We develop methods for reducing emissions from buildings by <u>31%</u> with <u>zero</u> pre-training. Such approaches could scale to **every building in the world.**

Take a photo for the full paper, talk and code.

Low Emission Building Control with Zero Shot Reinforcement Learning

Scott Jeen {srj38@cam.ac.uk}, Alessandro Abate & Jonathan M. Cullen

1 Background

Motivation. Buildings account for 31% of global emissions. Past works have shown RL can reduce these by ~ 35%, but do so relying on pre-training in simulation.

Assumption. To scale to every building in the world, we must bypass expensive-to-obtain simulators and perform **zero-shot control**.

This Work. Can we find methods that can be deployed zero-shot, yet perform as well as SOTA pre-trained agents?

2 Problem Formulation

Standard MDP.

- state space $\mathcal{S} \in \mathbb{R}^{d_s}$
- action space $\mathcal{A} \in \mathbb{R}^{d_a}$
- reward function $\mathcal{R}(s_t, a_t)$
- transition dynamics $p(s_{s+1}|s_t, a_t)$

$\mathcal{R}(s_t, a_t)$ takes the form:

$$r[t] = \underbrace{r_E[t]}_{\text{emissions}} + \overbrace{r_T[t]}^{\text{temp.}} \quad \text{with,}$$

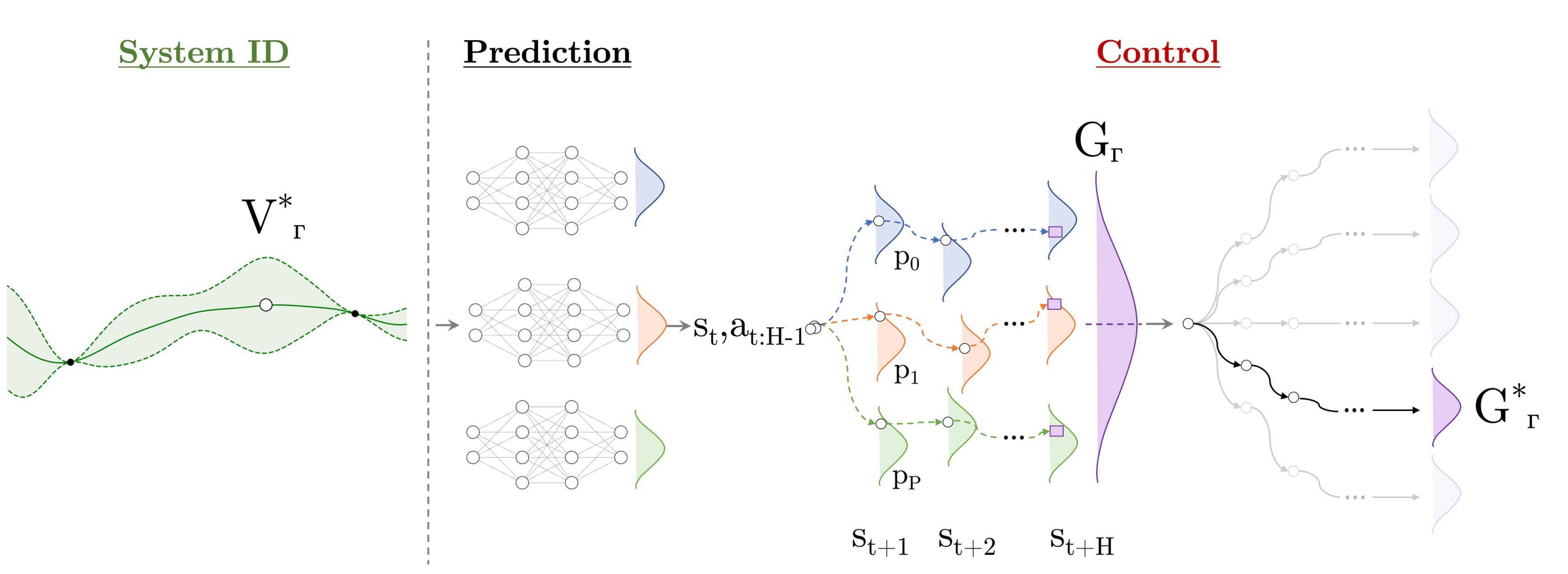
$$r_E[t] = -\phi(E[t]C[t]) \quad \text{and,}$$

$$r_T^i[t] = \begin{cases} 0: & T_{low} \leq T_{obs}^i \leq T_{high} \\ -\theta \min[(T_{low} - T_{obs}^i[t])^2,\ (T_{high} - T_{obs}^i)^2]: & \text{otherwise} \end{cases}$$

