M Wyder, J Nathan Kutz, and David Donoho. Driving science faster and more credibly with the common task framework. Computing in Science & Engineering, 2026.
- [21] Judah A Goldfeder and John A Sipple. A lightweight calibrated simulation enabling efficient offline learning for optimal control of real buildings. In Proceedings of the 10th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, pages 352–356, 2023.
- [22] Sang woo Ham, Donghun Kim, Michael Rossetti, and John Sipple. Tensorized radiative heat transfer for a scalable and calibrated building energy simulator. In Proceedings of the 13th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, pages 230–233, 2026.
- [23] A Husaunndee, R Lahrech, H Vaezi-Nejad, and JC Visier. Simbad: A simulation toolbox for the design and test of hvac control systems. In Proceedings of the 5th international IBPSA conference, volume 2, pages 269–276. International Building Performance Simulation Association (IBPSA) Prague . . . , 1997.
- [24] Zixin Jiang, Xuezheng Wang, Han Li, Tianzhen Hong, Fengqi You, Ján Drgoňa, Draguna Vrabie, and Bing Dong. Physics-informed machine learning for building performance simulation-a review of a nascent field. Advances in Applied Energy, 18:100223, 2025.

- [25] Zixin Jiang, Ruizhi Song, Guowen Li, Yuhang Zhang, Zheng O'Neill, Xuezheng Wang, Judah Goldfeder, and Bing Dong. Bestopt: A modular, physics-informed machine learning based building modeling, control and optimization framework. arXiv preprint arXiv:2601.16283, 2026.
- [26] Dongsu Kim, Jongman Lee, Sunglok Do, Pedro J Mago, Kwang Ho Lee, and Heejin Cho. Energy modeling and model predictive control for hvac in buildings: a review of current research trends. Energies, 15(19):7231, 2022.
- [27] Peter Klanatsky, François Veynandt, and Christian Heschl. Grey-box model for model predictive control of buildings. Energy and Buildings, 300:113624, 2023.
- [28] Thomas Krug, Fabian Raisch, Dominik Aimer, Markus Wirnsberger, Ferdinand Sigg, Benjamin Schäfer, and Benjamin Tischler. Builda: A thermal building data generation framework for transfer learning. In 2025 Annual Modeling and Simulation Conference (ANNSIM), pages 1–13. IEEE, 2025.
- [29] Karl Mason and Santiago Grijalva. A review of reinforcement learning for autonomous building energy management. Computers & Electrical Engineering, 78:300–312, 2019.
- [30] Devon Myers, Rami Mohawesh, Venkata Ishwarya Chellaboina, Anantha Lakshmi Sathvik, Praveen Venkatesh, Yi-Hui Ho, Hanna Henshaw, Muna Alhawawreh, David Berdik, and Yaser Jararweh. Foundation and large language models: fundamentals, challenges, opportunities, and social impacts. Cluster Computing, 27(1):1–26, 2024.
- [31] Cheol Park, Daniel R Clark, and George E Kelly. An overview of hvacsim+, a dynamic building/hvac/control systems simulation program. In Proceedings of the 1st Annual Building Energy Simulation Conference, Seattle, WA, pages 21–22, 1985.
- [32] Luis Pérez-Lombard, José Ortiz, and Christine Pout. A review on buildings energy consumption information. Energy and buildings, 40(3):394–398, 2008.
- [33] Fabian Raisch, Thomas Krug, Christoph Goebel, and Benjamin Tischler. Gentl: A general transfer learning model for building thermal dynamics. In Proceedings of the 16th ACM International Conference on Future and Sustainable Energy Systems, pages 322–333, 2025.
- [34] Fabian Raisch, Max Langtry, Felix Koch, Ruchi Choudhary, Christoph Goebel, and Benjamin Tischler. Adapting to change: A comparison of continual and transfer learning for modeling building thermal dynamics under concept drifts. Energy and Buildings, page 116868, 2025.
- [35] Peter Riederer. Matlab/simulink for building and hvac simulation-state of the art. In Ninth International IBPSA Conference, pages 1019–1026, 2005.
- [36] Vincent Taboga, Justin Veilleux, Doseok Jang, Anushree Rankawat, and Pierre-Luc Bacon. Building2building: A large scale benchmark for generalizable real-world reinforcement learning. arXiv preprint arXiv:2607.16534, 2026.
