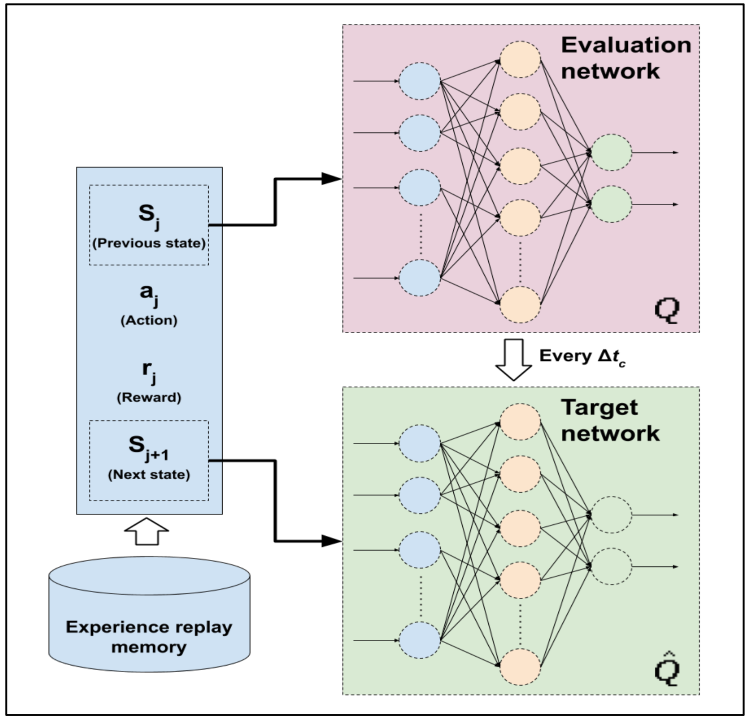
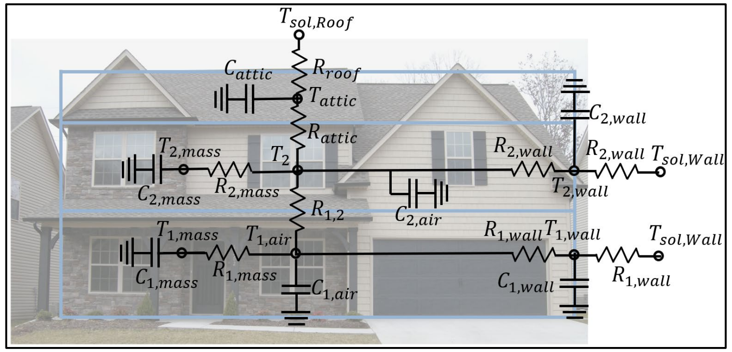
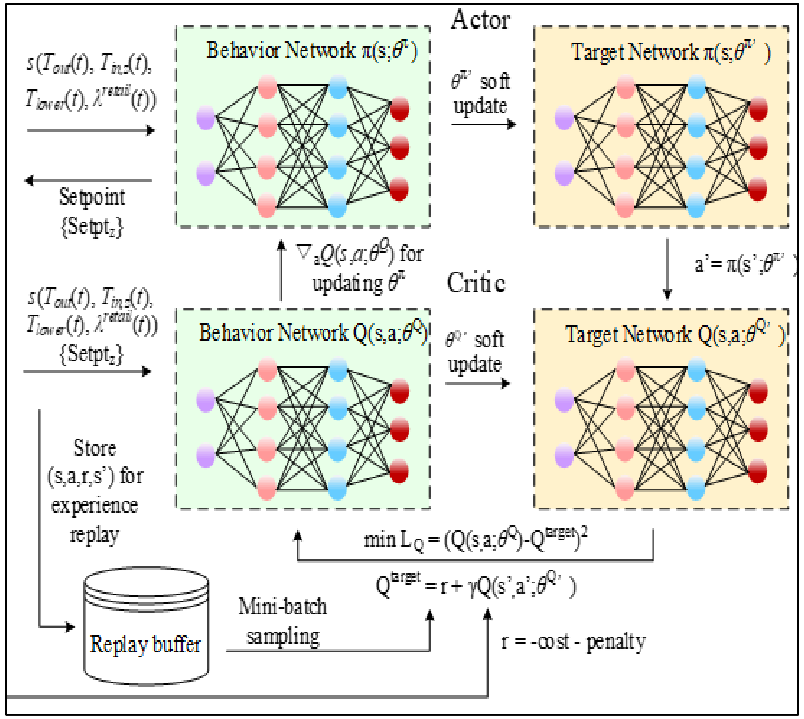
Scalable Load Management using Reinforcement Learning







Oak Ridge National Laboratory (ORNL)
Helia Zandi, R&D Staff
zandih@ornl.gov

Project Summary

Timeline:
Start date: 10/9/2018
Planned end date: 9/30/2021

Key Milestones
1. Develop, formulate, and test the RL algorithm in simulation;9/30/2019
2. Test the scalability of the developed load management system;9/30/2020
3. Compare RL performance with a golden standard optimization technique;6/30/2021

Budget:

Total Project $ to Date:
- DOE: $1,824 K
- Cost Share: $0

Total Project $:
- DOE: $2,100 K
- Cost Share: $0

Key Partners:

Southern Company	University of Tennessee, Knoxville

Oak Ridge National Laboratory · Southern Company · The University of Tennessee Knoxville

Project Outcome:

- ❑ A scalable load management system that can be deployed by utilities on grid existing infrastructure
- ❑ A learning-based optimization algorithm that can be applied to existing homes and new construction with minimal effort and minimal additional devices

Team

Oak Ridge National Laboratory

Principal Investigator

Helia Zandi

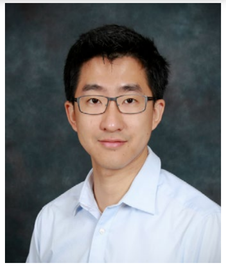
RL Algorithm Development

Kuldeep Kurte

Kadir Amasyali

Building & Equipment Modeling

Jeffrey Munk



Borui Cui

Software Development

Michael Walsh

Jesse McGaha

Subprogram Manager

Teja Kuruganti

Southern Company

Justin Hill

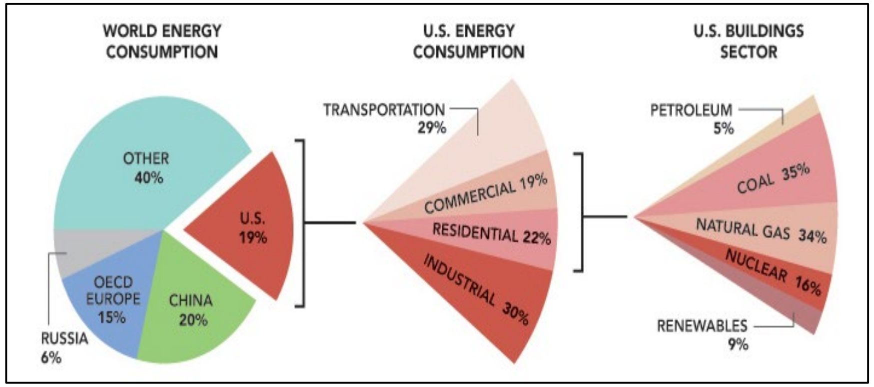
The University of Tennessee Knoxville

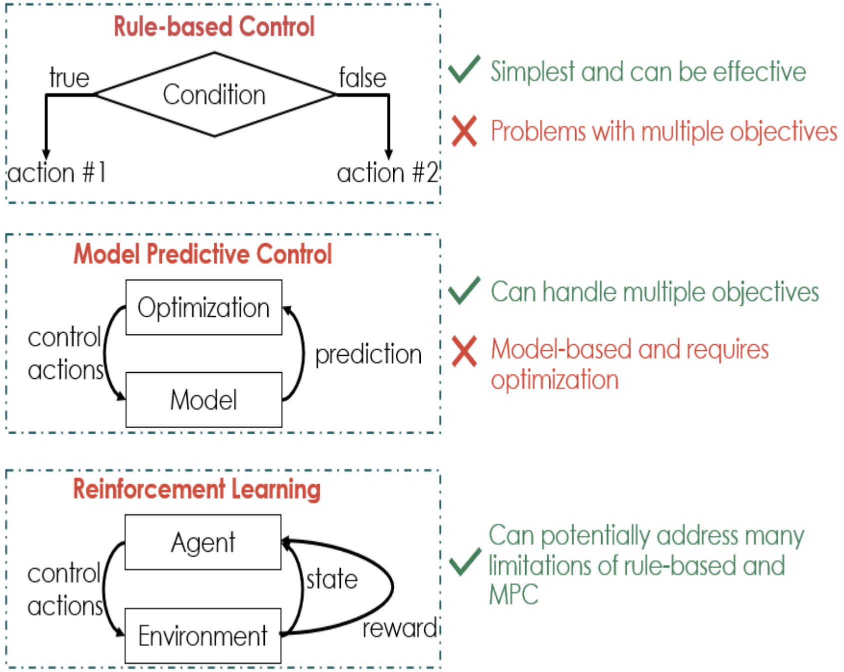
Fran Li

Challenge: The Effect of Emerging Technologies on the Electric Grid

- Over **75 billion** connected devices predicted to be in use by 2025
- There is an immense opportunity for a management system which can **control and coordinate the power use** of these devices
- **41%** of the energy consumption in the United States is from buildings
- Advanced sensing and controls have the potential to save energy in buildings up to **40%**

What does the electric grid need?





