Online Comfort-Constrained HVAC Control via Feature Transfer

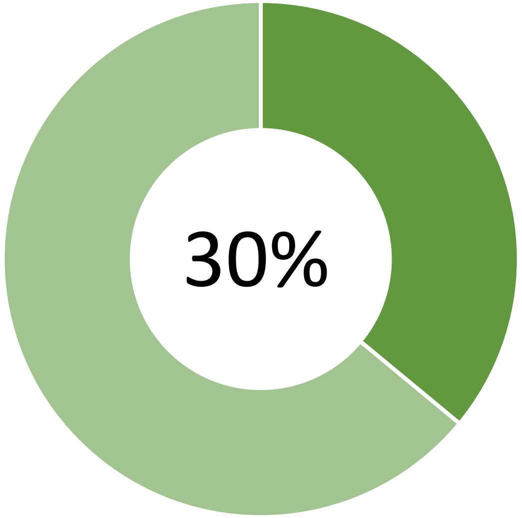
Jing Yu, Tianyu Zhang, Omid Ardakanian, Adam Wierman

ACM e-Energy

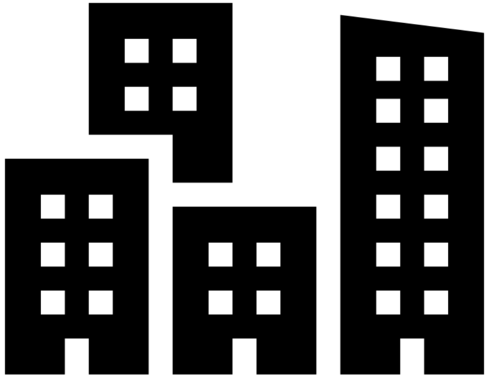
June 2025

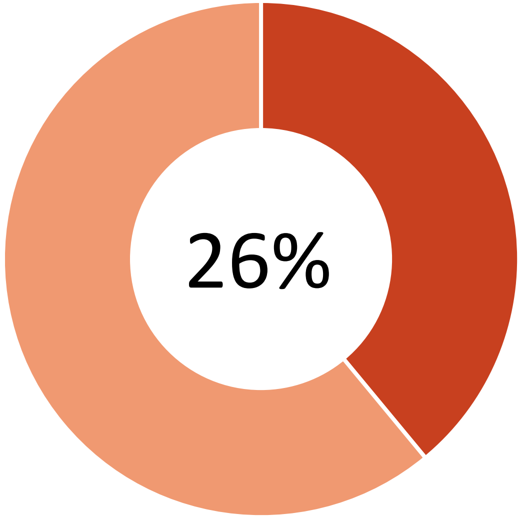
Buildings are a Major Driver of Global Energy Use and Emissions

The operations of buildings account for



of global final energy consumption





of global energy-related carbon emissions

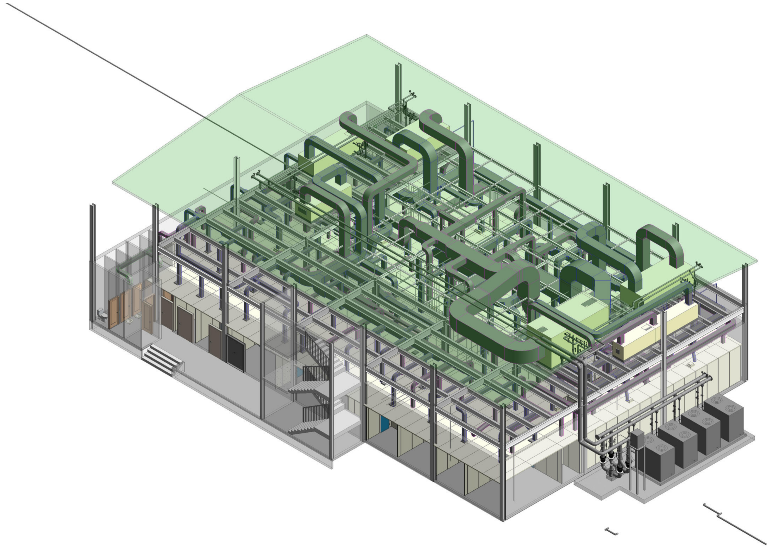
https://www.iea.org/energy-system/buildings

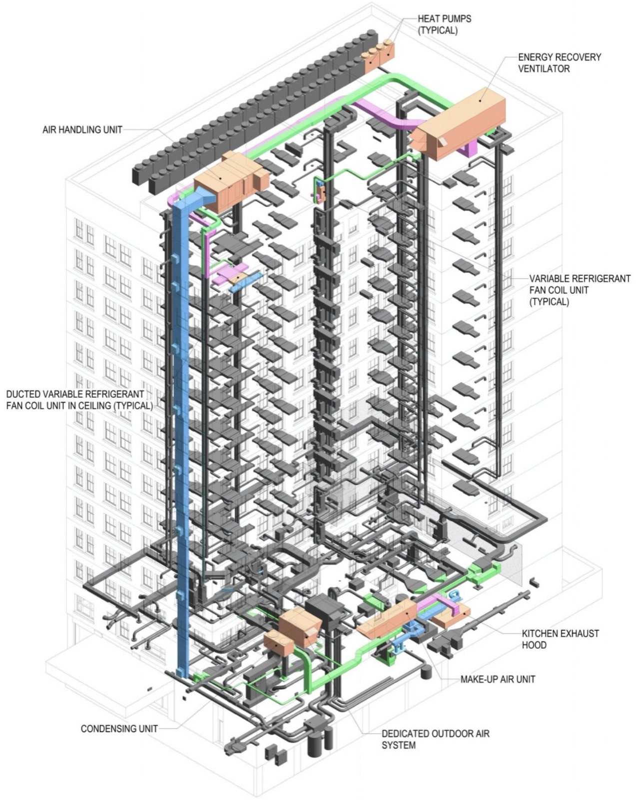
Buildings are a Major Driver of Global Energy Use and Emissions

HVAC is responsible for a significant proportion of a building's energy consumption