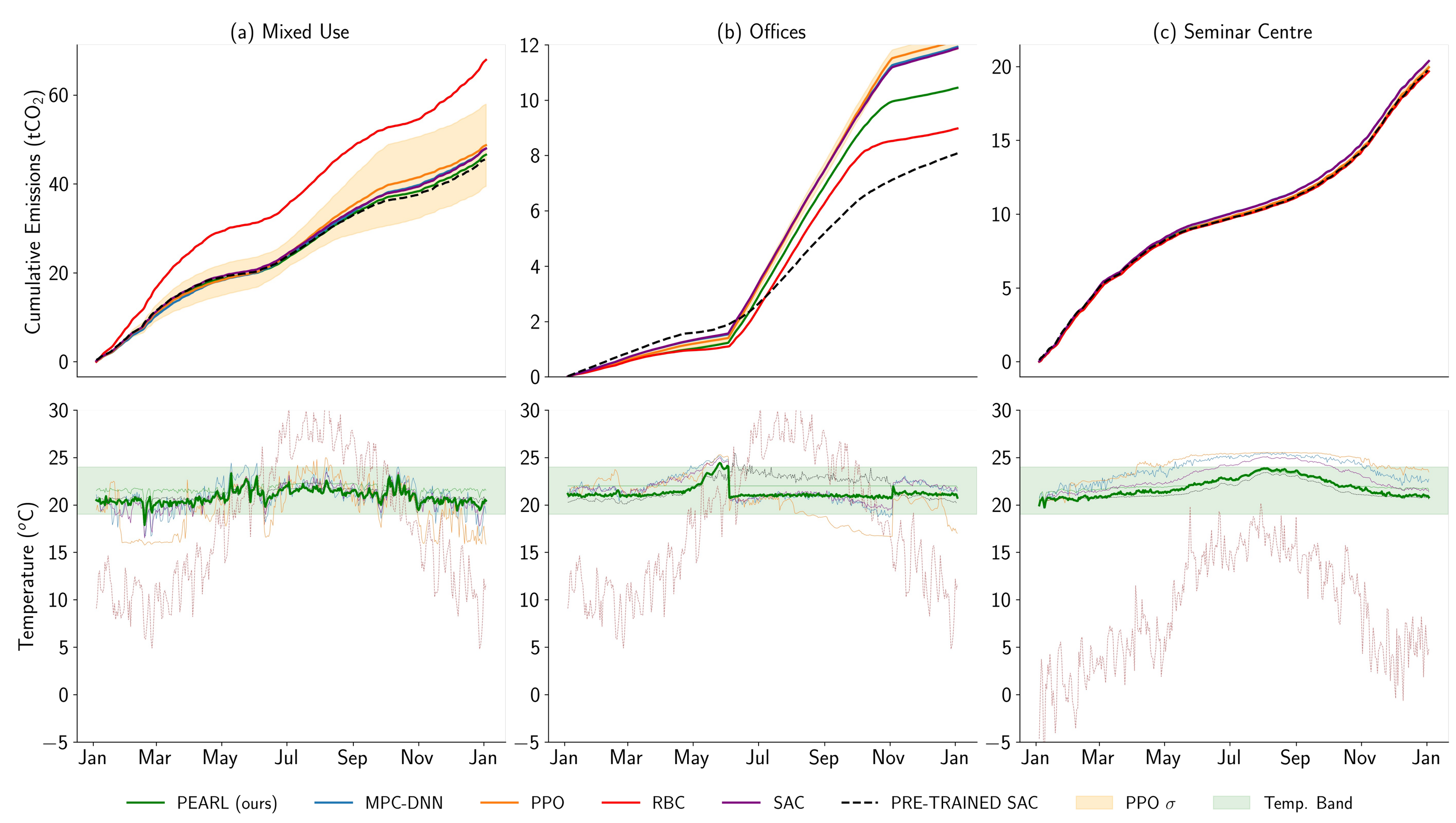
	Mixed-Use	Office	Seminar Centre
Location	Athens, Greece	Athens, Greece	Billund, Denmark
Floor Area (m^2)	566.38	643.78	1278.94
Action-dim	$\mathbb{R}_{12}$	$\mathbb{R}_{14}$	$\mathbb{R}_{18}$
State-dim	$\mathbb{R}_{37}$	$\mathbb{R}_{56}$	$\mathbb{R}_{59}$
Thermal Zones	13	25	27
Sampling Period	15 mins.	15 mins.	10 mins.
Equipment	Thermostats AHU	Thermostats	Thermostats

Task. Minimise emissions whilst ensuring thermal comfort for one year without prior knowledge or access to the simulator *a priori*.