Approach: Project Overview

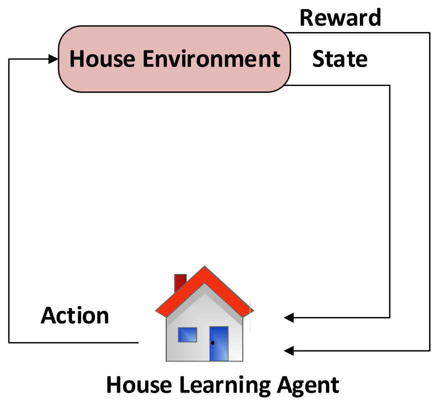
Design, develop, and field evaluate a *scalable* and *cost-effective* load management system using **Reinforcement Learning (RL)**

Project Objectives

Objective 1: Develop Reinforcement Learning-based optimization and control methods for understanding energy use patterns and for load scheduling

Objective 2: Develop a scalable load management system to access flexibility in loads

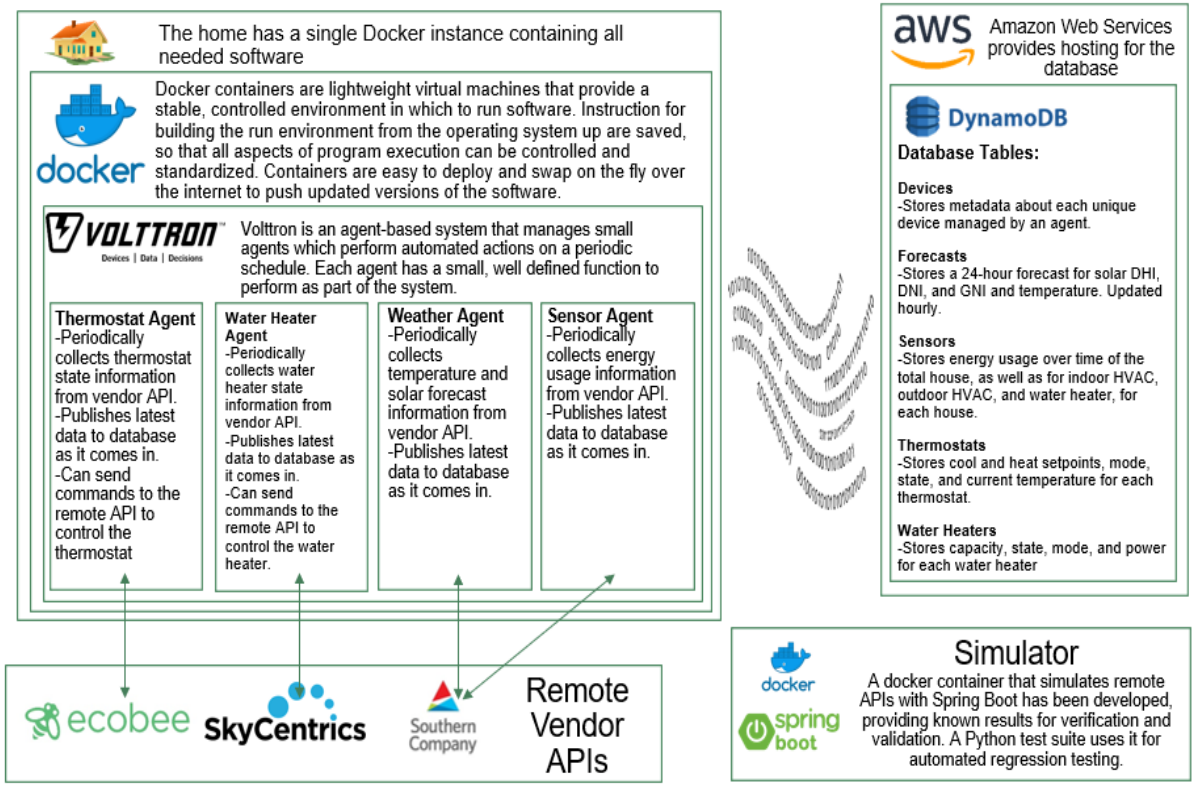
Objective 3: Perform field validation of the software framework and demonstrate benefits of running RL-based optimization and control in residential buildings

