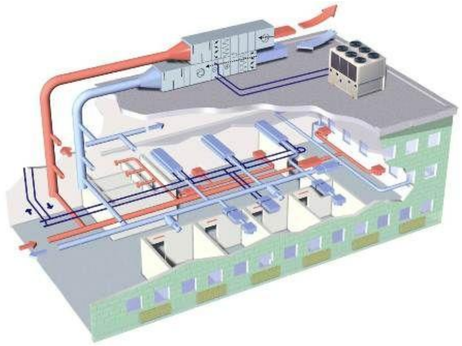
CLUE: Safe Model-Based RL HVAC ControL Using Epistemic Uncertainty Estimation

Zhiyu (Ryan) An, Xianzhong Ding, Arya Rathee, Wan Du

University of California, Merced

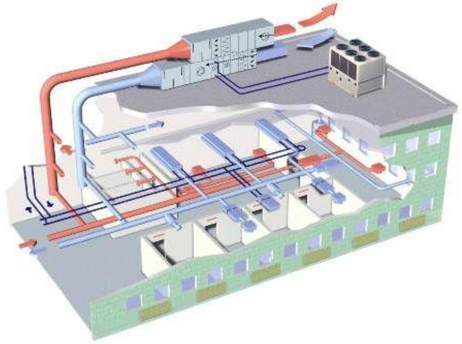
ACM BuildSys 2023, Istanbul, Turkey

HVAC control in smart buildings



Heating, Ventilation, and Air Conditioning (HVAC) system

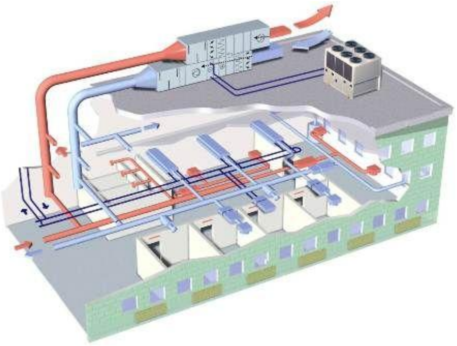
HVAC control in smart buildings



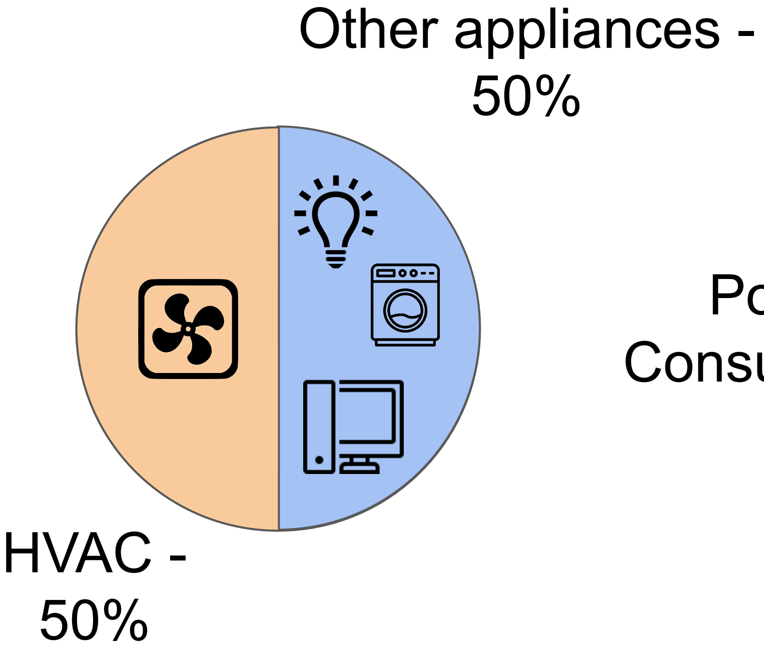
Heating, Ventilation, and Air Conditioning (HVAC) system

Thermal Comfort

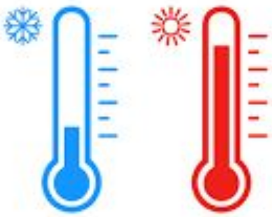
HVAC control in smart buildings



Heating, Ventilation, and Air Conditioning (HVAC) system

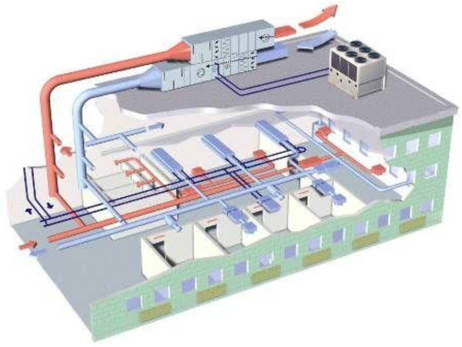


Power Consumption

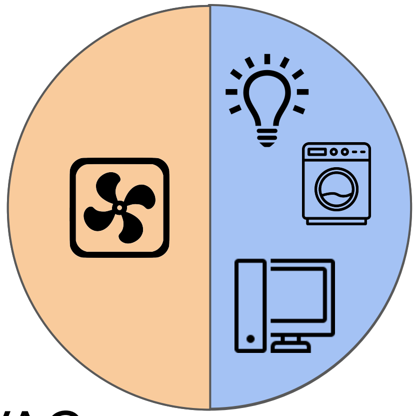


Thermal Comfort

HVAC control in smart buildings



Heating, Ventilation, and Air Conditioning (HVAC) system



Power Consumption

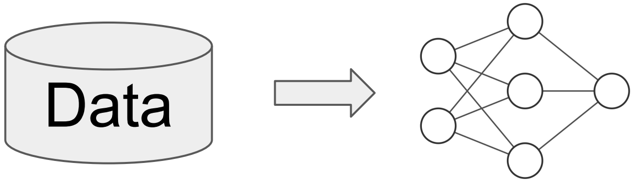
Dual objectives

Thermal Comfort

Current State-of-the-Art

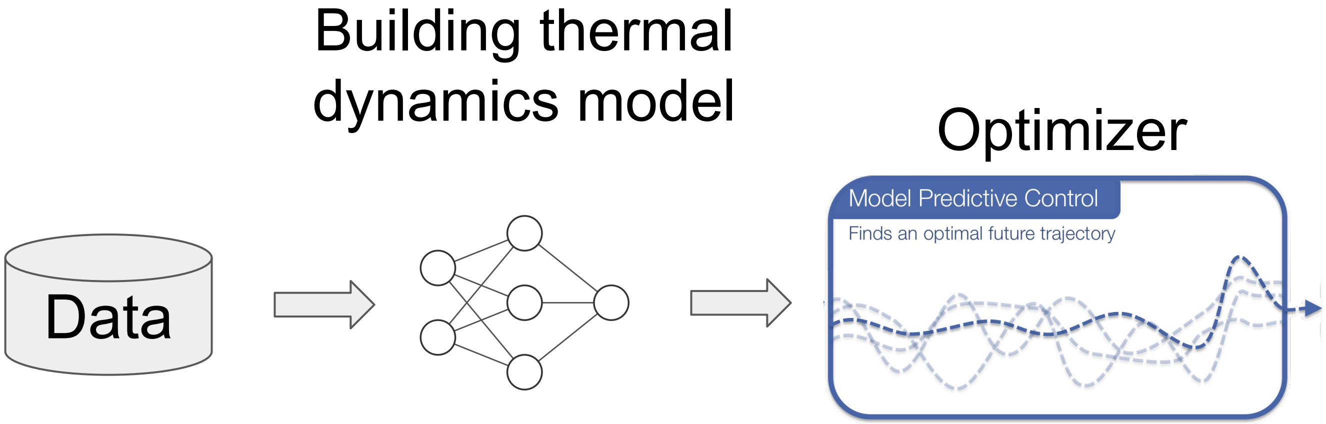
Data

Current State-of-the-Art



Building thermal dynamics model

Current State-of-the-Art



Building thermal dynamics model

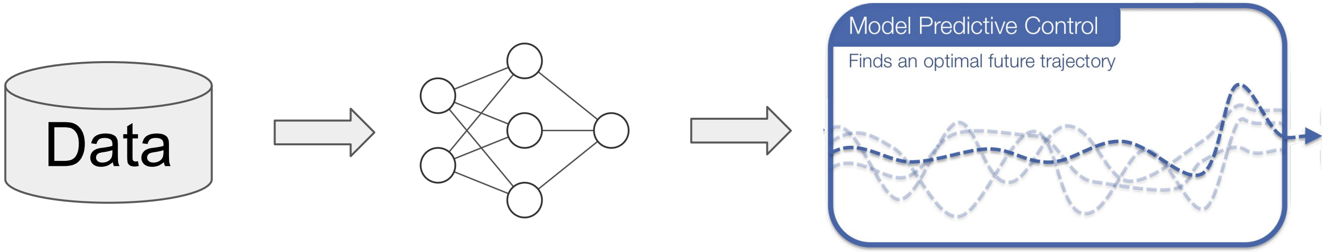
Optimizer

Data

Model Predictive Control

Finds an optimal future trajectory

Current State-of-the-Art



Building thermal dynamics model

Optimizer

Data

180 days of data collection[1]

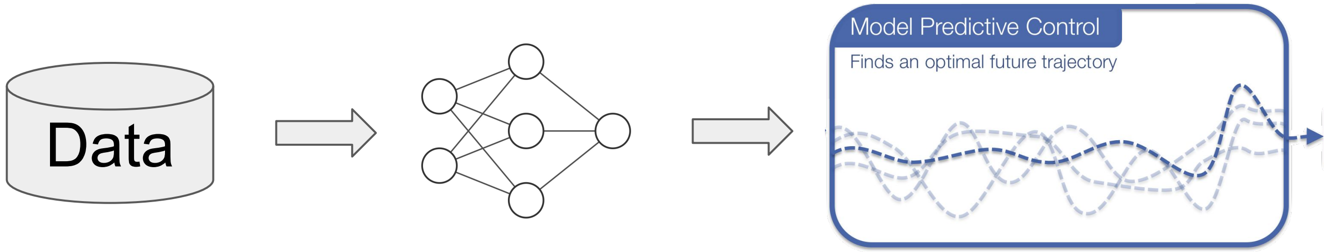
---

[1] Ding et al. MB2C: Model-based Deep Reinforcement Learning for Multi-zone Building Control. ACM BuildSys. 2020

Current State-of-the-Art

Building thermal dynamics model

Optimizer



180 days of data collection[1]