Learning-Based Strategies Yield Major Energy Savings

But they require extensive offline data or costly online interaction





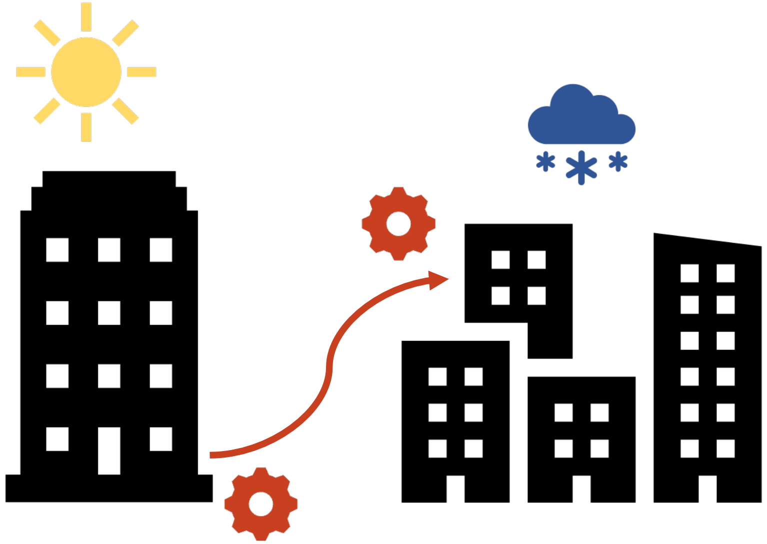
Can we transfer and reuse a model (or a control policy) trained for a different building?

Transfer Learning for HVAC Control

Learn relevant features from source building(s) and adapt them to target building(s)

Pros:

- Easy to adapt a dynamical model (or a control policy) to real-time conditions in a new building
- Significant reduction in energy consumption can be readily achieved [1]



Cons:

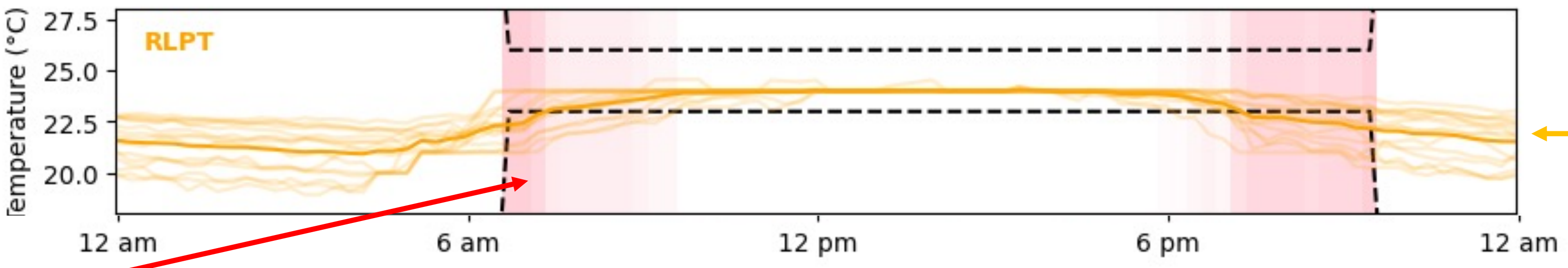
- Requires large and representative datasets to build high-quality candidate system models (or candidate control policies)
- Difficult to enforce the new building's constraints

[1] Aakash Krishna, Tianyu Zhang, Omid Ardakanian, Matthew E. Taylor, "Mitigating an Adoption Barrier of Reinforcement Learning-Based Control Strategies in Buildings", Energy and Buildings, vol.285, pp.112878, April 2023

We proposed an online algorithm to augment a black-box policy with safety

A glimpse of the empirical result

Thermal comfort constraints



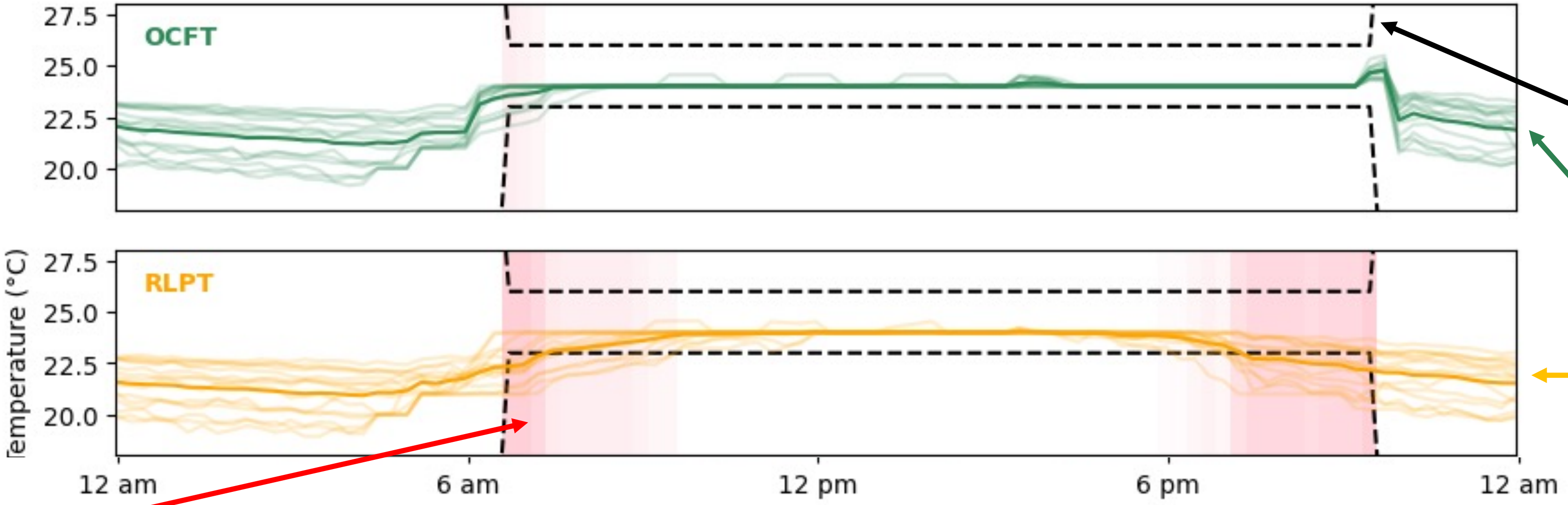
Temperature profile averaged over many days

constraint violation

Behavior of control algorithms over a single day

A glimpse of the empirical result

Proposed method significantly reduces constraint violation while minimizing energy consumption



Thermal comfort constraints

Temperature profile averaged over many days

constraint violation

Behavior of control algorithms over a single day

OCFT: Online Comfort-Constrained HVAC Control with Feature Transfer

1. **Feature adaptation** + **augmentation** for transfer of **black-box** policies
   - Tunable trade-off between energy efficiency & thermal comfort violations
2. **Theoretical** insights
   - Finite constraint violations & finite-time model learning guarantees
3. **Empirical evaluations** in EnergyPlus
   - Comprehensive experiments in 19 U.S. climate zones
   - OCFT reduces constraint violations by 81.28% while using 11.23% more energy