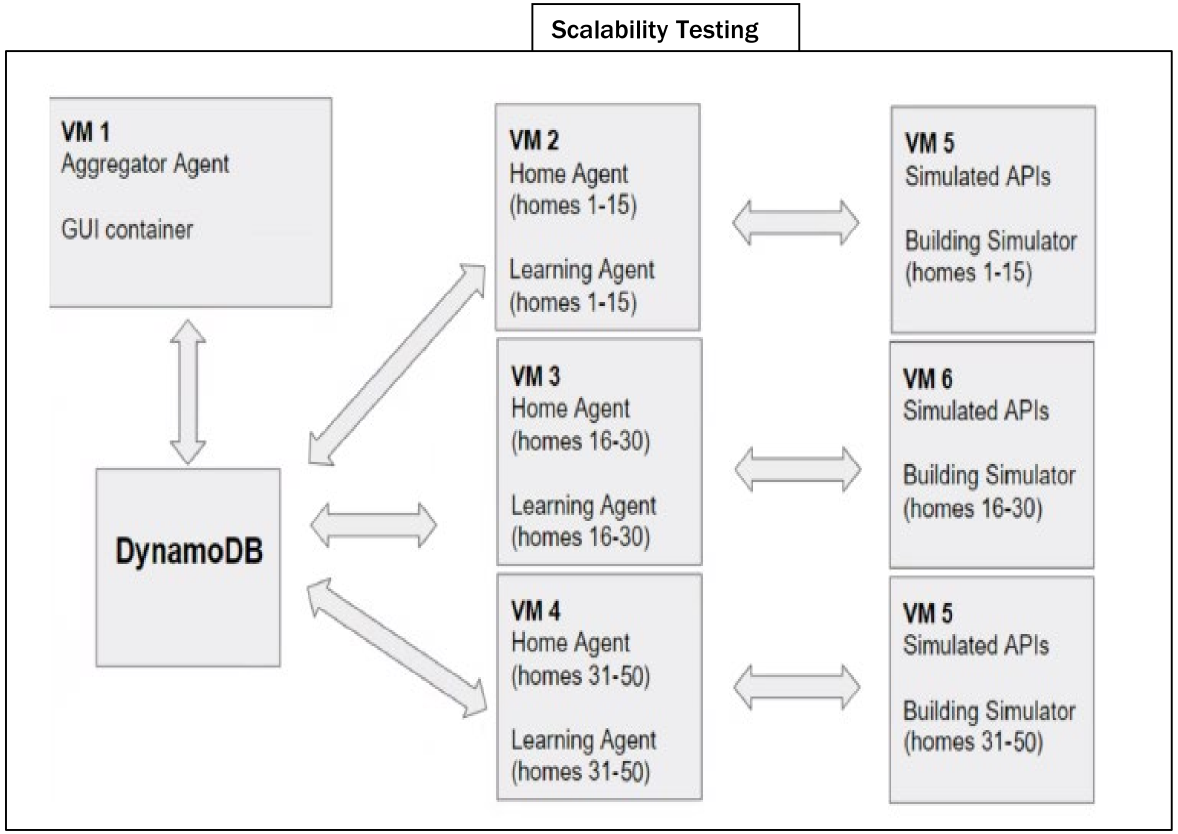


ORNL Yarnell Station Research House

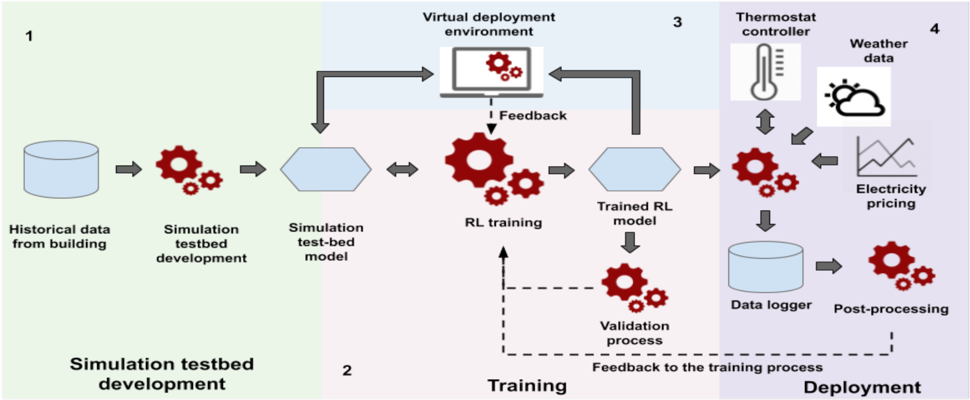
Load Management System Software Architecture

- Hierarchical cloud-based multi-agent system (MAS) architecture
- Critical data are identified



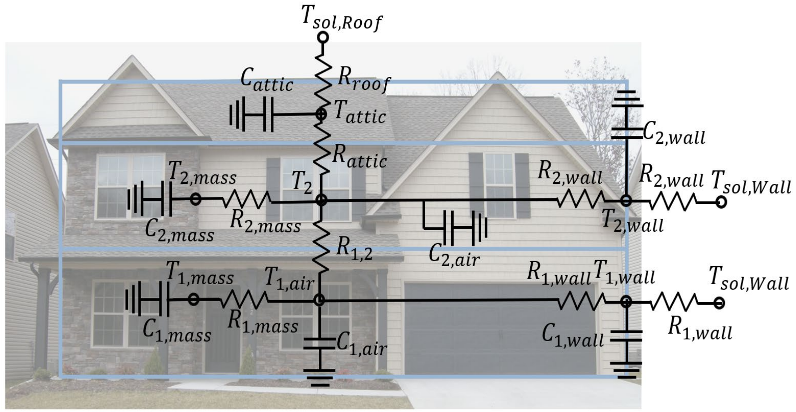


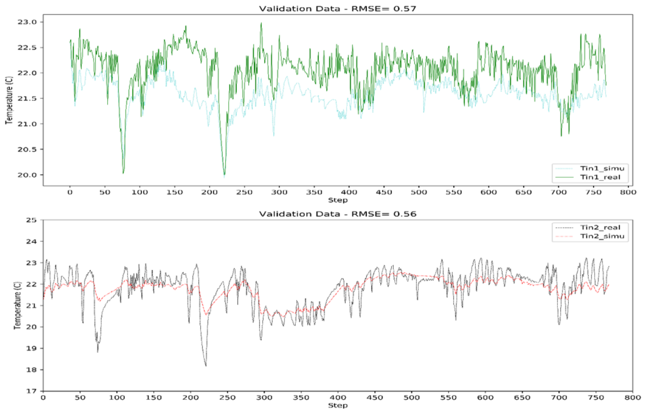
Approach: Control Development Approach



Model Development

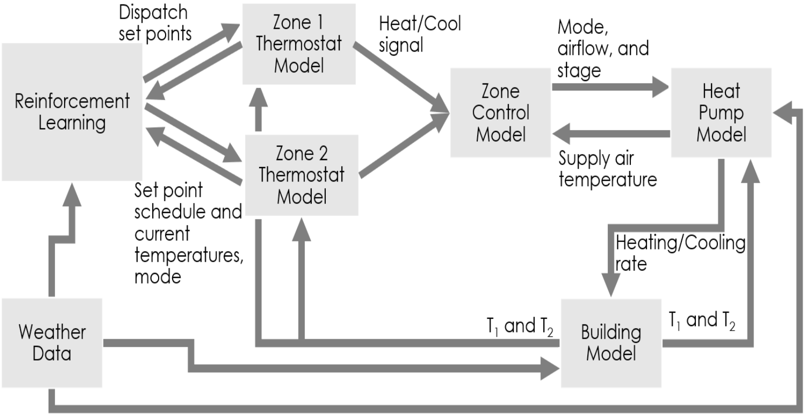
1. Single-zone single-family house
2. Two-zone single-family house
3. Single-stage HVAC
4. 2-zone HVAC
5. Water heater





Simulation Testbed

1. Single-zone building, single-stage HVAC simulation testbed
2. 2-zone building, 2-zone HVAC simulation testbed



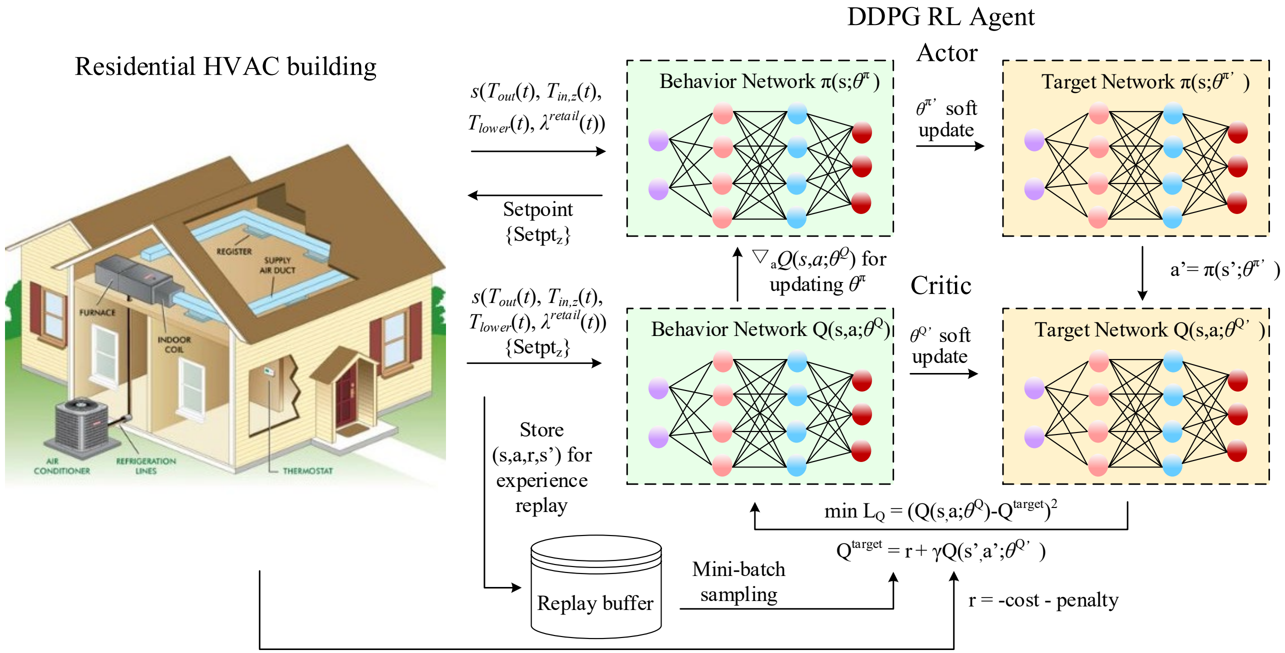
Algorithm Development

Developed and evaluate two different model-free deep reinforcement learning approach :