---

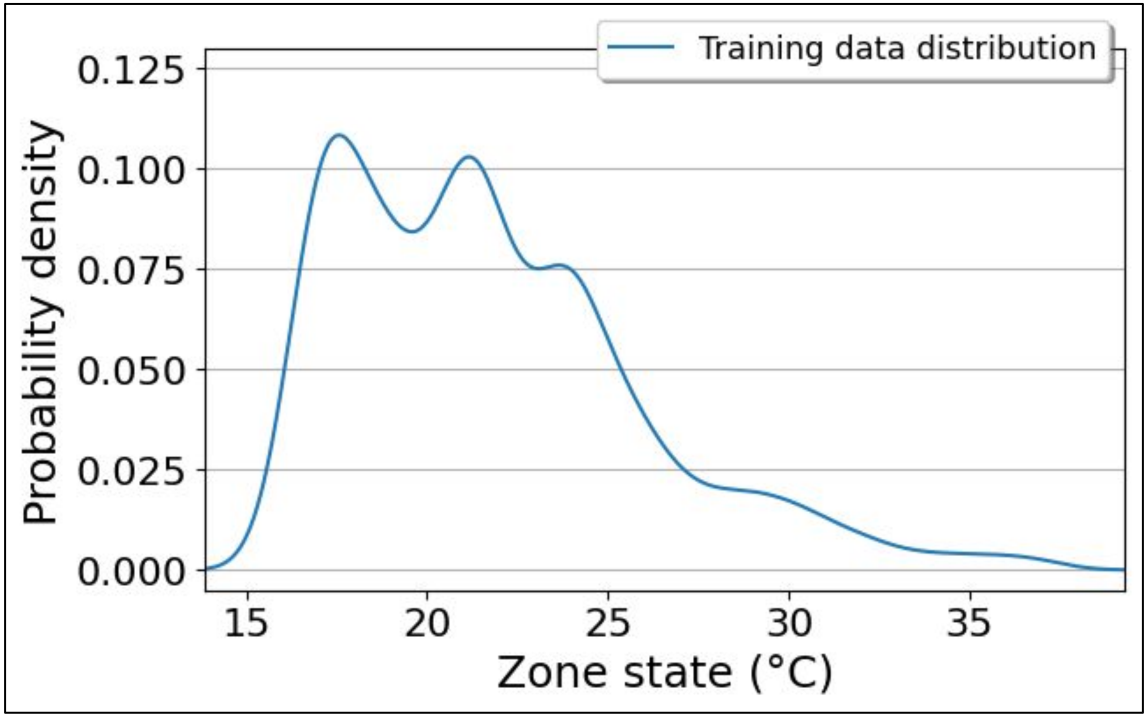
[1] Ding et al. MB2C: Model-based Deep Reinforcement Learning for Multi-zone Building Control. ACM BuildSys. 2020

Preliminary Experiment Setting

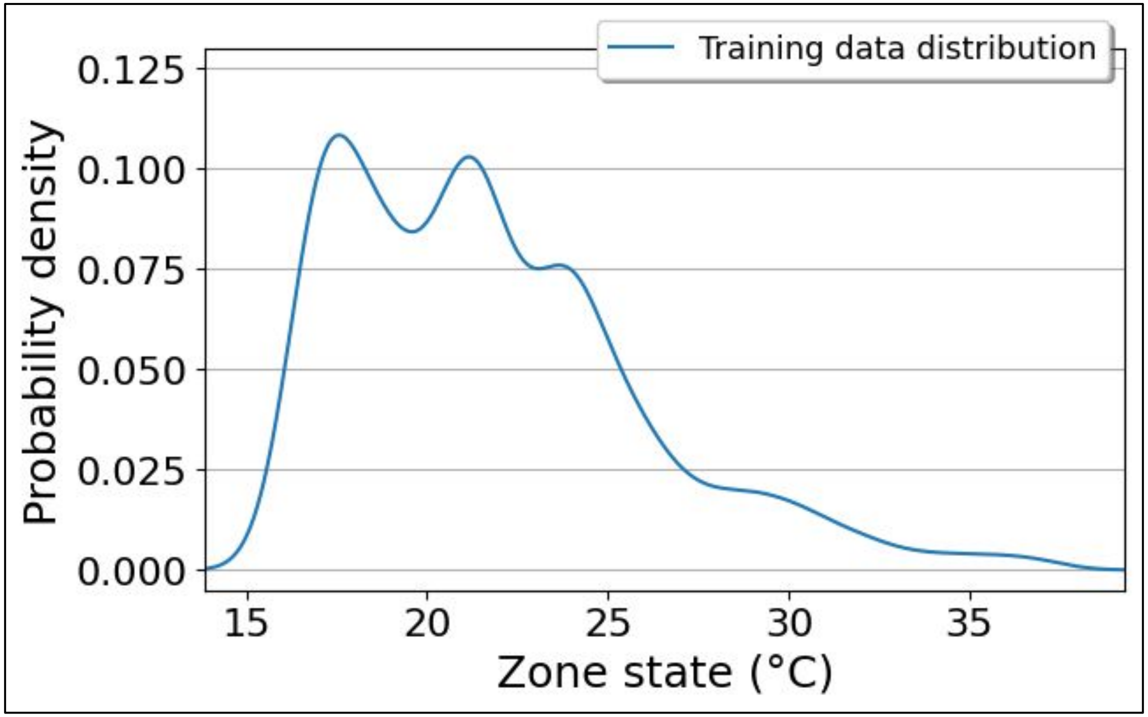
- Single thermal zone
- Five-layer Neural Network [1]

[1] Ding et al. MB2C: Model-based Deep Reinforcement Learning for Multi-zone Building Control. ACM BuildSys. 2020

Preliminary Experiment Result

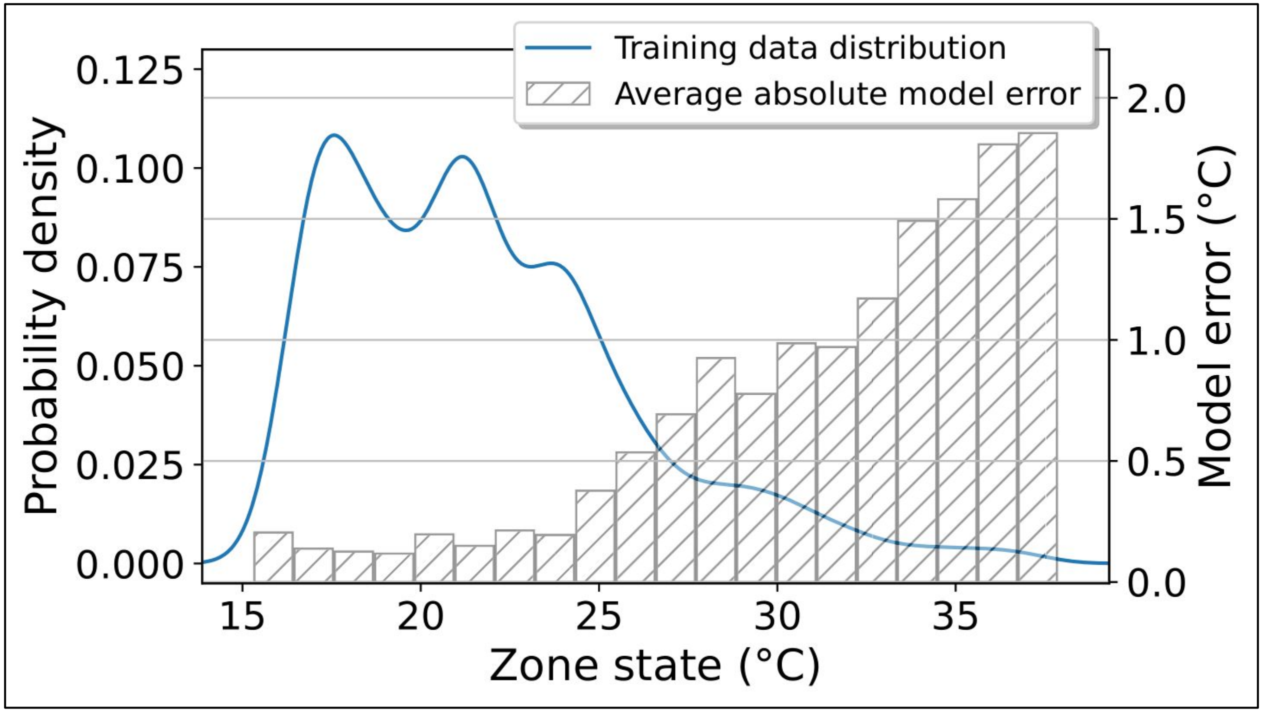


Preliminary Experiment Result



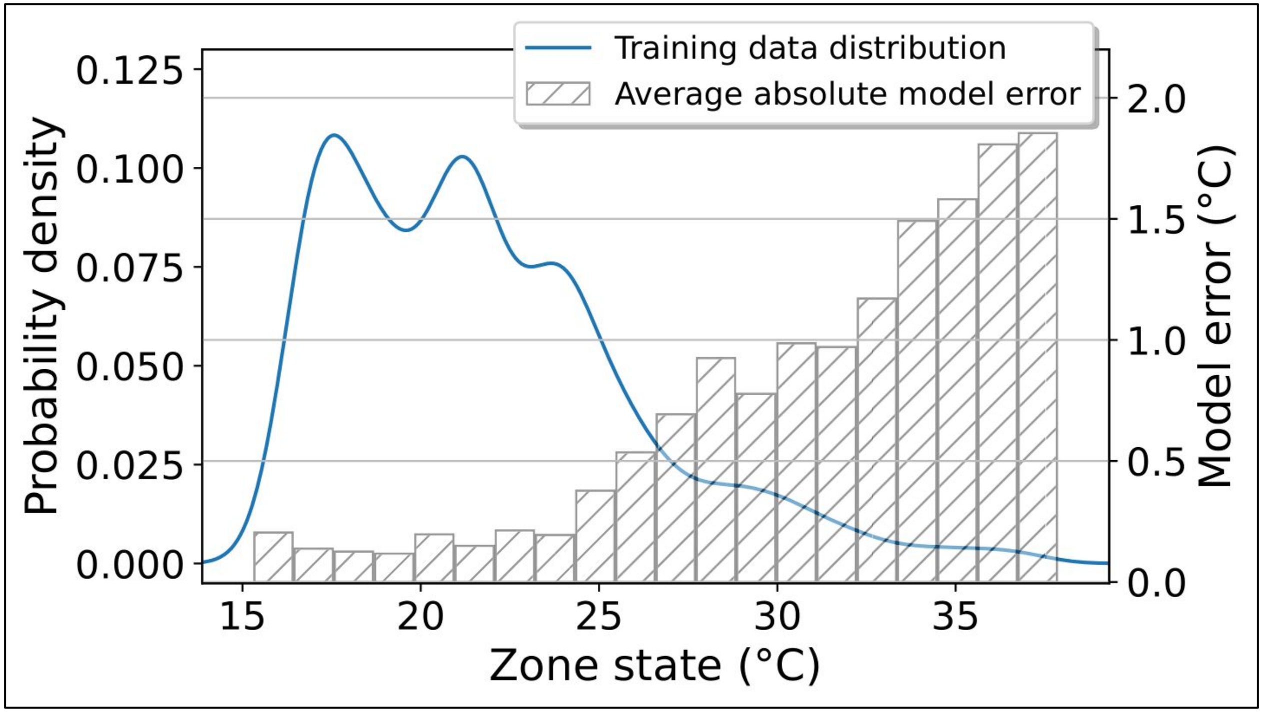
1. Building data is intrinsically biased