System Model

Linear combination of nonlinear feature functions

- **State-space Representation:**

system parameter

features learnt in source building

zone temperature

$$s_{t+1} = (\theta^*)^\top \Phi(s_t, z_t, a_t) + d_t$$

disturbances

measurable states

- **Thermal Comfort Constraints:**

$$\underline{T_t} \leq s_t \leq \overline{T_t}$$

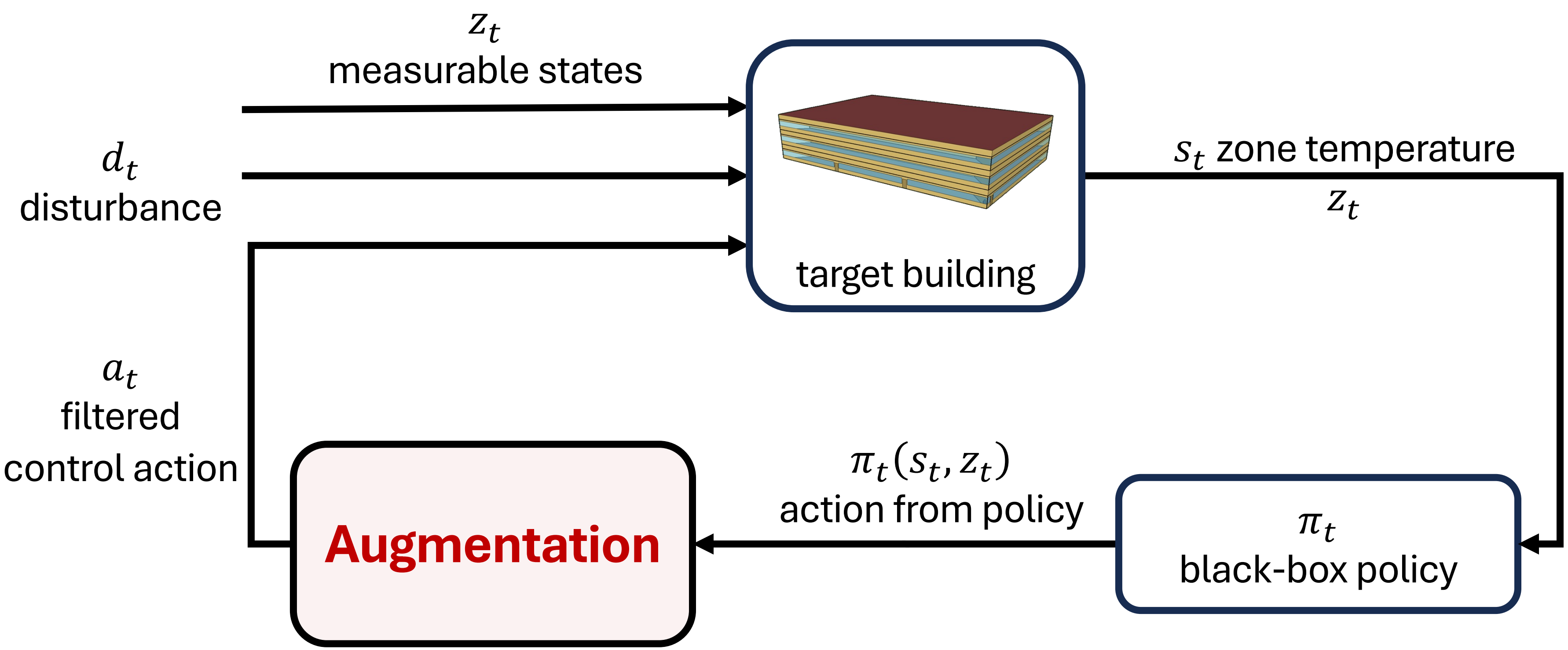
- **Exogenous Disturbances:**

$$\|d_t\|_\infty \leq W$$

Goal: **minimize energy** consumption subject to **thermal comfort constraints**

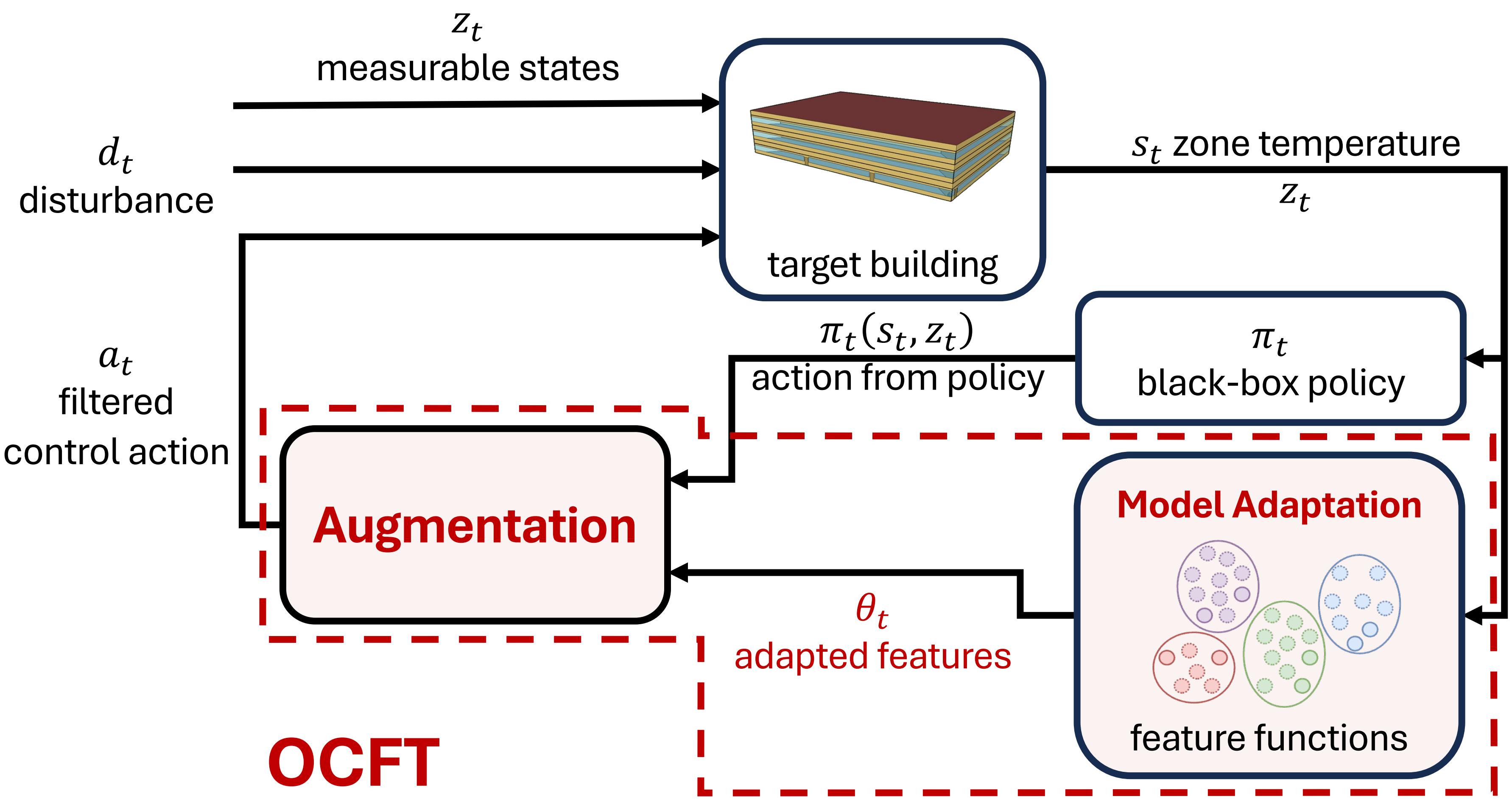
Approach

Augmentation of black-box control policy with safety