- [37] Saman Taheri, Paniz Hosseini, and Ali Razban. Model predictive control of heating, ventilation, and air conditioning (hvac) systems: A state-of-the-art review. Journal of Building Engineering, 60:105067, 2022.
- [38] Zheng Tang, Shuo Wang, David C Anastasiu, Ming-Ching Chang, Anuj Sharma, Quan Kong, Munkhjargal Gochoo, Jun-Wei Hsieh, Tomasz Kornuta, Zhedong Zheng, et al. The 10th ai city challenge. arXiv preprint arXiv:2608.17044, 2026.
- [39] Marija Trčka and Jan LM Hensen. Overview of hvac system simulation. Automation in construction, 19(2):93–99, 2010.
- [40] Marija Trcka, Michael Wetter, and JLM Hensen. Comparison of co-simulation approaches for building and hvac/r system simulation. In 10th International IBPSA Building Simulation Conference (BS 2007), September 3-6, 2007, Beijing, China, pages 1418–1425, 2007.
- [41] Marija Trčka, Jan LM Hensen, and Michael Wetter. Co-simulation of innovative integrated hvac systems in buildings. Journal of Building Performance Simulation, 2(3):209–230, 2009.

- [42] José R Vázquez-Canteli and Zoltán Nagy. Reinforcement learning for demand response: A review of algorithms and modeling techniques. Applied energy, 235:1072–1089, 2019.
- [43] Marshall Wang, John Willes, Thomas Jiralerspong, and Martin Moezzi. A comparison of classical and deep reinforcement learning methods for hvac control. arXiv preprint arXiv:2308.05711, 2023.
- [44] Mubashir Wani, Akshya Swain, and Abhisek Ukil. Control strategies for energy optimization of hvac systems in small office buildings using energyplus tm. In 2019 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia), pages 2698–2703. IEEE, 2019.
- [45] Tianshu Wei, Yanzhi Wang, and Qi Zhu. Deep reinforcement learning for building hvac control. In Proceedings of the 54th annual design automation conference 2017, pages 1–6, 2017.
- [46] Philippe Wyder, Judah Goldfeder, Alexey Yermakov, Yue Zhao, Stefano Riva, Jan Williams, David Zoro, Amy Rude, Matteo Tomasetto, Joe Germany, et al. Common task framework for a critical evaluation of scientific machine learning algorithms. Advances in Neural Information Processing Systems, 38, 2026.
- [47] Sherry Yang, Ofir Nachum, Yilun Du, Jason Wei, Pieter Abbeel, and Dale Schuurmans. Foundation models for decision making: Problems, methods, and opportunities. arXiv preprint arXiv:2303.04129, 2023.
- [48] Liang Yu, Yi Sun, Zhanbo Xu, Chao Shen, Dong Yue, Tao Jiang, and Xiaohong Guan. Multi-agent deep reinforcement learning for hvac control in commercial buildings. IEEE Transactions on Smart Grid, 12(1):407–419, 2020.
- [49] Liang Yu, Shuqi Qin, Meng Zhang, Chao Shen, Tao Jiang, and Xiaohong Guan. A review of deep reinforcement learning for smart building energy management. IEEE Internet of Things Journal, 8(15):12046–12063, 2021.
- [50] Chi Zhang, Yuanyuan Shi, and Yize Chen. Bear: Physics-principled building environment for control and reinforcement learning. In Proceedings of the 14th ACM International Conference on Future Energy Systems, pages 66–71, 2023.
- [51] Zhiang Zhang, Adrian Chong, Yuqi Pan, Chenlu Zhang, and Khee Poh Lam. Whole building energy model for hvac optimal control: A practical framework based on deep reinforcement learning. Energy and Buildings, 199:472–490, 2019.
- [52] Huan Zhao, Junhua Zhao, Ting Shu, and Zibin Pan. Hybrid-model-based deep reinforcement learning for heating, ventilation, and air-conditioning control. Frontiers in Energy Research, 8: 610518, 2021.
- [53] Jie Zhao, Khee Poh Lam, B Erik Ydstie, and Omer T Karaguzel. Energyplus model-based predictive control within design–build–operate energy information modelling infrastructure. Journal of Building Performance Simulation, 8(3):121–134, 2015.