Preliminary Experiment Result



1. Building data is intrinsically biased

Preliminary Experiment Result

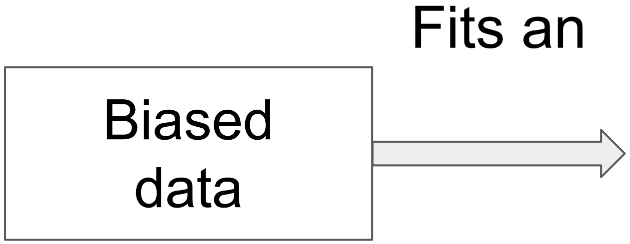


1. Building data is intrinsically biased
2. Biased data result in inaccurate model

Our Insight

Biased data

Our Insight



Fits an

Biased data

Our Insight

Biased data

Fits an

Inaccurate model

Our Insight

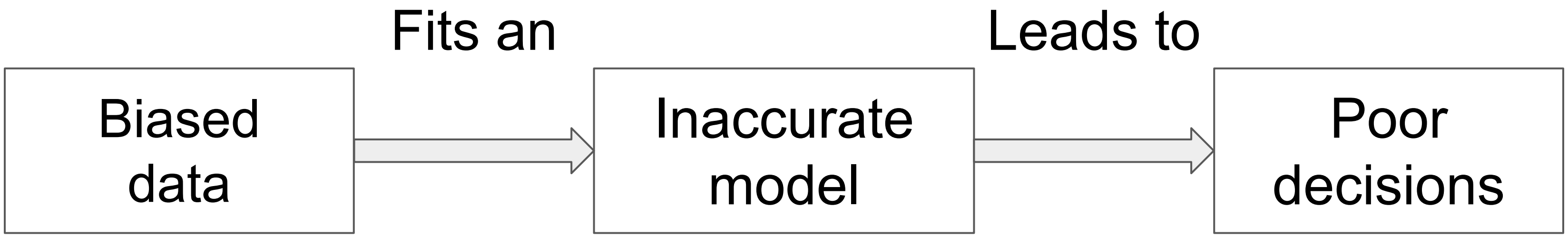
Biased data

Fits an

Inaccurate model

Leads to

Our Insight



Biased data

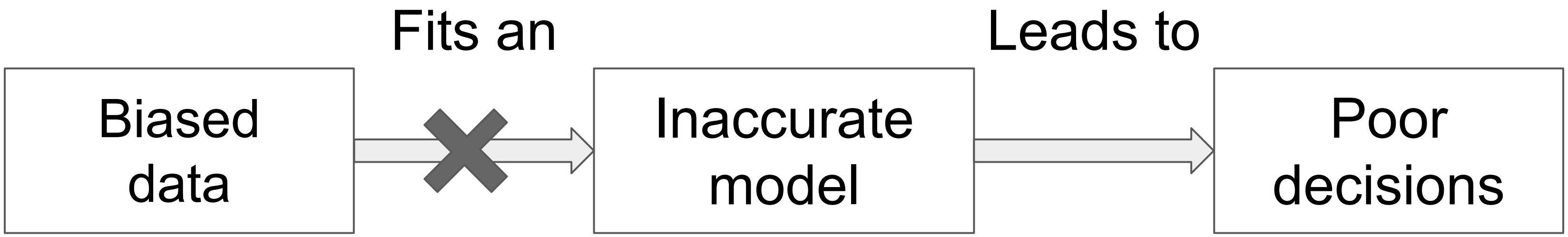
Fits an

Inaccurate model

Leads to

Poor decisions

Our Insight



Fits an

Leads to

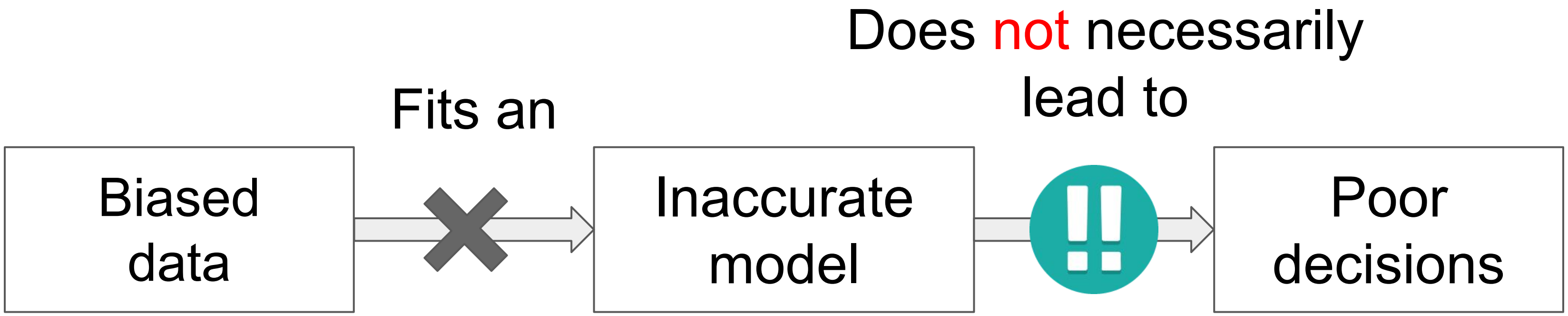
Biased data

Inaccurate model

Poor decisions

Not enough information!