OCFT Framework

An overview



OCFT Framework

Warm start

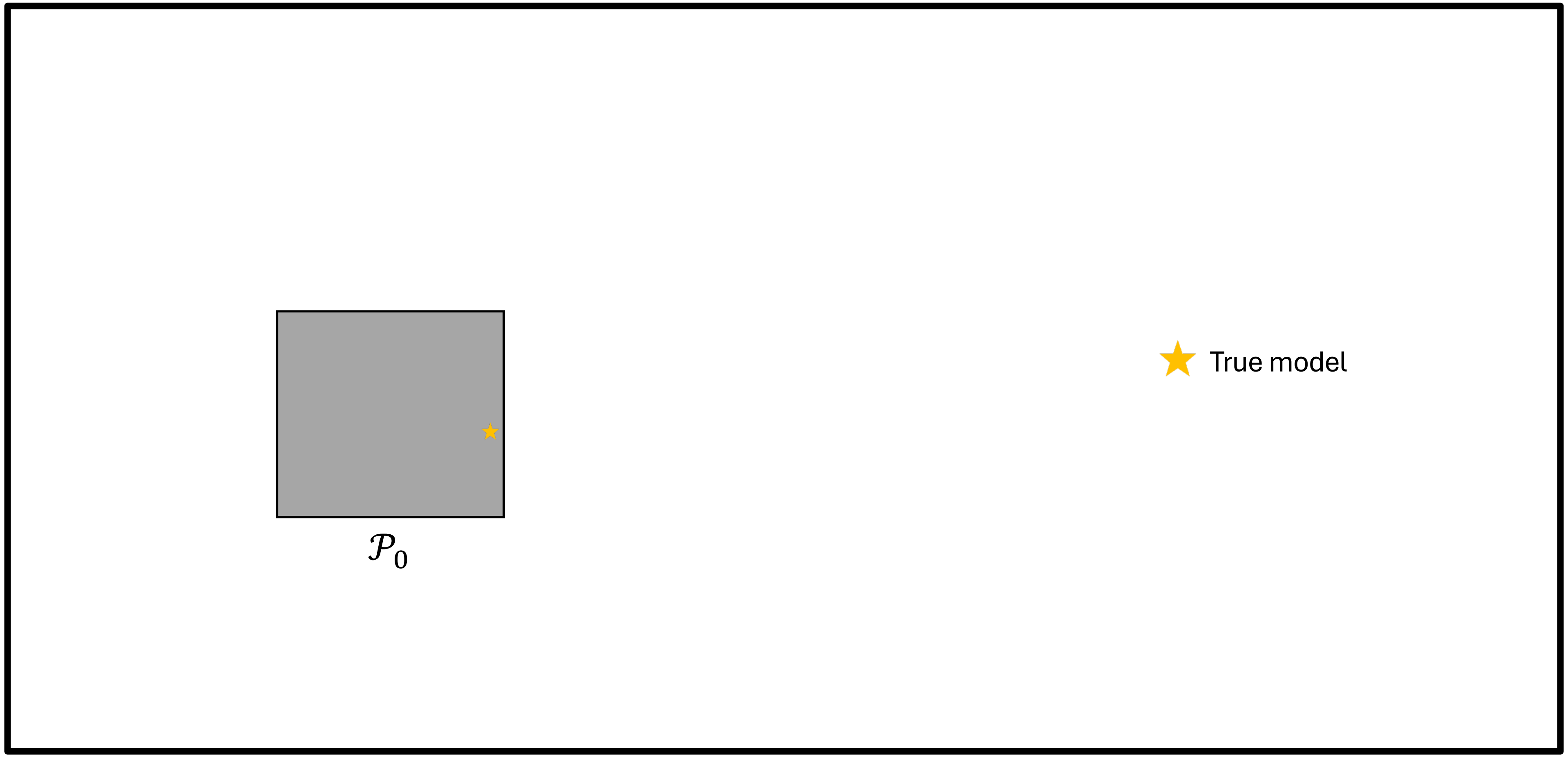
- Construct the **initial parameter uncertainty set** based on the available log data $\mathcal{D}$:

$$\mathcal{P}_0 = \{\theta \in \Theta \colon \|s_{k+1} - \theta^\top \Phi(s_k, z_k, a_k)\|_\infty \leq W, \forall (s_k, z_k, a_k) \in \mathcal{D}\}.$$

- **Log data**: 14 days of data generated under default controller in the target building
- **Interpretation**: $\mathcal{P}_0$ contains **all plausible models** that could have generated the log data given the disturbance bound and feature functions