1. Deep Q-Network(DQN)
2. Deep Deterministic Policy Gradient (DDPG)

Approach: Reinforcement Learning

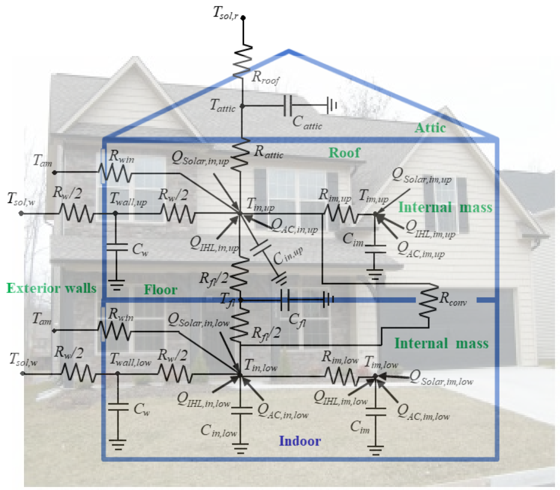
- ❑ Reinforcement learning (RL) method is a type of machine learning method that optimizes the decision-making strategy of an agent within an unknown environment
- ❑ RL uses a Markov Decision Process (MDP)
- ❑ There are mainly two types of RL method:
  - Value-based RL method: estimates Q-value of a state-action pair
  - Policy-based RL method: generate probability of all feasible action for current state

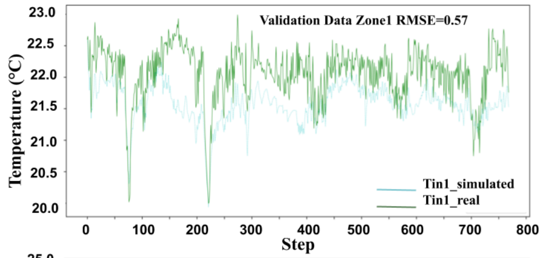


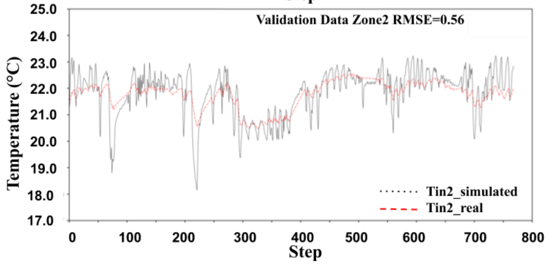
•Yan Du, Fangxing Li, Jeffrey Munk, Kuldeep Kurte, , Kadir Amasyali, Olivera Kotevska, Helia Zandi "Intelligent Multi-zone Residential HVAC Control Strategy based on Deep Reinforcement Learning" Applied Energy 2021, https://doi.org/10.1016/j.apenergy.2020.116117
•Yan Du, Fangxing Li, Jeffrey Munk, Kuldeep Kurte, , Kadir Amasyali, Olivera Kotevska, Helia Zandi "Multi-task Deep Reinforcement Learning for Intelligent Multi-zone Residential HVAC Control," Electric Power Systems Research 2021, https://doi.org/10.1016/j.epsr.2020.106959
•Kuldeep Kurte, Jeffrey Munk, Kadir Amasyali, Olivera Kotevska, Robert Smith, Evan McKee, Yan Du, Bori Cui, Teja Kuruganti, Helia Zandi "Evaluating the Adaptability of Reinforcement Learning based HVAC Control for Residential Houses" *MDPI Sustainability as part of the Special Issue Building and Urban Energy Prediction-Big Data Analysis and Sustainable 2020,* https://doi.org/10.3390/su12187727

Approach: HVAC Optimization for 2-Zone Single Family Building

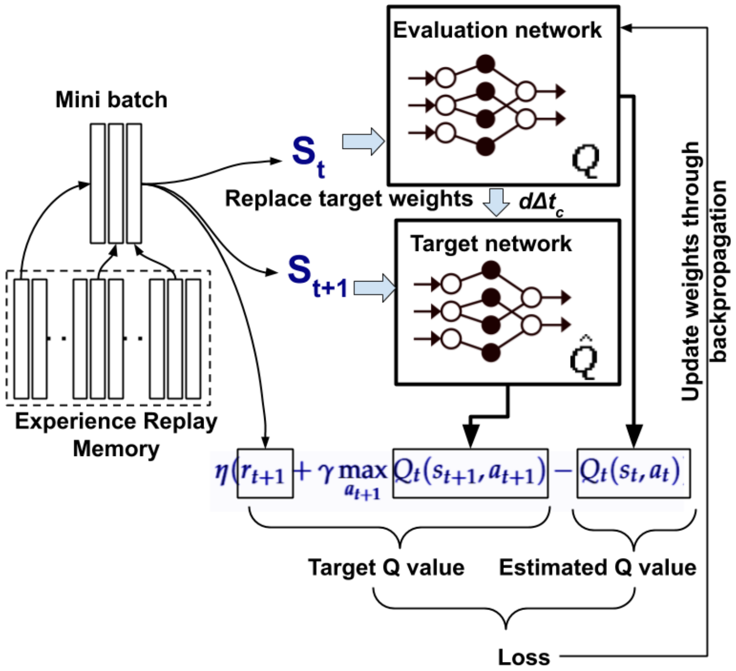
Building model







Neural network-based Q-learning



RL Model Development Based on DQN

State space

$$S_t = \{t, T_t^{in}, T_t^{out}, P_t\}$$

$$P_t = \{pt, pt_{+1,\ldots}, pt_{+k}, \ldots\}$$

Action space

$$A = \{a_0, a_1, a_2, a_3\}$$

$$SetPoint_t = \begin{cases} T_t^{in} - \Delta T & \text{if } action_t \equiv 1 \\ T_t^{in} + \Delta T & \text{if } action_t \equiv 0 \end{cases}$$

Actions	AC ON/OFF Zone-1	AC ON/OFF Zone-2	Set Point-1	Set Point-2
a_0	0	0	$T_t^{in} + \Delta T$	$T_t^{in} + \Delta T$
a_1	0	1	$T_t^{in} + \Delta T$	$T_t^{in} - \Delta T$
a_2	1	0	$T_t^{in} - \Delta T$	$T_t^{in} + \Delta T$
a_3	1	1	$T_t^{in} - \Delta T$	$T_t^{in} - \Delta T$

Reward function

$$r_t = -w_1 * cost_{\{t-\Delta t_c\}}^{avg} - w_2 * \left[\left(T_{[t-\Delta t_c,t]}^{avg} - UT\right) + \left(LT - T_{[t-\Delta t_c,t]}^{avg}\right)\right]$$