Our Insight



Biased data

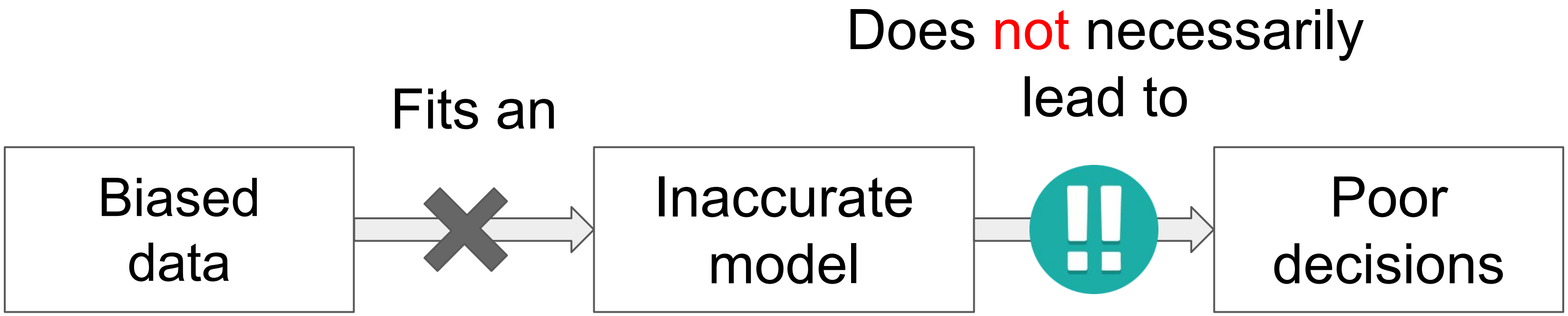
Fits an

Inaccurate model

Does not necessarily lead to

Poor decisions

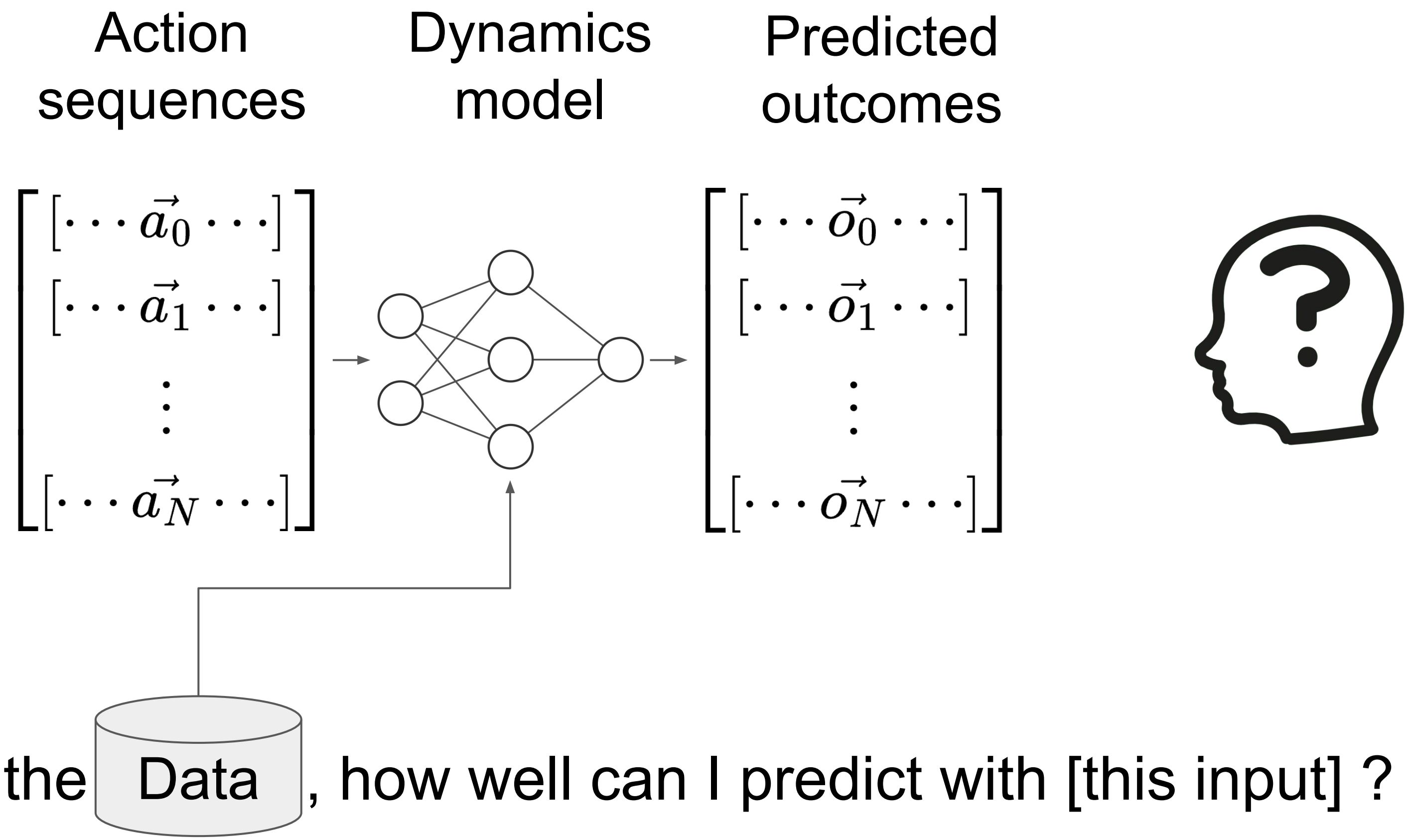
Our Insight



Uncertainty-Aware Model

Make conservative decision
when model is uncertain

Our Method



Action sequences

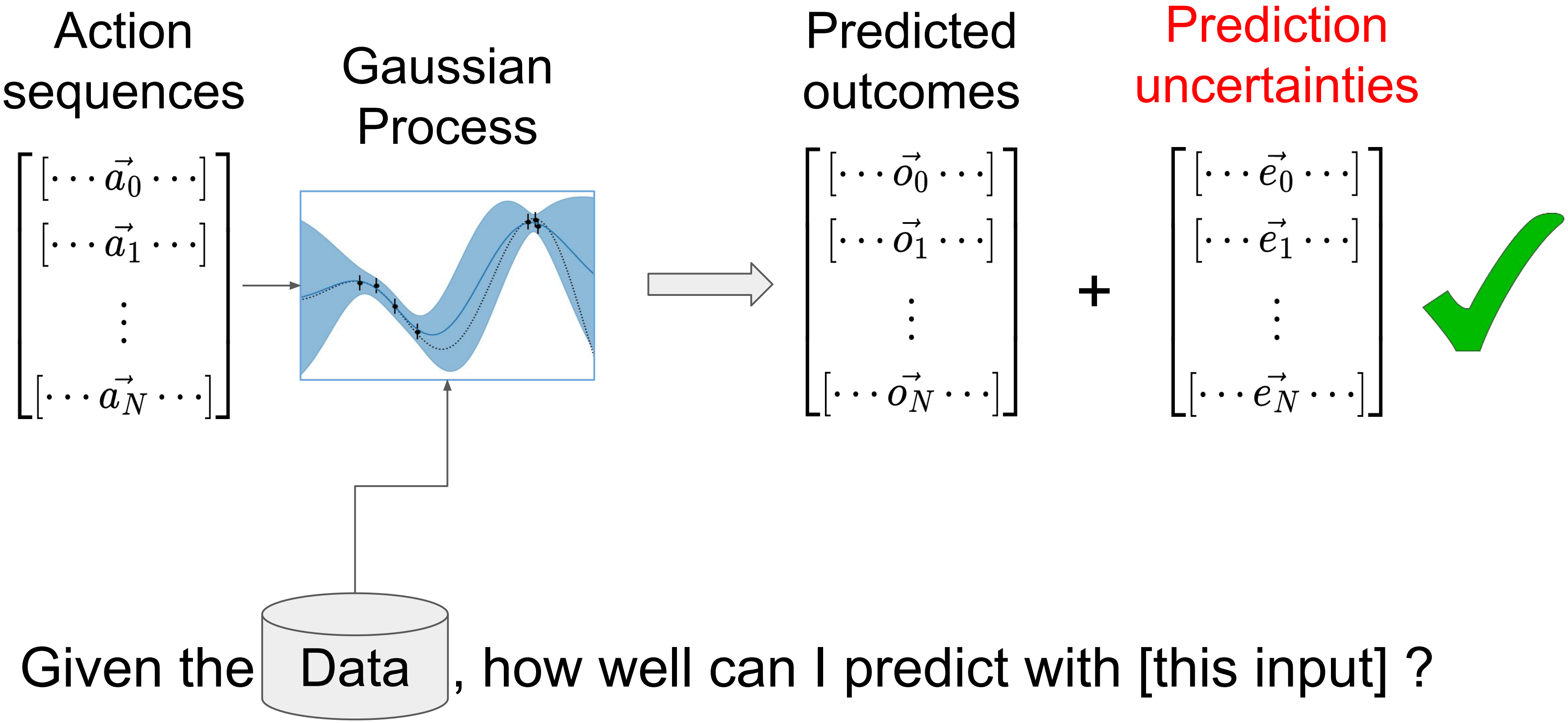
Dynamics model

Predicted outcomes

$$\begin{bmatrix} [\cdots \vec{a}_0 \cdots] \\ [\cdots \vec{a}_1 \cdots] \\ \vdots \\ [\cdots \vec{a}_N \cdots] \end{bmatrix} \rightarrow \quad \rightarrow \begin{bmatrix} [\cdots \vec{o}_0 \cdots] \\ [\cdots \vec{o}_1 \cdots] \\ \vdots \\ [\cdots \vec{o}_N \cdots] \end{bmatrix}$$

Given the Data, how well can I predict with [this input] ?

Our Method



Action sequences

$$\begin{bmatrix} [\cdots \vec{a}_0 \cdots] \\ [\cdots \vec{a}_1 \cdots] \\ \vdots \\ [\cdots \vec{a}_N \cdots] \end{bmatrix}$$

Gaussian Process

Predicted outcomes

$$\begin{bmatrix} [\cdots \vec{o}_0 \cdots] \\ [\cdots \vec{o}_1 \cdots] \\ \vdots \\ [\cdots \vec{o}_N \cdots] \end{bmatrix}$$

+

Prediction uncertainties

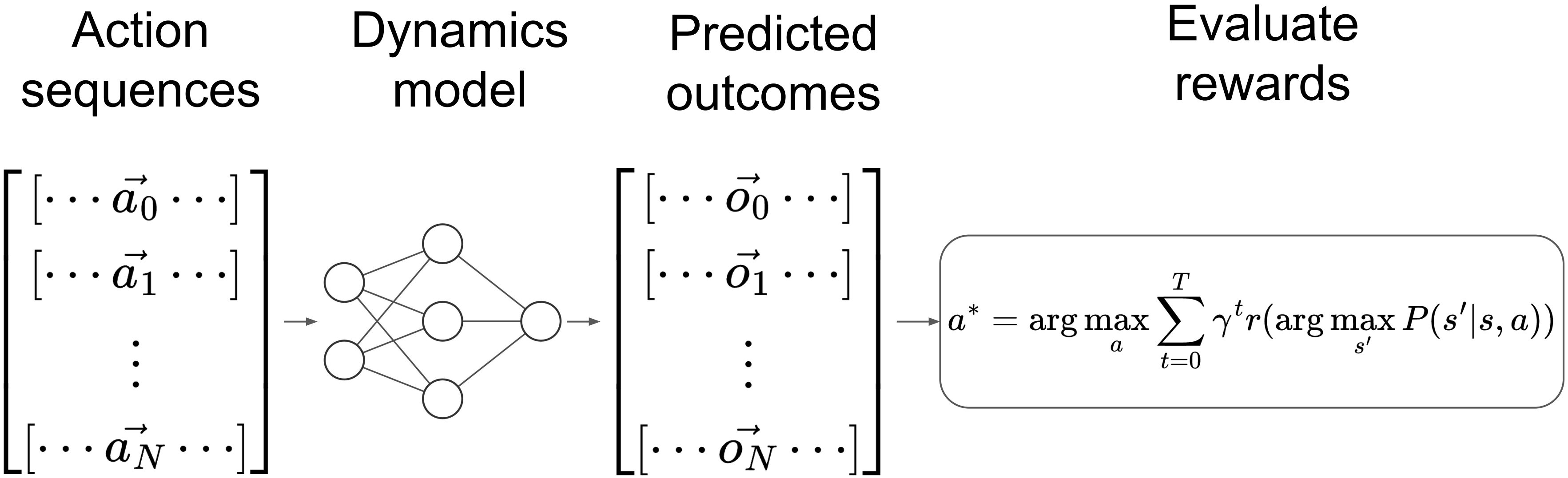
$$\begin{bmatrix} [\cdots \vec{e}_0 \cdots] \\ [\cdots \vec{e}_1 \cdots] \\ \vdots \\ [\cdots \vec{e}_N \cdots] \end{bmatrix}$$

✓

Given the Data, how well can I predict with [this input] ?

Our Method: Uncertainty-Aware Controller

Instead of using a traditional controller like this…..



Action sequences

Dynamics model

Predicted outcomes

Evaluate rewards

$$\begin{bmatrix}[\cdots \vec{a}_0 \cdots] \\ [\cdots \vec{a}_1 \cdots] \\ \vdots \\ [\cdots \vec{a}_N \cdots]\end{bmatrix} \rightarrow \begin{bmatrix}[\cdots \vec{o}_0 \cdots] \\ [\cdots \vec{o}_1 \cdots] \\ \vdots \\ [\cdots \vec{o}_N \cdots]\end{bmatrix}$$

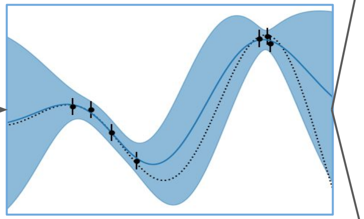
$$a^* = \arg\max_a \sum_{t=0}^{T} \gamma^t r(\arg\max_{s'} P(s'|s,a))$$

Our Method: Uncertainty-Aware Controller

Action sequences

$$\begin{bmatrix}[\cdots \vec{a_0} \cdots] \\ [\cdots \vec{a_1} \cdots] \\ \vdots \\ [\cdots \vec{a_N} \cdots]\end{bmatrix}$$