OCFT Framework

Model adaptation



OCFT Framework

Model adaptation

1. Observe state transition to $s_{t+1} = (\theta^*)^\top \Phi(s_t, z_t, a_t) + d_t$

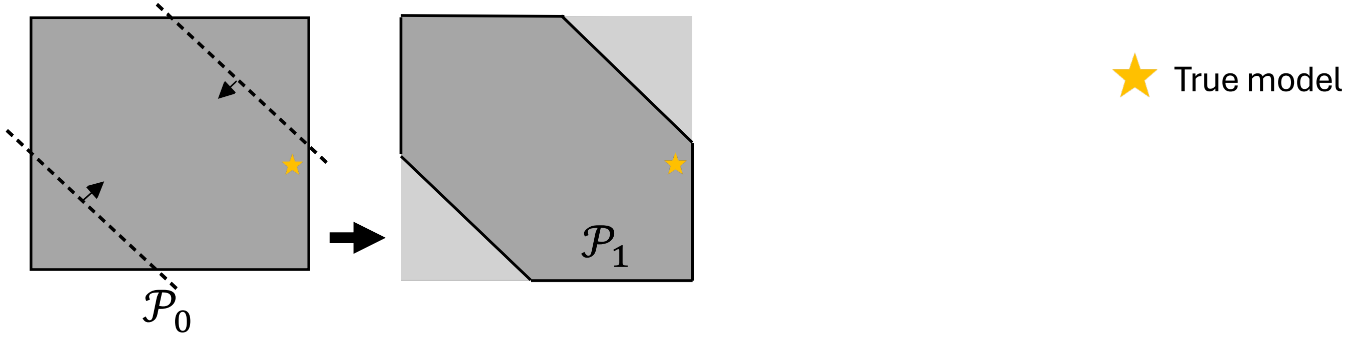
OCFT Framework

Model adaptation

1. Observe state transition to $s_{t+1} = (\theta^*)^\top \Phi(s_t, z_t, a_t) + d_t$

2. Update the **consistent model set with new data**:

$$\mathcal{P}_t = \{\theta \in \Theta\colon \|s_t - \theta^\top \Phi(s_{t-1}, z_{t-1}, a_{t-1})\|_\infty \le W\} \cap \mathcal{P}_{t-1}$$



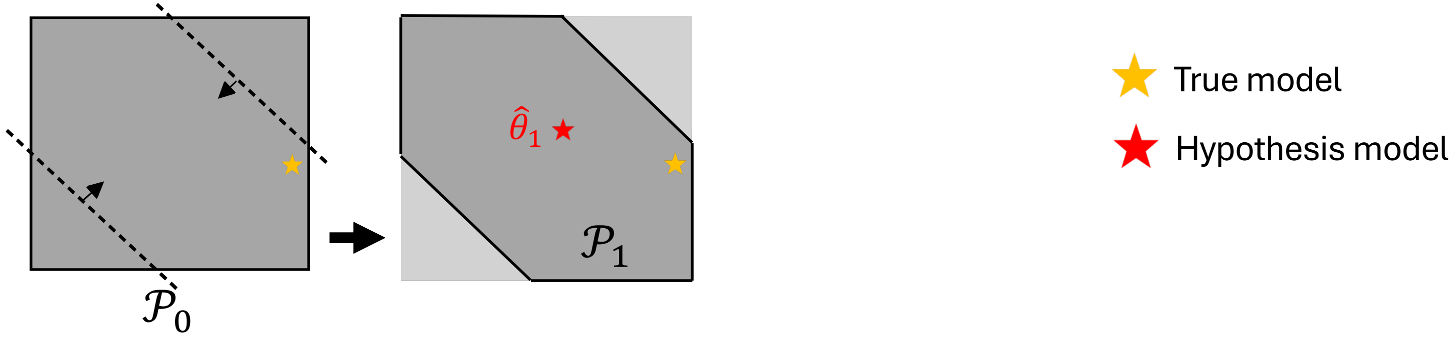
OCFT Framework

Model adaptation

1. Observe state transition to $s_{t+1} = (\theta^*)^\top \Phi(s_t, z_t, a_t) + d_t$

2. Update the **consistent model set with new data**:

$$\mathcal{P}_t = \{\theta \in \Theta\colon \|s_t - \theta^\top \Phi(s_{t-1}, z_{t-1}, a_{t-1})\|_\infty \le W\} \cap \mathcal{P}_{t-1}$$



3. Pick a ***hypothesis model*** (potentially wrong): $\hat{\theta}_t \in \mathcal{P}_t$ using **nested convex body chasing (NCBC)**

OCFT Framework

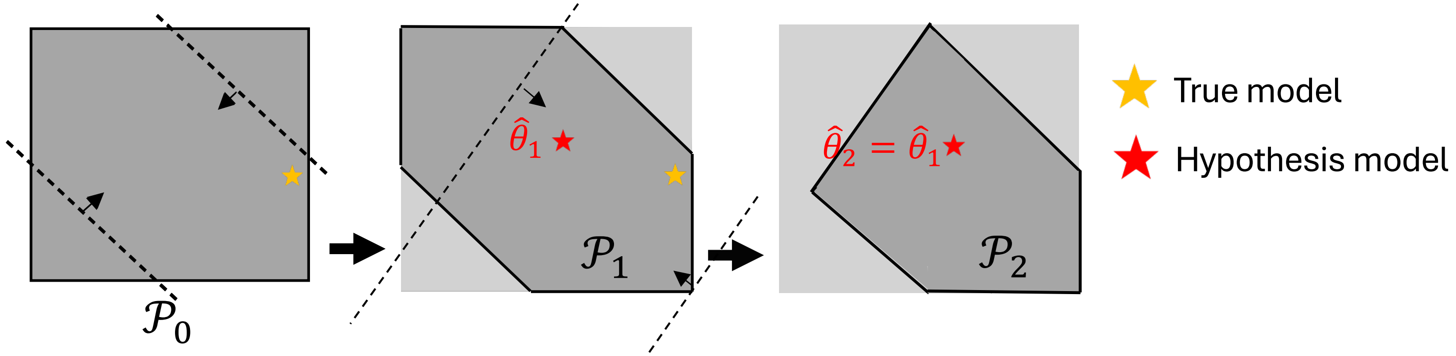
Model adaptation

$t+1$

1. Observe state transition to $s_{t+1} = (\theta^*)^\top \Phi(s_t, z_t, a_t) + d_t$

2. Update the **consistent model set with new data**:

$$\mathcal{P}_t = \{\theta \in \Theta\colon \|s_t - \theta^\top \Phi(s_{t-1}, z_{t-1}, a_{t-1})\|_\infty \leq W\} \cap \mathcal{P}_{t-1}$$



3. Pick a ***hypothesis model*** (potentially wrong): $\hat{\theta}_t \in \mathcal{P}_t$ using **nested convex body chasing (NCBC)**