$$R_t = \sum_{t'=1}^{\infty} \gamma^{t'-t} r_{t'} \quad \ldots..(1)$$

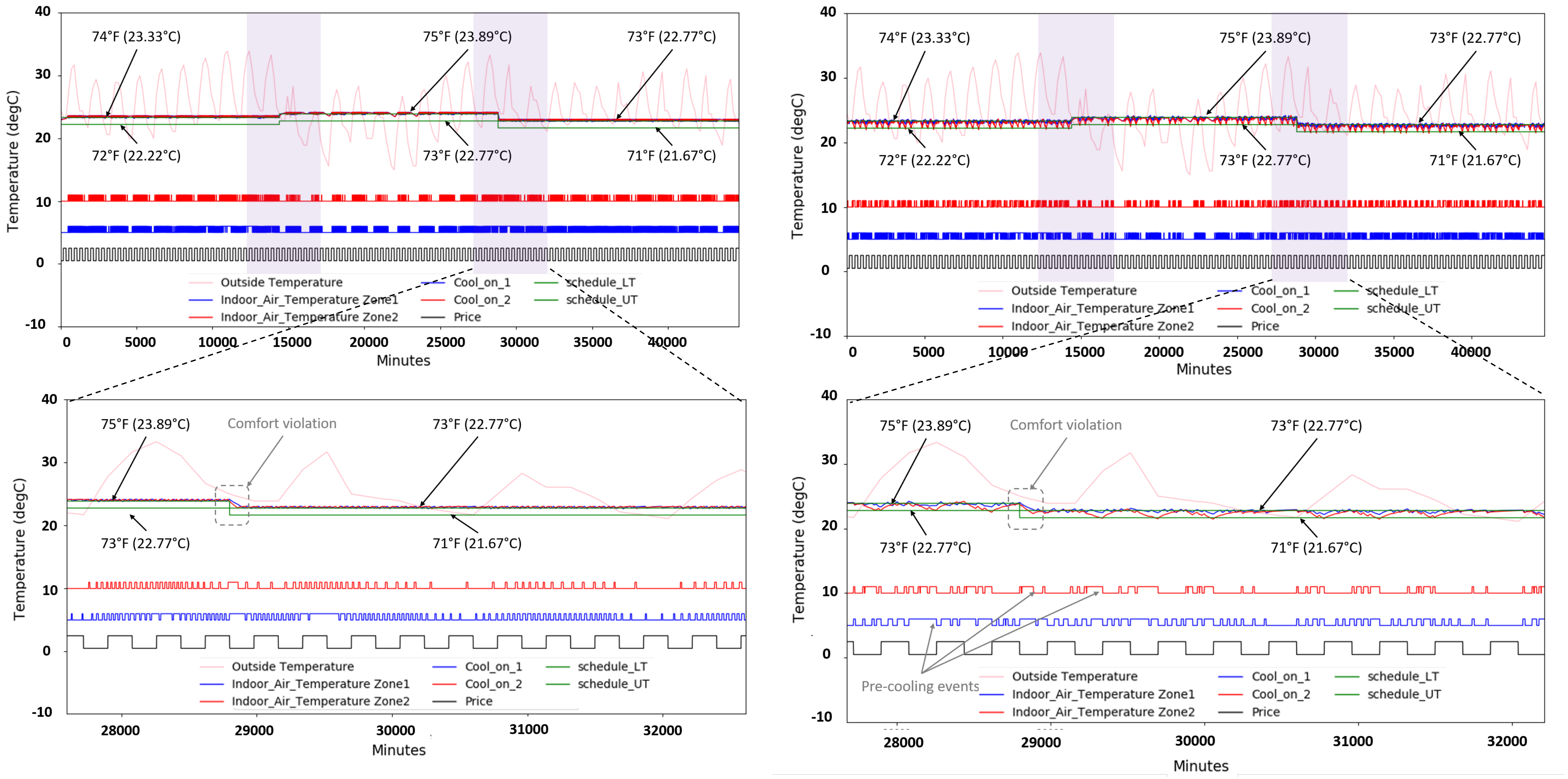
$$Q^*(s_t, a_t) = \max_{\pi} E[R_t | s_t, a_t] \quad \ldots..(2)$$

$$Q^*(s_t, a_t) = E\left[r_{t+1} + \gamma \max_{a_{t+1}} Q^*(s_{t+1}, a_{t+1}) \middle| s_t, a_t\right] \ldots..(3)$$

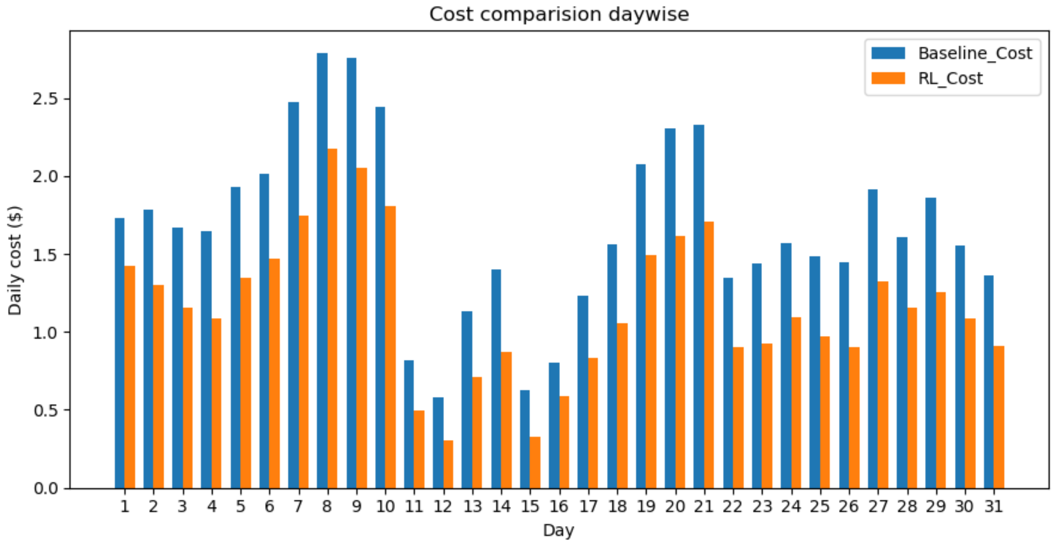
$$Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + \eta\left(r_{t+1} + \gamma \max_{a_{t+1}} Q_t(s_{t+1}, a_{t+1}) - Q_t(s_t, a_t)\right) \ldots...(4)$$

Kuldeep Kurte, Jeffrey Munk, Kadir Amasyali, Olivera Kotevska, Robert Smith, Evan McKee, Yan Du, Borui Cui, Teja Kuruganti, Helia Zandi "Evaluating the Adaptability of Reinforcement Learning based HVAC Control for Residential Houses" *Sustainability as part of the Special Issue Building and Urban Energy Prediction-Big Data Analysis and Sustainable 2020,* https://doi.org/10.3390/su12187727

Progress: Validation-1: with the Same House as a Simulation



Progress: Validation-1: with the Same House as a Simulation



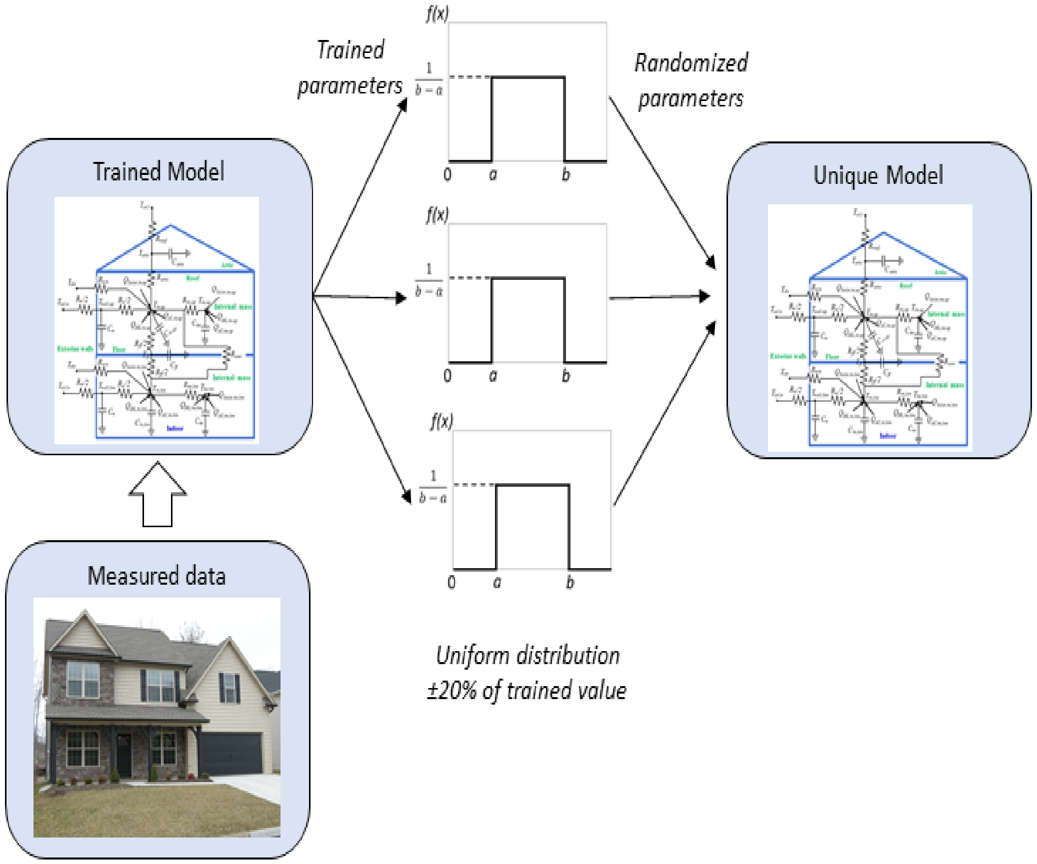
Day wise cost comparison of operating HVAC with cooling set point (baseline) and pre-trained RL model