GP model



Predicted outcomes

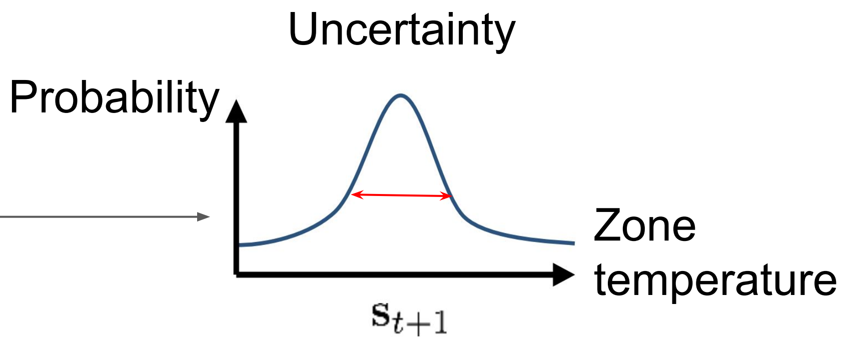
$$\begin{bmatrix}[\cdots \vec{o_0} \cdots] \\ [\cdots \vec{o_1} \cdots] \\ \vdots \\ [\cdots \vec{o_N} \cdots]\end{bmatrix}$$

Prediction uncertainties

$$\begin{bmatrix}[\cdots \vec{e_0} \cdots] \\ [\cdots \vec{e_1} \cdots] \\ \vdots \\ [\cdots \vec{e_N} \cdots]\end{bmatrix}$$

considers uncertainty in the objective function

$$a^* = \arg\max_a \sum_{t=0}^{T} \gamma^t \Big(r\big(\arg\max_{s'} P(s'|s,a)\big) + \lambda \mathbb{E}\big[|P(s'|s,a) - P\widehat{(s'|s,a)}|\big]\Big)$$

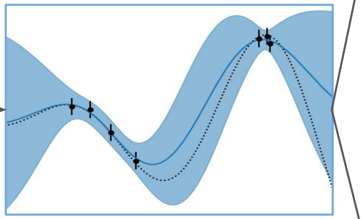


Our Method: Uncertainty-Aware Controller

Action sequences

$$\begin{bmatrix}[\cdots \vec{a}_0 \cdots]\\ [\cdots \vec{a}_1 \cdots]\\ \vdots \\ [\cdots \vec{a}_N \cdots]\end{bmatrix}$$

GP model



GP is not data efficient!

Predicted outcomes

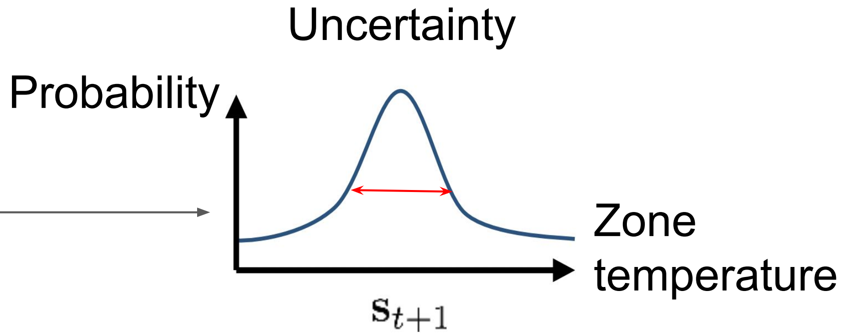
$$\begin{bmatrix}[\cdots \vec{o}_0 \cdots]\\ [\cdots \vec{o}_1 \cdots]\\ \vdots \\ [\cdots \vec{o}_N \cdots]\end{bmatrix}$$

Prediction uncertainties

$$\begin{bmatrix}[\cdots \vec{e}_0 \cdots]\\ [\cdots \vec{e}_1 \cdots]\\ \vdots \\ [\cdots \vec{e}_N \cdots]\end{bmatrix}$$

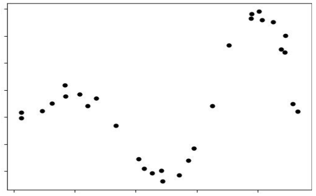
considers uncertainty in the objective function

$$a^* = \arg\max_a \sum_{t=0}^{T} \gamma^t \big(r(\arg\max_{s'} P(s'|s,a)) + \lambda \mathbb{E}[|P(s'|s,a) - P(\widehat{s'|s,a})|]\big)$$



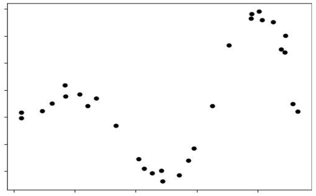
Gaussian Process Data Efficiency

Data



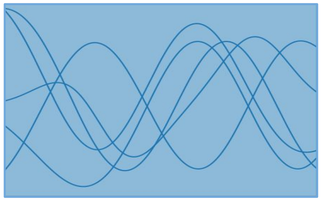
Gaussian Process Data Efficiency

Data

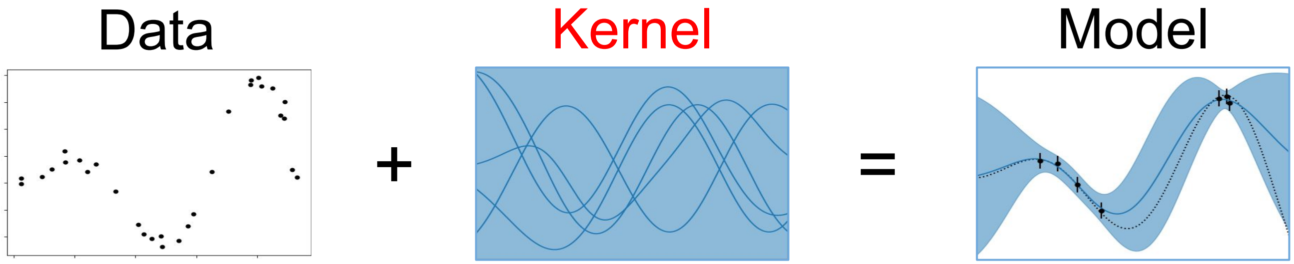


\+

Kernel



Gaussian Process Data Efficiency



RBF kernel:

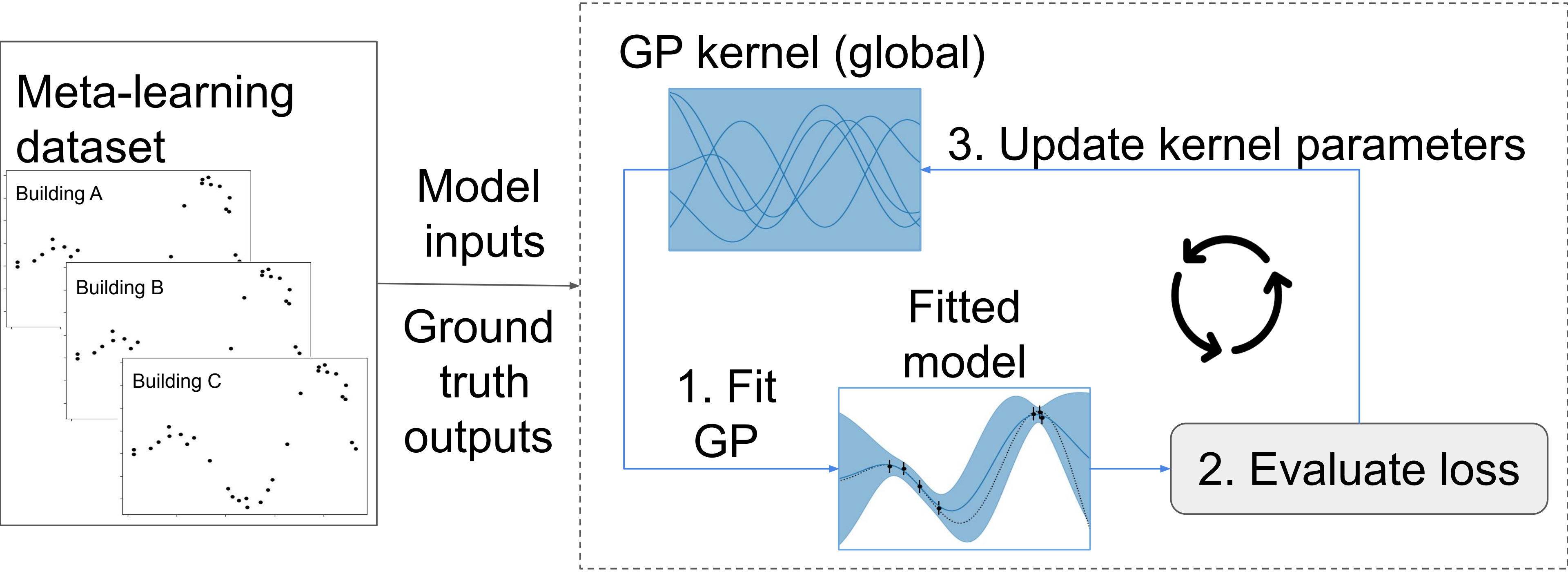
$$k(x, x') = \underset{\uparrow}{\theta_{\text{scale}}} \exp\left(-\frac{1}{2}(x - x')^\top \underset{\uparrow}{\Theta}^{-2}(x - x')\right)$$

$$\theta_{\mathcal{GP}} : \{\theta_{\text{scale}}, \Theta \in \mathbb{R}^{|\mathcal{X}|\times|\mathcal{X}|}\}$$

Challenge: how to efficiently tune the GP kernel hyperparameters?

Our Solution: Meta Kernel Learning

Use data from similar buildings to tune the hyperparameters!