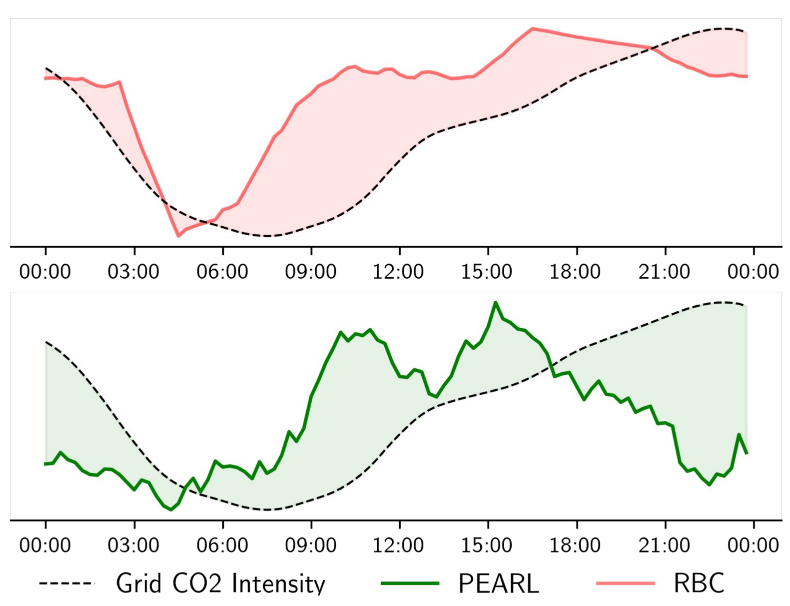
3 PEARL: Probabilistic Emission-Abating RL



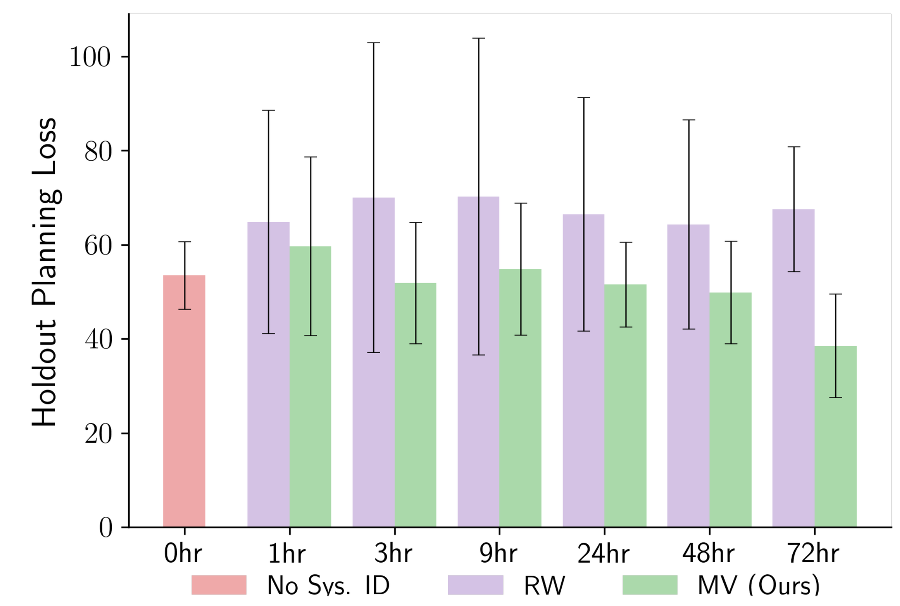
4 Results



Performance. Top: Cumulative emissions; Bottom: Mean daily building temp. PEARL outperforms all existing RL works, but is outperformed by the RBC in the *Office* environment.



Load Shifting. Power consumption w.r.t. grid carbon intensity on an exemplar day in the *Office* environment. We wish to maximise the shaded area to minimise emissions..



System ID. Planning MSE post-commissioning on a holdout set of 100 randomly sampled state-action trajectories, given varying system ID duration.