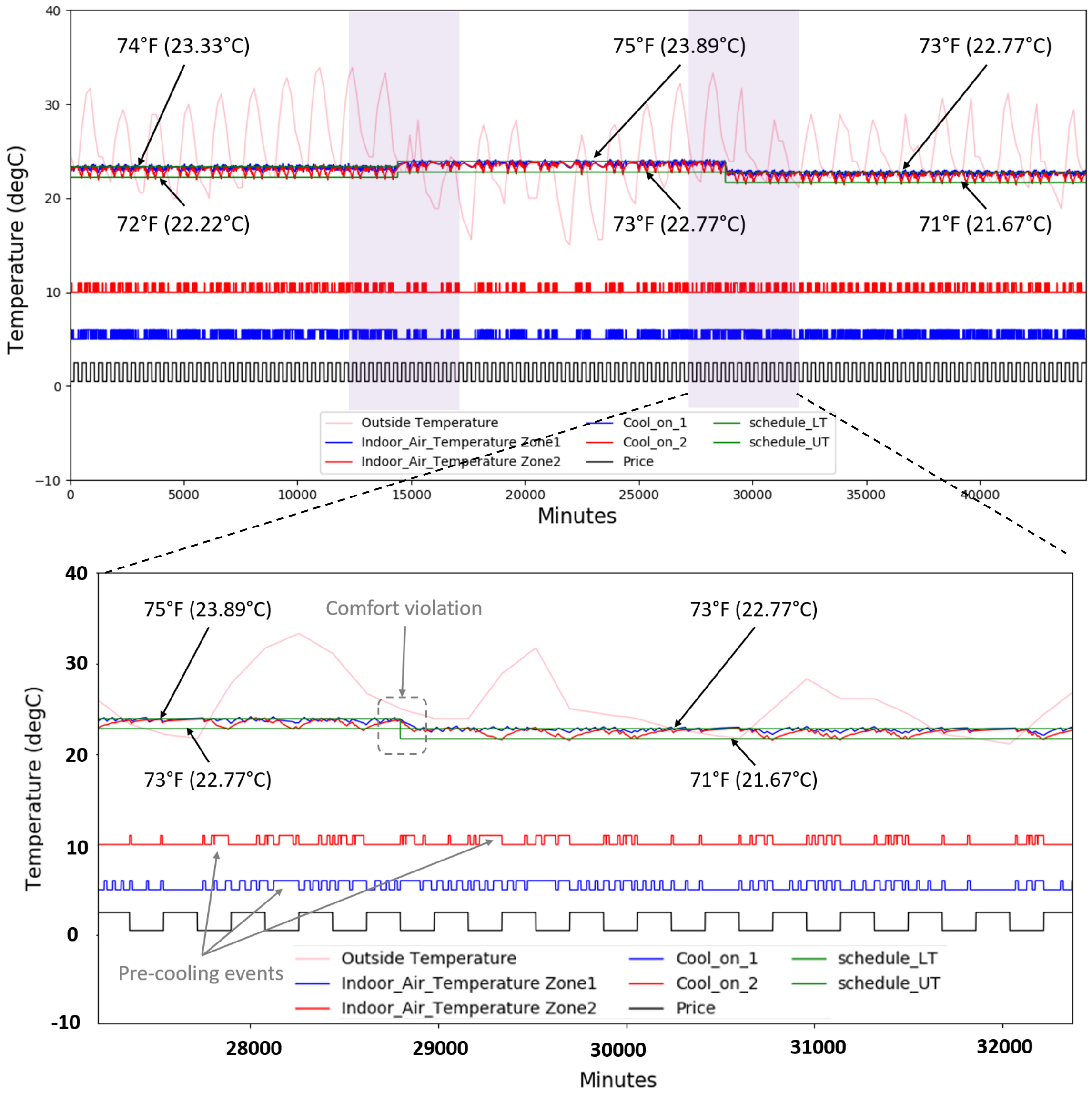
Cooling Set Point →		73 °F	74 °F	75 °F	(74, 75, 73) °F
Baseline	Energy (kWh)	368.01	333.49	301.37	334.84
	Cost ($)	56.94	51.91	46.74	51.69
	MOC (min)	0	0	0	97
RL	Energy (kWh)	430.89	390.96	352.72	393.97
	Cost ($)	40.10	35.46	31.58	36.09
	MOC (min)	0	0	0	79
% Cost reduction		29.57%	31.68%	32.43%	30.17%

Comparison of the cost of operation, energy consumption and MOC for a fixed cooling setpoint baseline and pre-trained

Adaptability Testing

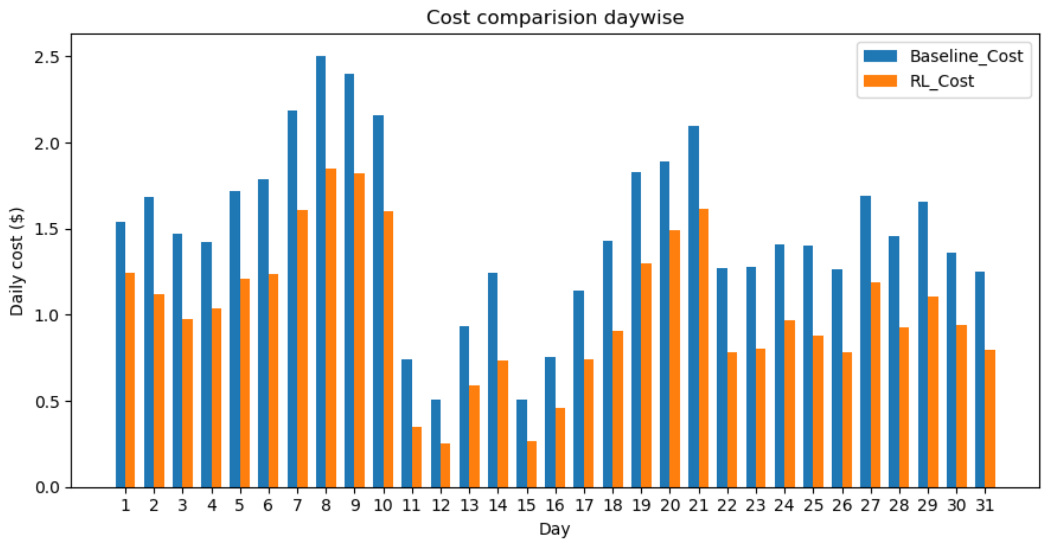
Generating unique building models
Validating with synthetic houses





Validation with synthetic house and real house

Validation with synthetic house



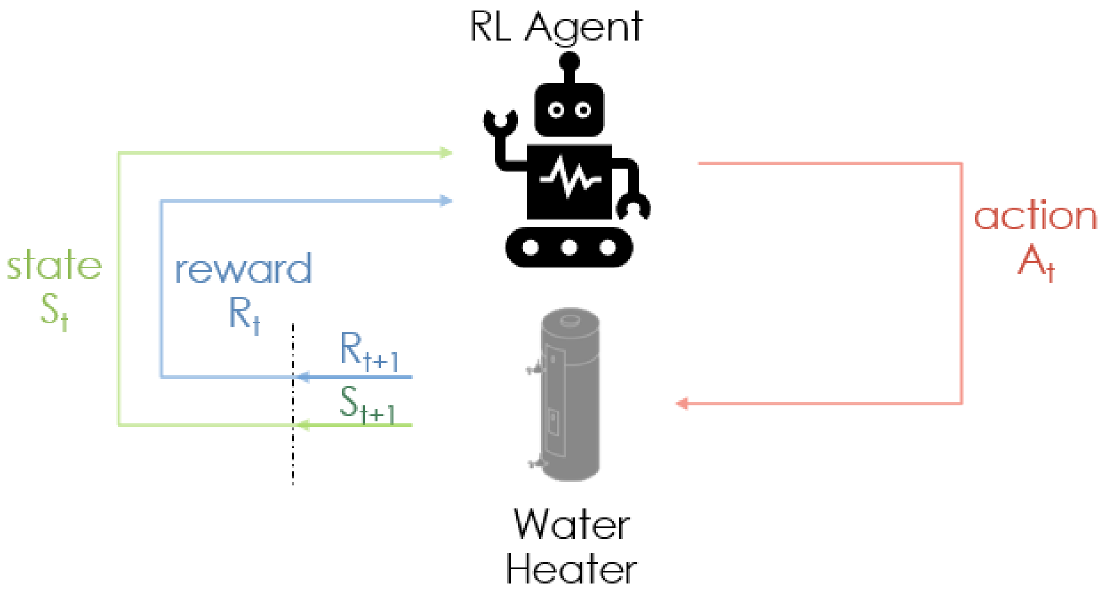
Cooling Set Point →		73 °F	74 °F	75 °F	(74, 75, 73) °F
Baseline	Energy (kWh)	330.15	297.72	269.20	299.61
	Cost ($)	50.94	46.00	41.81	45.98
	MOC (min)	0	0	0	97
RL	Energy (kWh)	390.75	352.36	317.71	358.73
	Cost ($)	35.17	30.94	27.41	31.57
	MOC (min)	0	0	0	69
% Cost reduction		30.9%	32.72%	34.44%	31.33%