Meta-learning dataset

Building A

Building B

Building C

Model inputs

Ground truth outputs

GP kernel (global)

3. Update kernel parameters

Fitted model

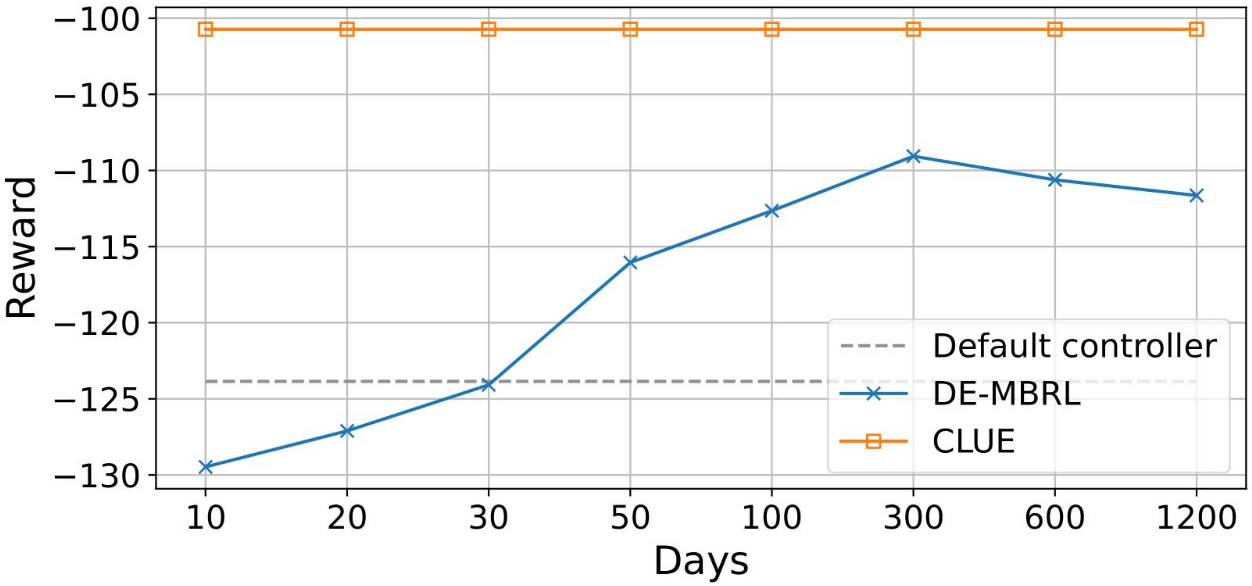
1. Fit GP

2. Evaluate loss

Experiment

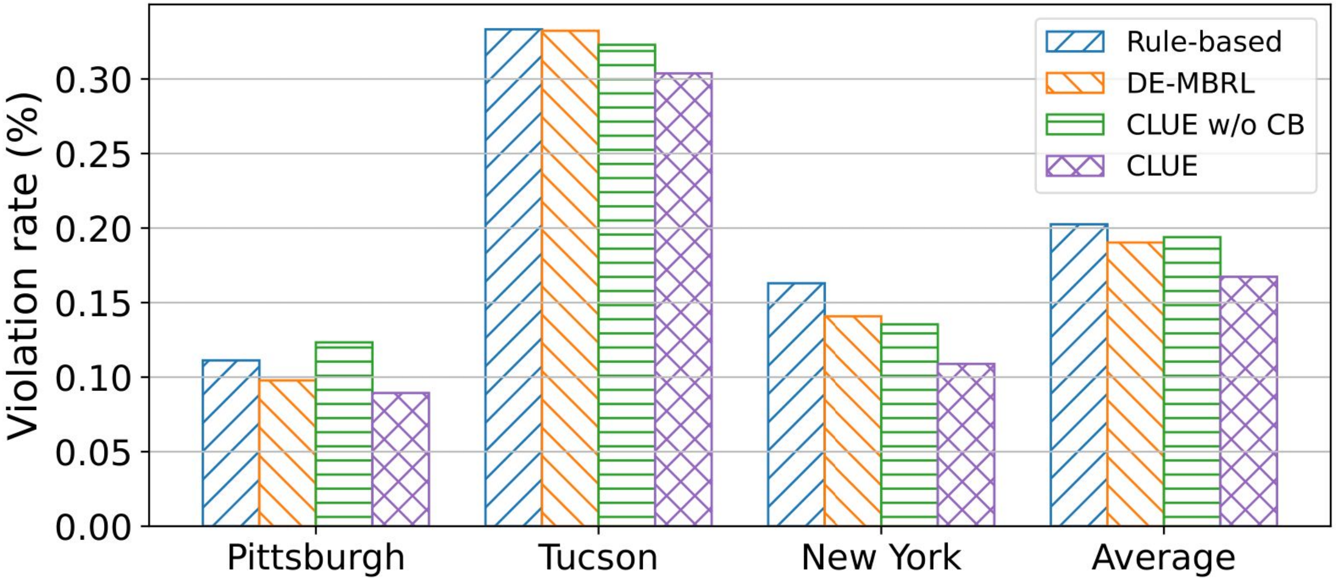
- Platform: EnergyPlus
- Three locations: Pittsburgh, Tucson, New York
- Data efficiency
- Building control performance
  - Energy usage
  - Comfort violation rate

Experiment - Data Efficiency



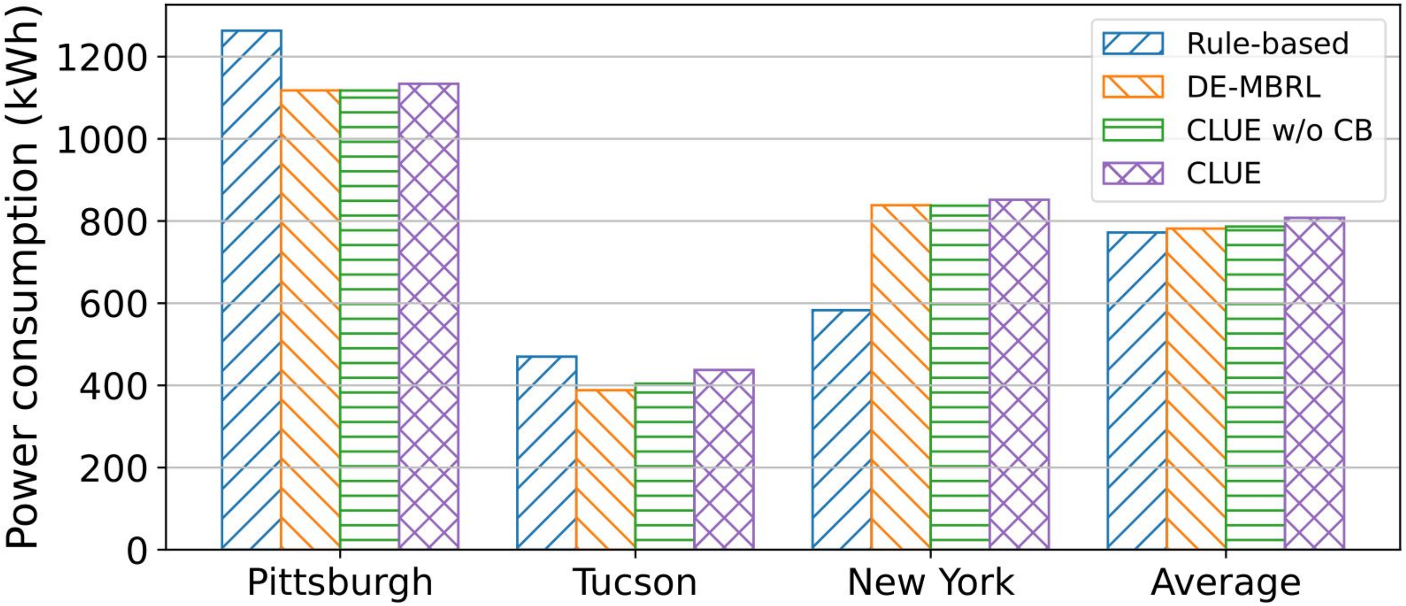
CLUE converges >30x faster than previous SOTA

Experiment - Building Control #1



CLUE produces 12.07% lower violation rates compared w/ previous SOTA

Experiment - Building Control #2



Similar energy saving w/ previous SOTA

Conclusion

- We are the first to include epistemic uncertainty estimation in MBRL for HVAC.
- We proposed *CLUE*, a data-efficient and safe MBRL control method for HVAC, consists of meta kernel learning and uncertainty-aware control.
- We evaluated *CLUE* with extensive simulation experiments in three different locations.