OCFT Framework

Policy augmentation

Given black-box policy π_t, pass the suggested action through a filter based on $\hat{\theta}_t$:

trade-off parameter

adherence to black-box policy for energy use reduction

thermal constraint violation

$$\min_{a \in \mathcal{A}, \delta_1, \delta_2 \in \mathbb{R}} \lambda \| a - \pi_t(s_t) \|_2^2 + \eta_1 |\delta_1|^2 + \eta_2 |\delta_2|^2 ,$$

$$s.t. \quad \boxed{\begin{aligned} &\underline{T_t} + k - \delta_1 \leq \hat{\theta}_t^\top \Phi(s_t, z_t, a) \leq \overline{T_t} - k + \delta_2, \\ &k = W + \epsilon. \end{aligned}}$$

robust satisfaction of constraints against worst-case disturbances

slack variables

Theoretical insights

- **Adversarial disturbances** (worst-case model mismatch between our model and the true HVAC dynamics)

    Guarantee: OCFT makes finitely many thermal constraint violations

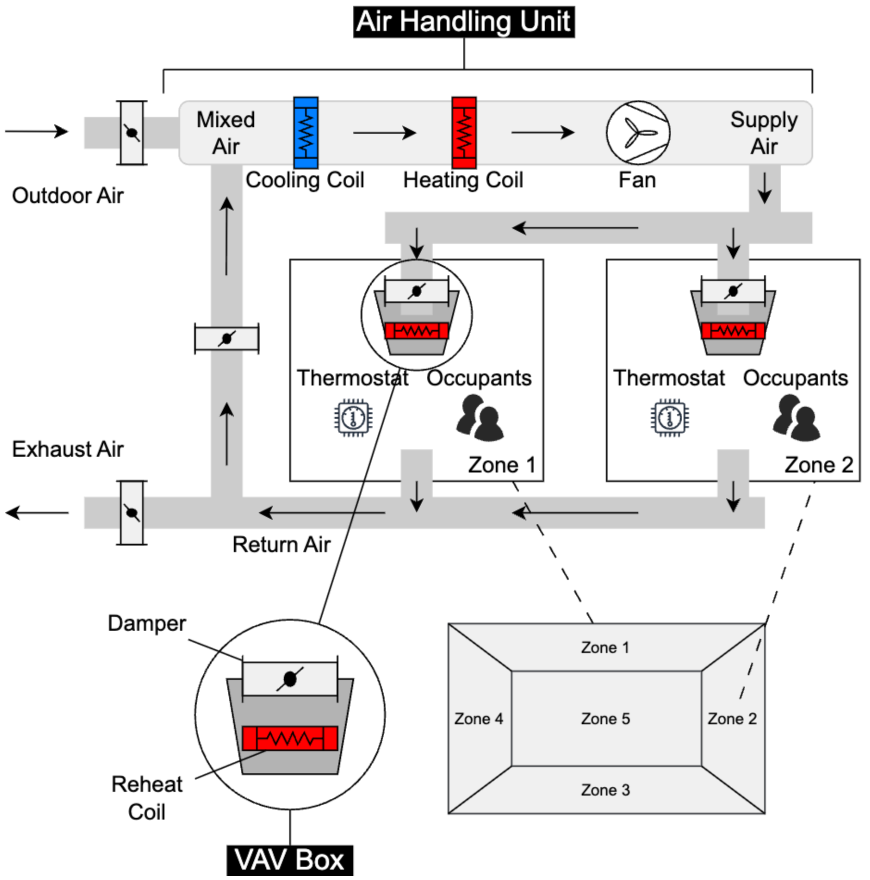
- **Stochastic disturbances** (when our model perfectly characterizes the true HVAC system dynamics)

    Guarantee: OCFT will learn the true system model with finite-time convergence guarantee

Experimental Results

Setup

Central AHU with zone-level VAV systems, each including
a damper and a reheat coil. We control only the damper position.



- **Target building:** 15-zone medium office building in 19 climates
- **Experiment period:** January and July
- **Warm-start data:** Two weeks of historical data from target building prior to each experiment period
- **Transfer policies:** 870 RL policies trained in a 5-zone building (Climate Zone 5B, January)
- **Feature functions:** ICNNs trained across various months in Climate Zone 5B