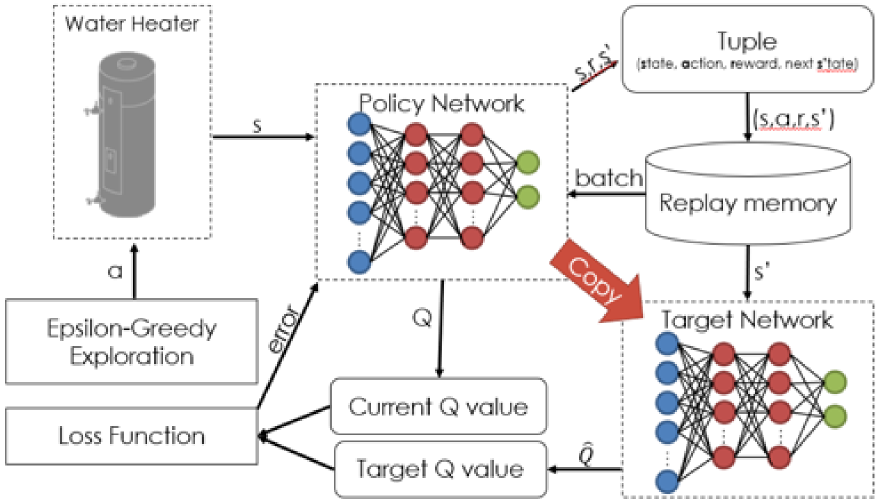
Validation with real house

- ❑ Deployed with Model Pre-Trained on the :
  - ▪ Deployed House Data
  - ▪ Synthetic house data
- ❑ Demonstrated an average saving of 20%

Approach: WH Optimization

<u>Focus:</u> Control a hybrid water heater using demand response (DR) commands for minimum electricity cost under a time-of-use (TOU) electricity pricing policy. rates





State (s)
- Water temperatures
- Electricity prices
- Hot water usages

Action (a)
- Load up
- Normal operation
- Shed

Reward (r)
- (-)Electricity cost

- Kadir Amasyali, Kuldeep Kurte, Jeffrey Munk, Olivera Kotevska, Robert Smith, Helia Zandi " Double Deep Q-Networks for Optimizing Electricity Cost of a Water Heater", *2021 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, Washington, DC, USA, 2021, pp. 1-5, doi: 10.1109/ISGT49243.2021. 9372205.

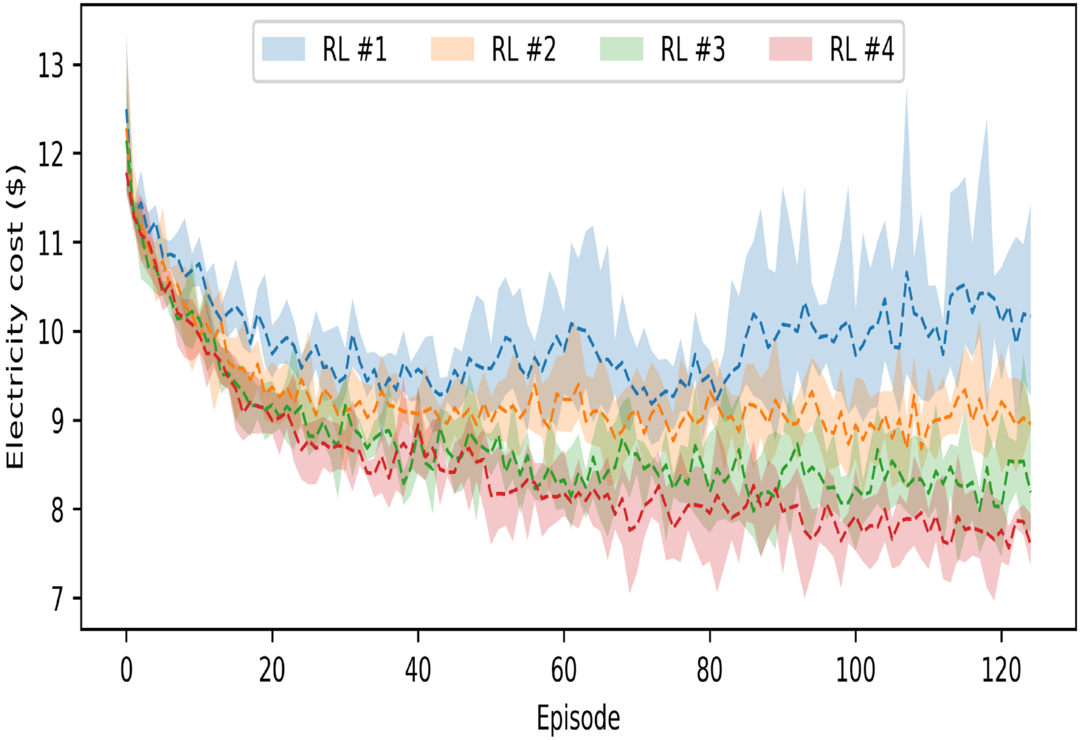
Progress: RL Training

RL agent #1 (RL #1)
- node temperatures
- Electricity prices of **next 30 minutes**

RL agent #2 (RL #2)
- node temperatures
- Electricity prices of next 30 minutes
- **Hot water usage volumes of next 30 minutes**

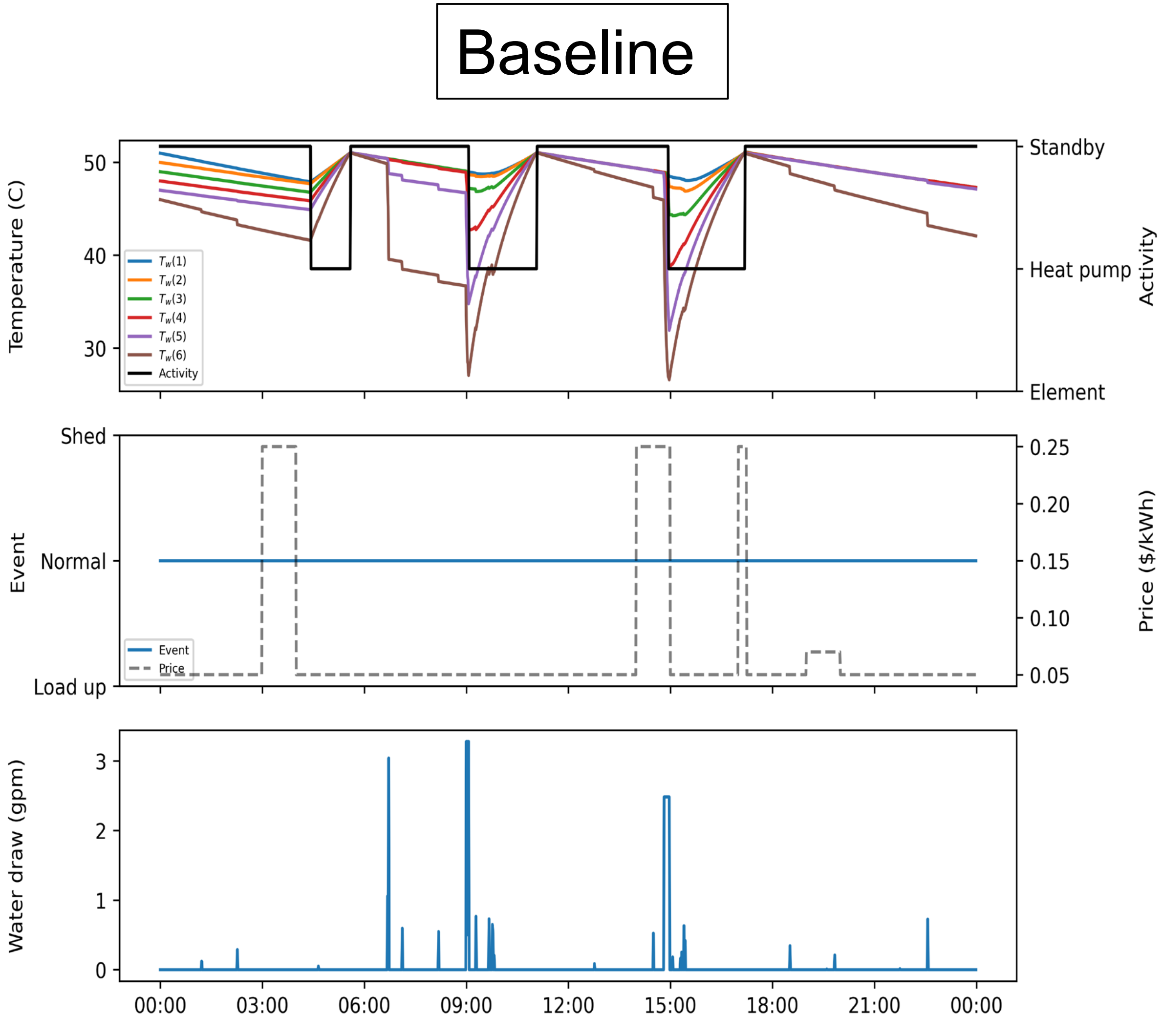
RL agent #3 (RL #3)
- node temperatures
- Electricity prices of **next hour**
- Hot water usage volumes of **next hour**