Code+data available at https://github.com/ryeii/CLUE/

Thank you

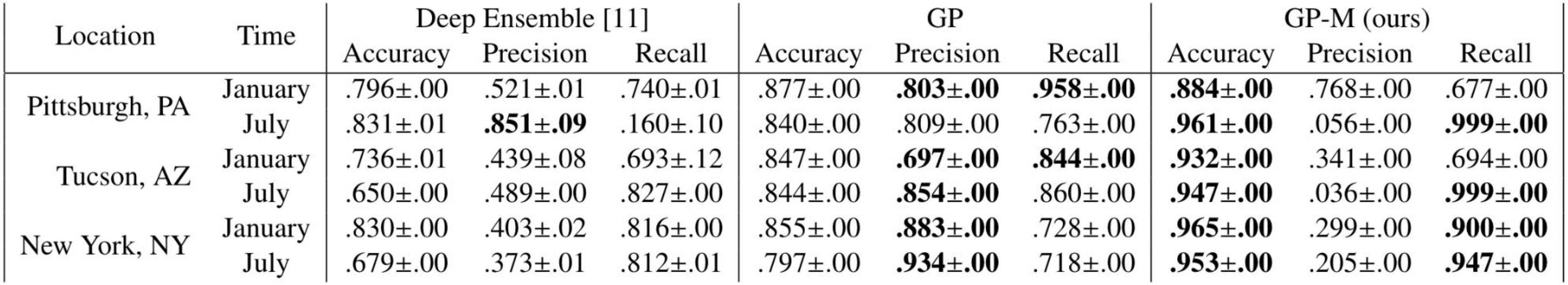
Q&A

Zhiyu (Ryan) An

zan7@ucmerced.edu

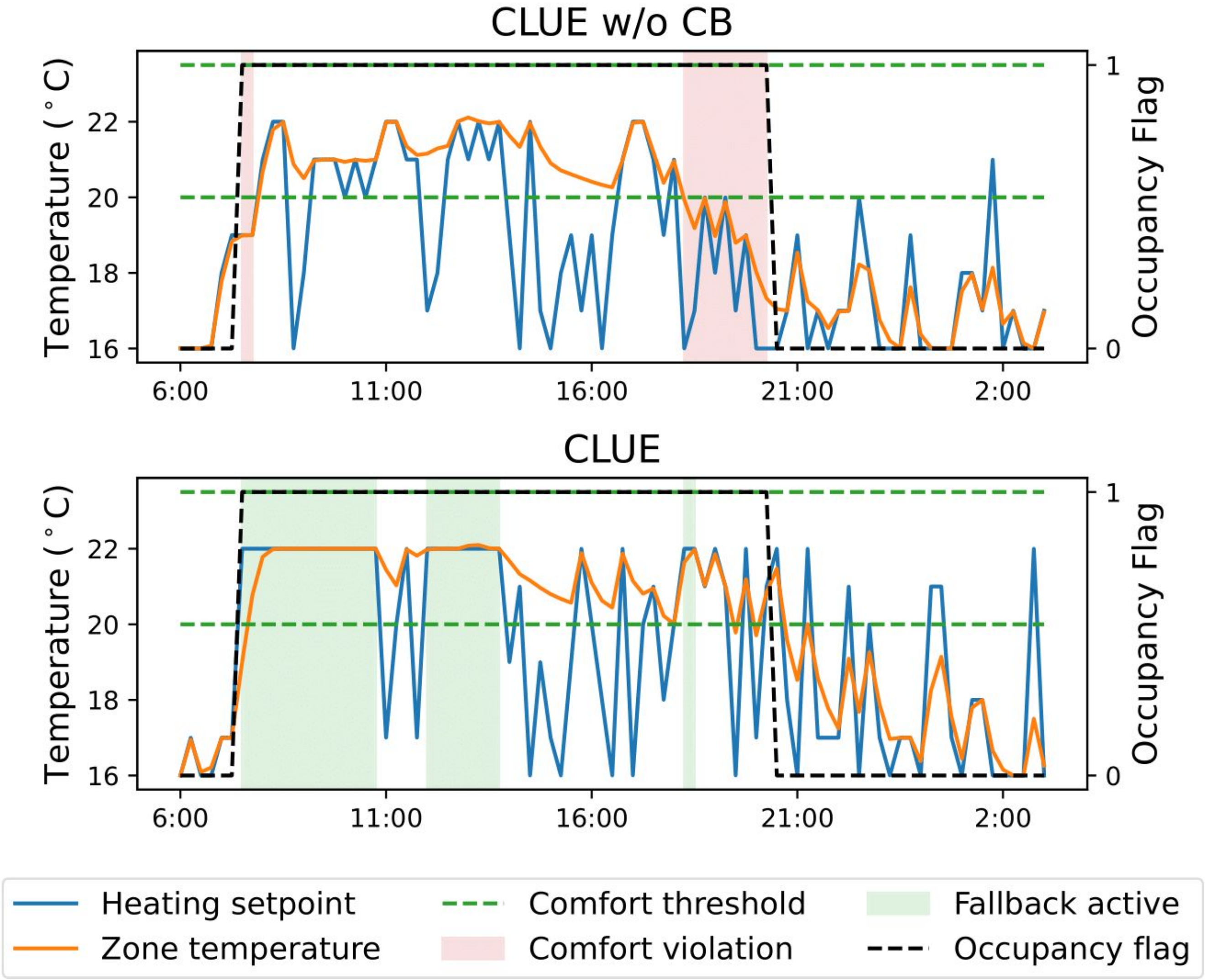
THE UNIVERSITY OF CALIFORNIA MERCED
2005

Uncertainty estimation

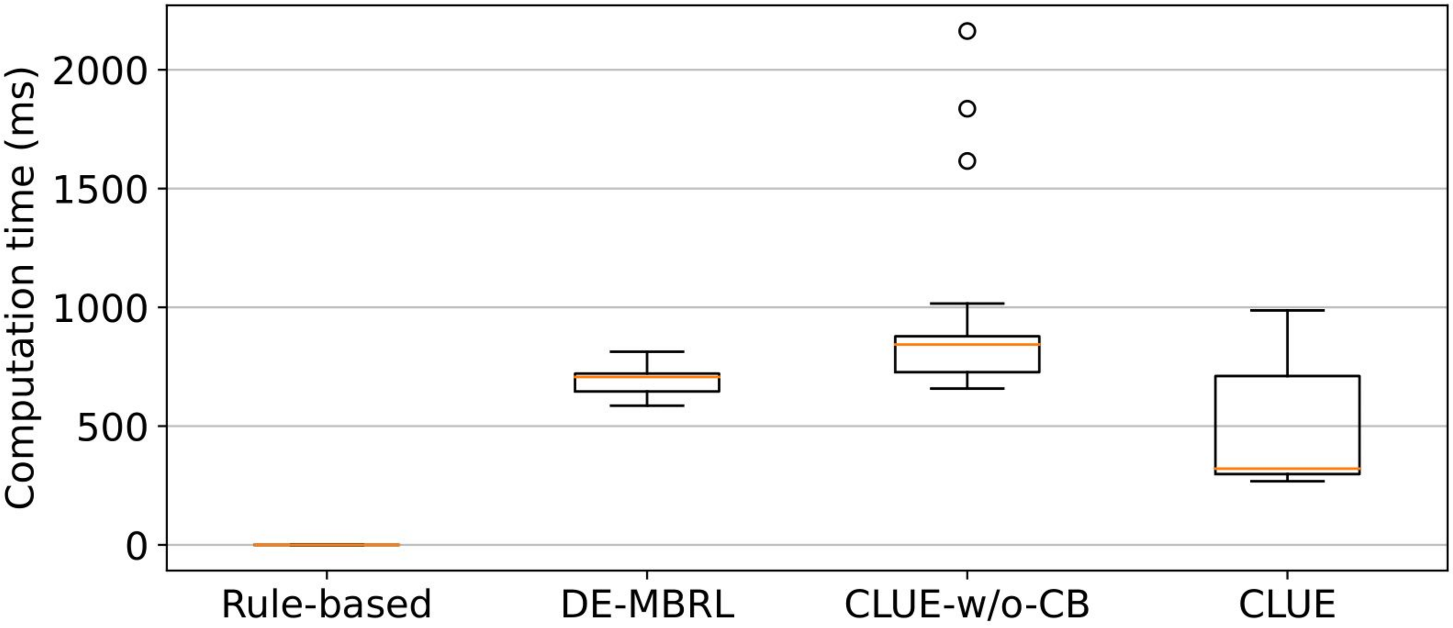


Location	Time	Deep Ensemble [11]			GP			GP-M (ours)		
		Accuracy	Precision	Recall	Accuracy	Precision	Recall	Accuracy	Precision	Recall
Pittsburgh, PA	January	.796±.00	.521±.01	.740±.01	.877±.00	**.803±.00**	**.958±.00**	**.884±.00**	.768±.00	.677±.00
	July	.831±.01	**.851±.09**	.160±.10	.840±.00	.809±.00	.763±.00	**.961±.00**	.056±.00	**.999±.00**
Tucson, AZ	January	.736±.01	.439±.08	.693±.12	.847±.00	**.697±.00**	**.844±.00**	**.932±.00**	.341±.00	.694±.00
	July	.650±.00	.489±.00	.827±.00	.844±.00	**.854±.00**	.860±.00	**.947±.00**	.036±.00	**.999±.00**
New York, NY	January	.830±.00	.403±.02	.816±.00	.855±.00	**.883±.00**	.728±.00	**.965±.00**	.299±.00	**.900±.00**
	July	.679±.00	.373±.01	.812±.01	.797±.00	**.934±.00**	.718±.00	**.953±.00**	.205±.00	**.947±.00**

Performance analysis

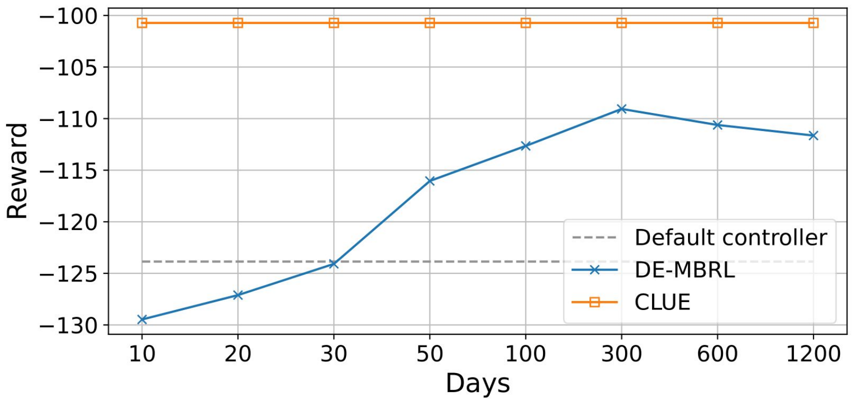


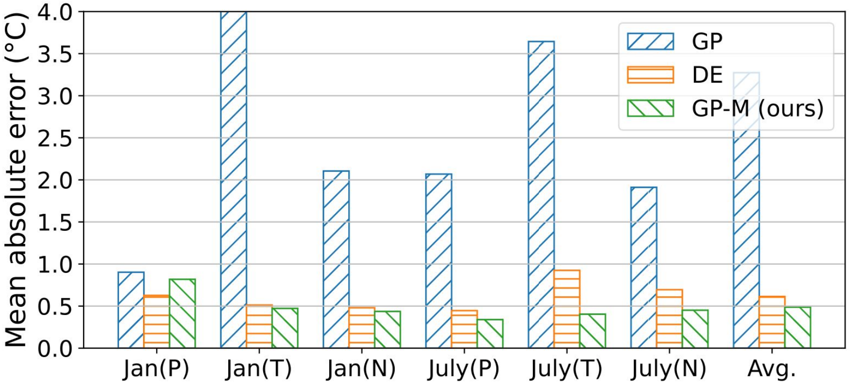
Computation overhead



Experiment - Data Efficiency

Setting | Result



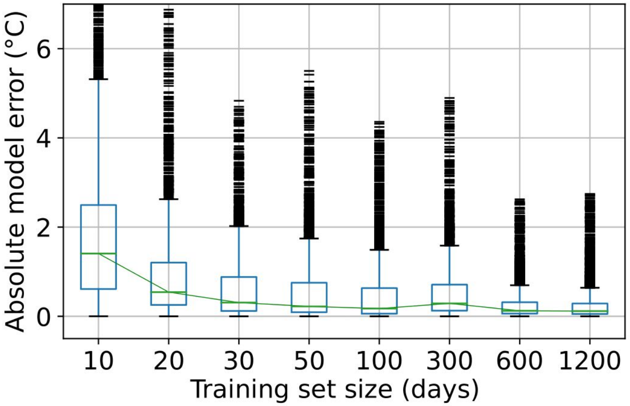


CLUE converges >30x faster than previous SOTA

Produces comparable accuracy given the same data

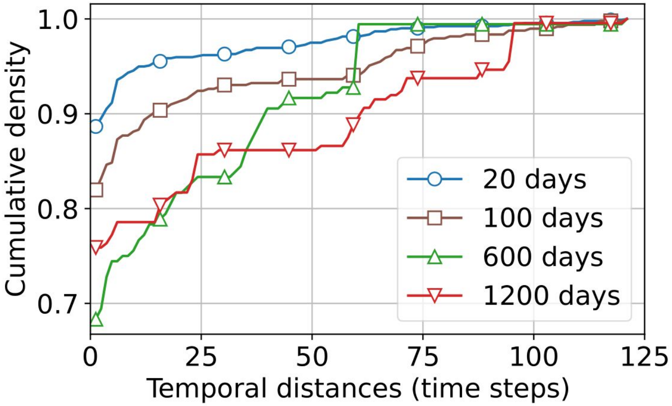
Motivation Experiments

1. Can we mitigate high model errors by training on more data?



No, high model errors persists even after very large dataset is used

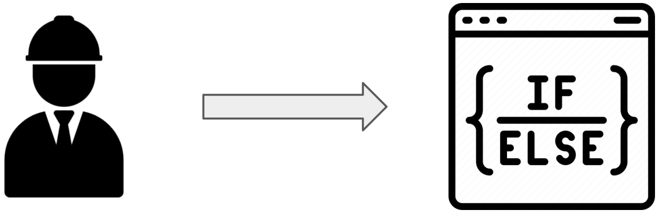
2. Can we let the building system tolerate short periods of controller glitches?