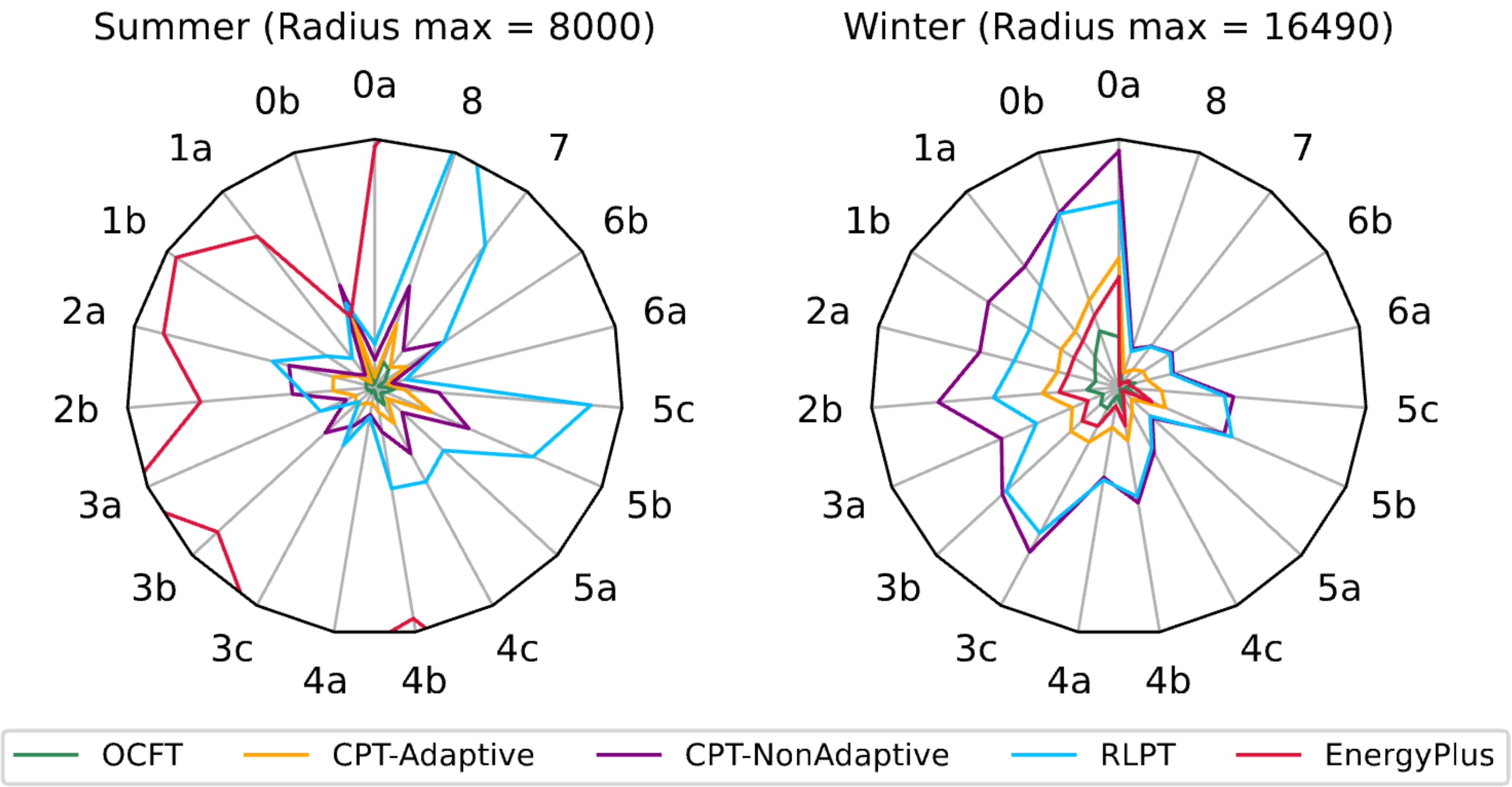
Experimental Results

Baseline algorithms

- Reinforcement Learning Policy Transfer (**RLPT**)
- Constrained Policy Transfer with Feature Function Adaptation (**CPT-Adaptive**)
  - adapted ICNN used in the projection layer for constraint satisfaction
- Constrained Policy Transfer without Feature Function Adaptation (**CPT-NonAdaptive**)
  - transferred ICNN used in the projection layer for constraint satisfaction
- Default Air System Control Strategy (**EnergyPlus**)