No, high model errors often appear in clusters

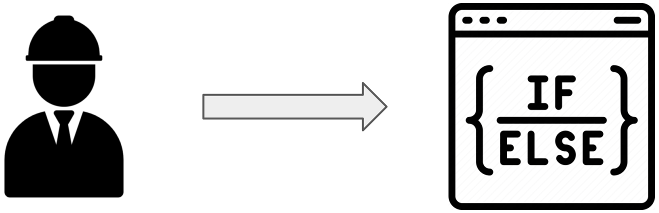
Preliminaries: HVAC control methods

1. Human designed rules



Preliminaries: HVAC control methods

1. Human designed rules

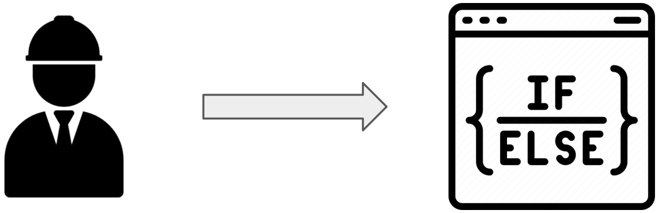


Baseline

Not energy efficient

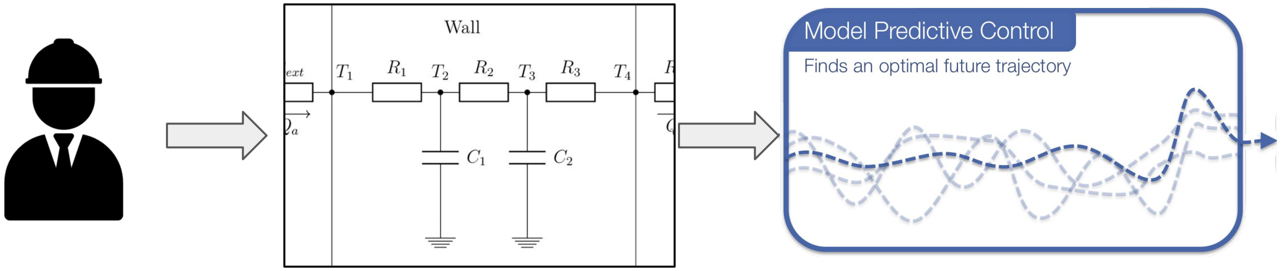
Preliminaries: HVAC control methods

1. Human designed rules



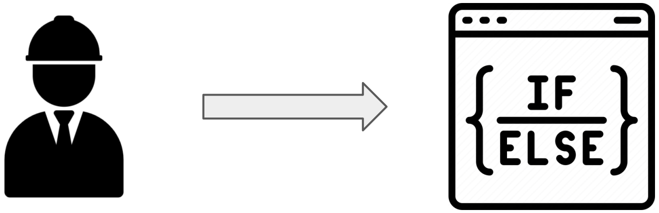
Baseline
Not energy efficient

2. Simple thermal dynamics model + optimizer



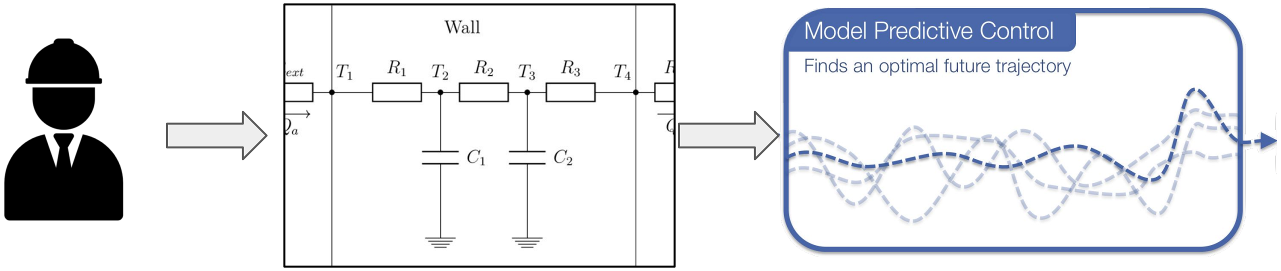
Preliminaries: HVAC control methods

1. Human designed rules



Baseline
Not energy efficient

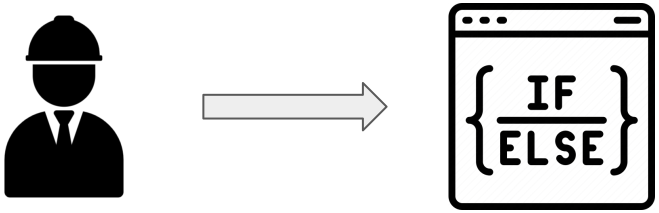
2. Simple thermal dynamics model + optimizer



Model not accurate
Laborious

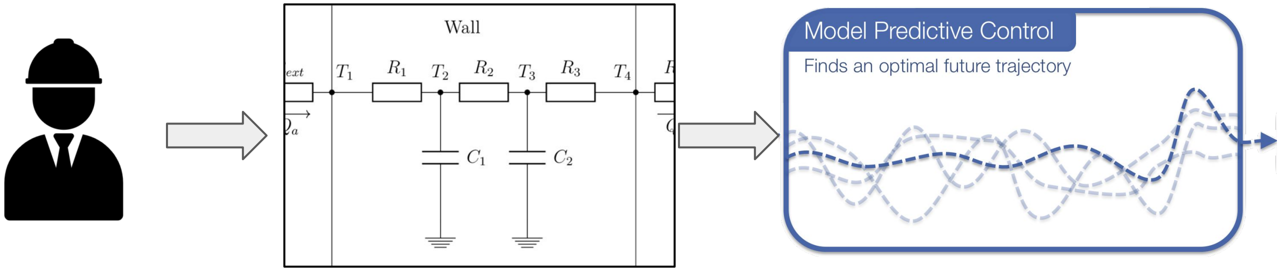
Preliminaries: HVAC control methods

1. Human designed rules



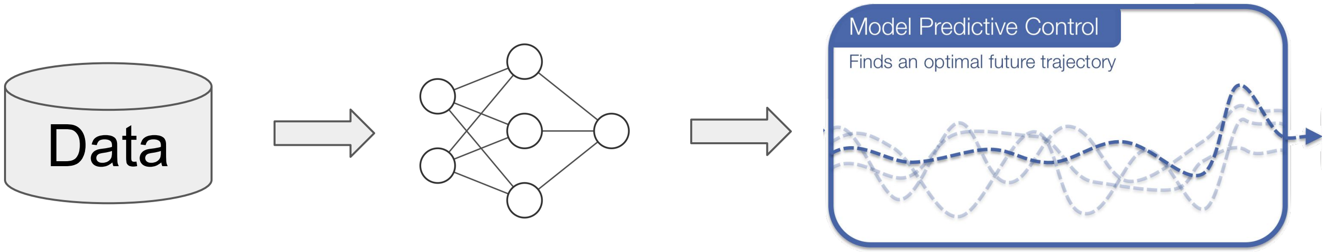
Baseline
Not energy efficient

2. Simple thermal dynamics model + optimizer



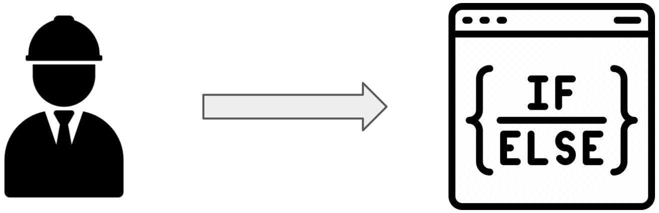
Model not accurate
Laborious

3. Neural dynamics model + optimizer



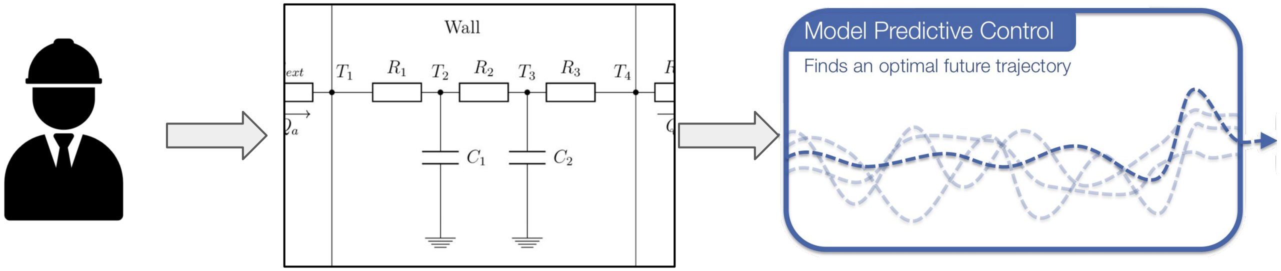
Preliminaries: HVAC control methods

1. Human designed rules



Baseline
Not energy efficient

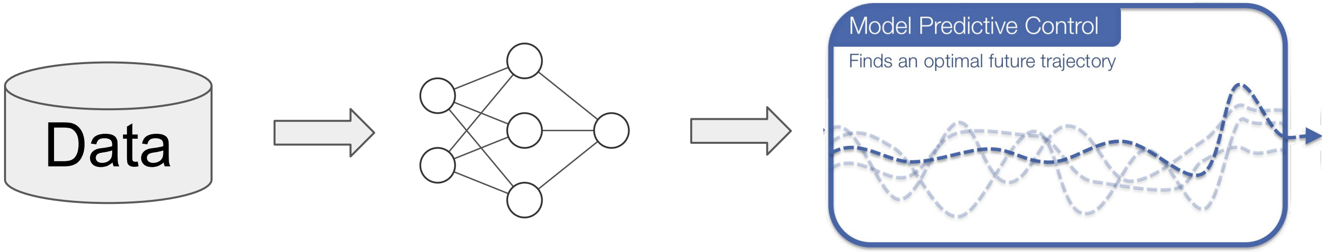
2. Simple thermal dynamics model + optimizer



Model Predictive Control
Finds an optimal future trajectory

Model not accurate
Laborious

3. Neural dynamics model + optimizer



Data

Model Predictive Control
Finds an optimal future trajectory

Autonomous
Current SOTA