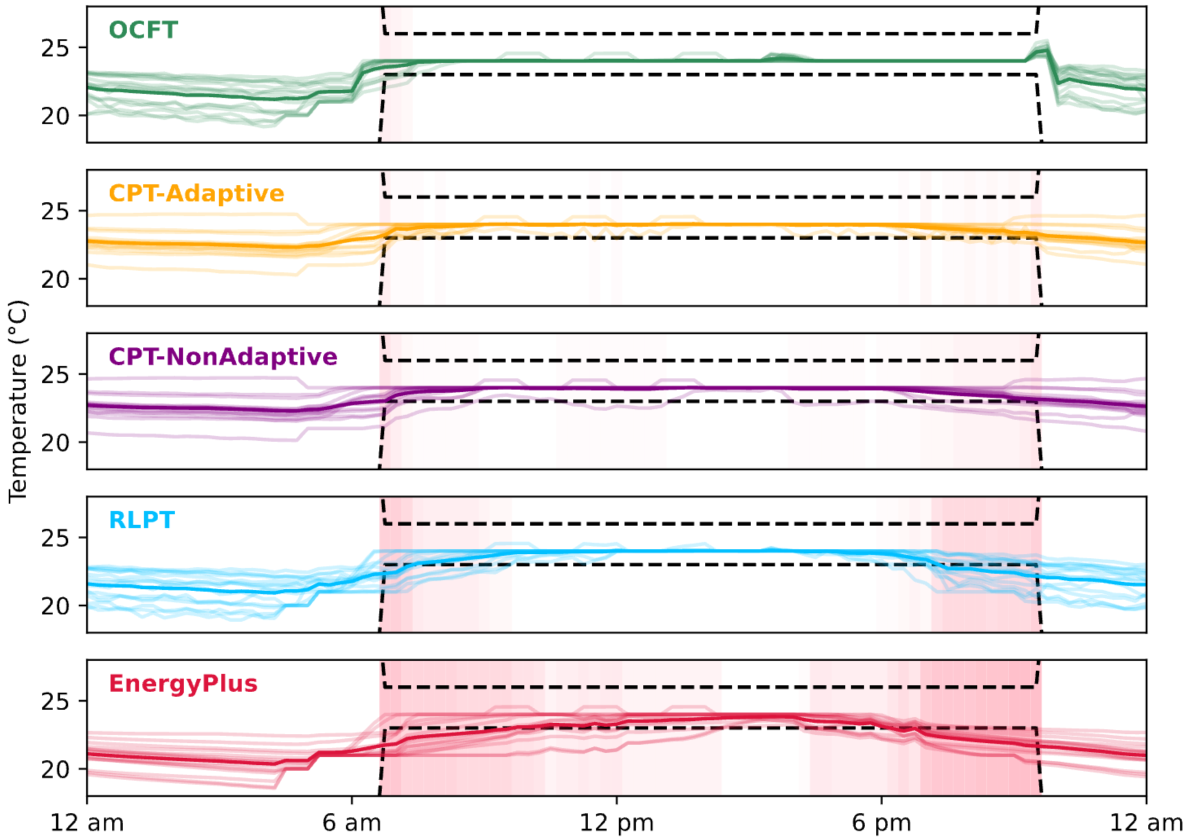
Experimental Results

Comfort constraint violations



Experimental Results

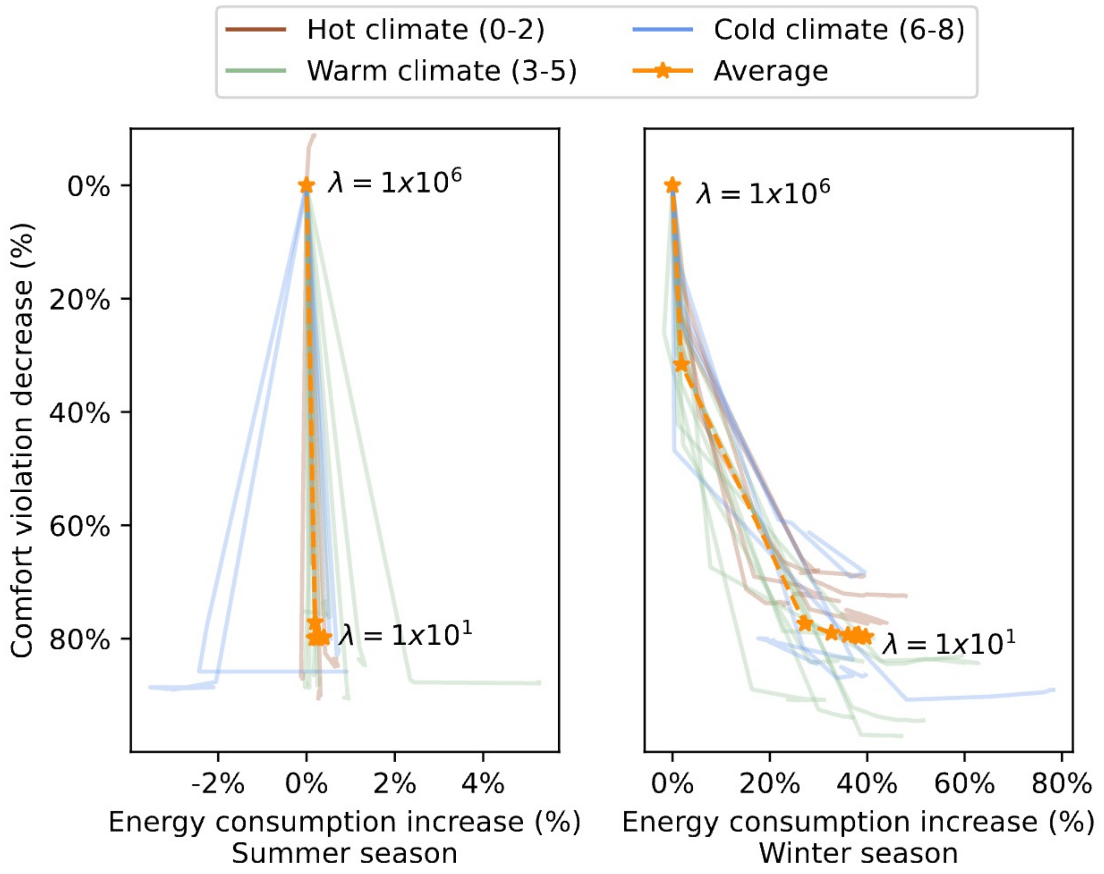
Temperature profile



Climate 5b in July

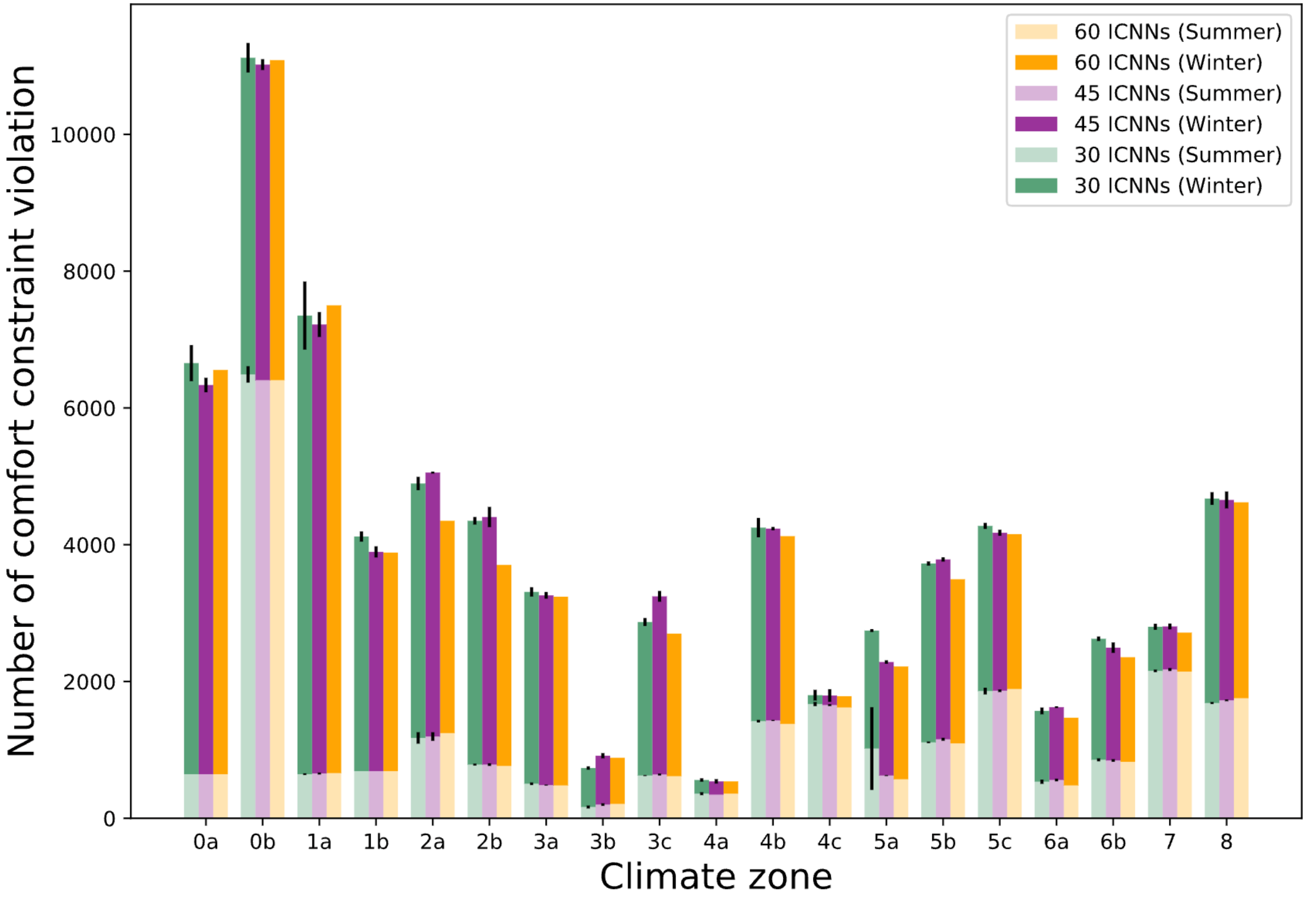
Experimental Results

Energy-comfort trade-off



Experimental Results

Robustness to feature function selection



$$s_{t+1} = (\theta^*)^\top \Phi(s_t, z_t, a_t) + d_t$$

ICNN : feature functions $\Phi(s_t, z_t, a_t)$ trained from the source building

Takeaways

- OCFT augments black-box transfer learning algorithms
  - Robust to small and simple feature functions
  - Requires little data to warm start and minimal tuning of parameters
  - Code: `github.com/sustainable-computing/ocft`
- OCFT trades off energy consumption & thermal comfort constraints
  - Reduces constraint violation by 81% while using 11% extra energy (vs. most energy-efficient baseline)
  - Reduces constraint violation by 59% while using 8% extra energy (vs. least constraint-violating baseline)