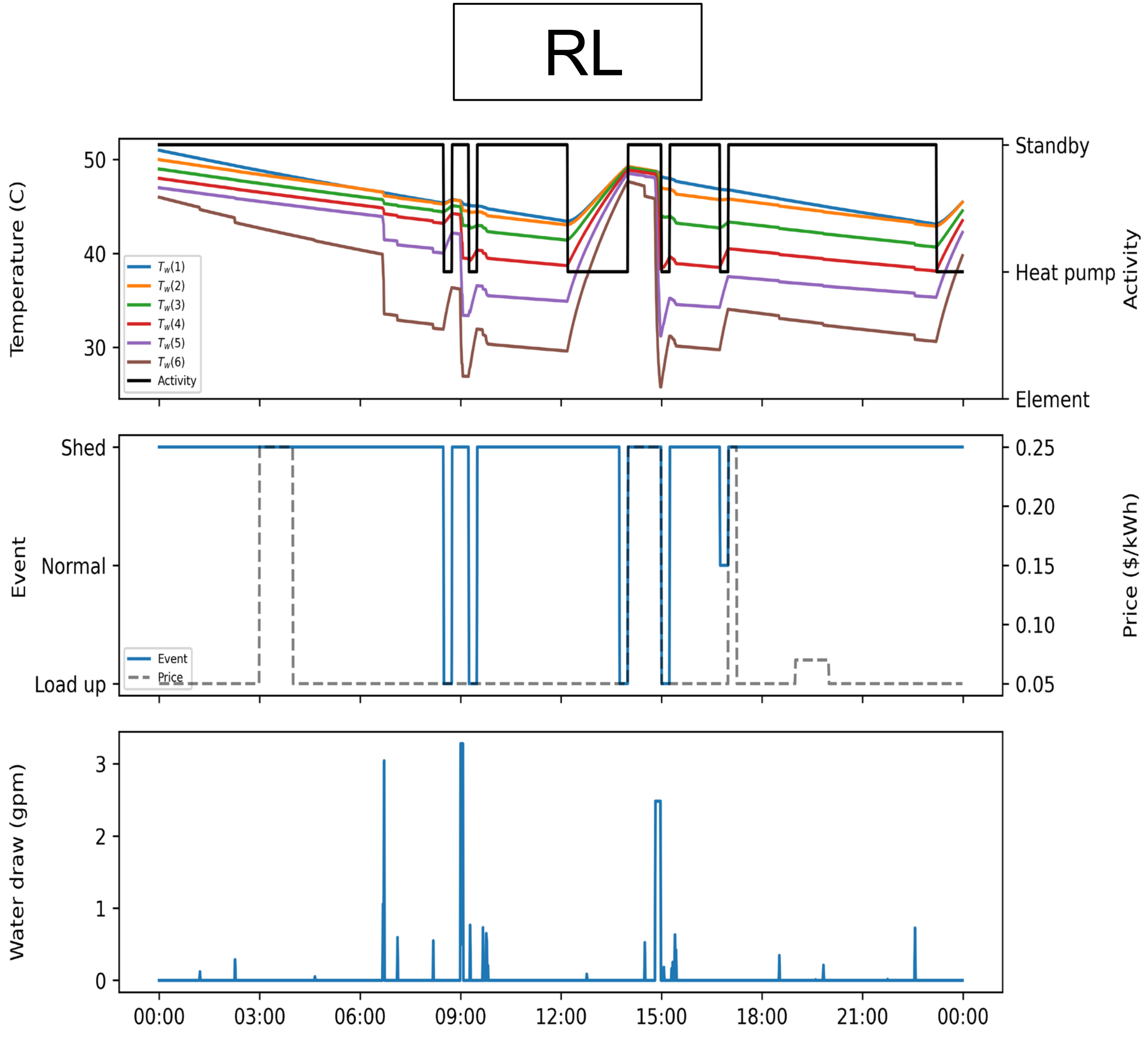
RL agent #4 (RL #4)
- node temperatures
- Electricity prices of **next two hours**
- Hot water usage volumes of **next two hours**



Operation strategy	Look Ahead	Electricity Cost	Element Usage	Average COP_{HP}
Baseline	N/A	$1.36	75 min	4.88
RL agent #2	30 min	$1.21	89 min	5.25
RL agent #3	1 hour	$1.03	82 min	5.47
RL agent #4	2 hours	$0.91	53 min	5.15
MPC #2	30 min	$1.31	157 min	5.76
MPC #3	1 hour	$1.04	104 min	5.70
MPC #4	2 hours	$0.94	84 min	5.61
Day-ahead* optimization	5 days	$0.81	47 min	5.48

Baseline vs RL





Progress: WH Deployment Results

Operation strategy	March 27	March 28	March 29	March 30
RL	**$0.07**	**$0.14**	**$0.09**	**$0.17**
Baseline	$0.11	$0.23	$0.11	$0.26

Cost

- In all days, the RL cost less than the baseline
- RL saved in the range of 18% to 39%
- RL achieved more savings when hot water demand is higher

Comfort

- The upper tank is reserved for users by the manufacturer and are not available for any DR command
- In both cases, the upper tank temperature never went below the comfort range

Impact

- ❑ Developed and demonstrated a learning-based load management system
  - Addressing the need for a scalable Grid-interactive Efficient Buildings (GEB) platform
  - Delivering towards BTO strategy for GEB - seamless connectivity between different devices and utilization of optimization and learning algorithm for optimal scheduling of residential loads.
- ❑ The data-driven RL algorithm addresses the need for:
  - Control algorithms that are self-aware and self-calibrating
  - Improve energy efficiency, reducing peak demand, and improving comfort.
- ❑ This project contributes towards the BTO emerging technology goal of reducing U.S. building portfolio's carbon footprint in half by 2035
- ❑ RL-based algorithm demonstrated an average demand cost savings of 25% while maintaining the occupant comfort

Stakeholder Engagement

- Weekly meetings
- ORNL has presented the project, findings, and lessons learned to other national laboratories, professional societies, workshops, conferences, seminars, industry representatives.
- Team members are active in professional societies
- Published 14 Journal and conference papers, 1 journal papers under review, and more in process.
- ORNL submits Quarterly Progress Report(QPR) to DOE

OAK RIDGE National Laboratory

Southern Company

THE UNIVERSITY OF TENNESSEE KNOXVILLE

IEEE

ASHRAE

acm Association for Computing Machinery

Remaining Project Work

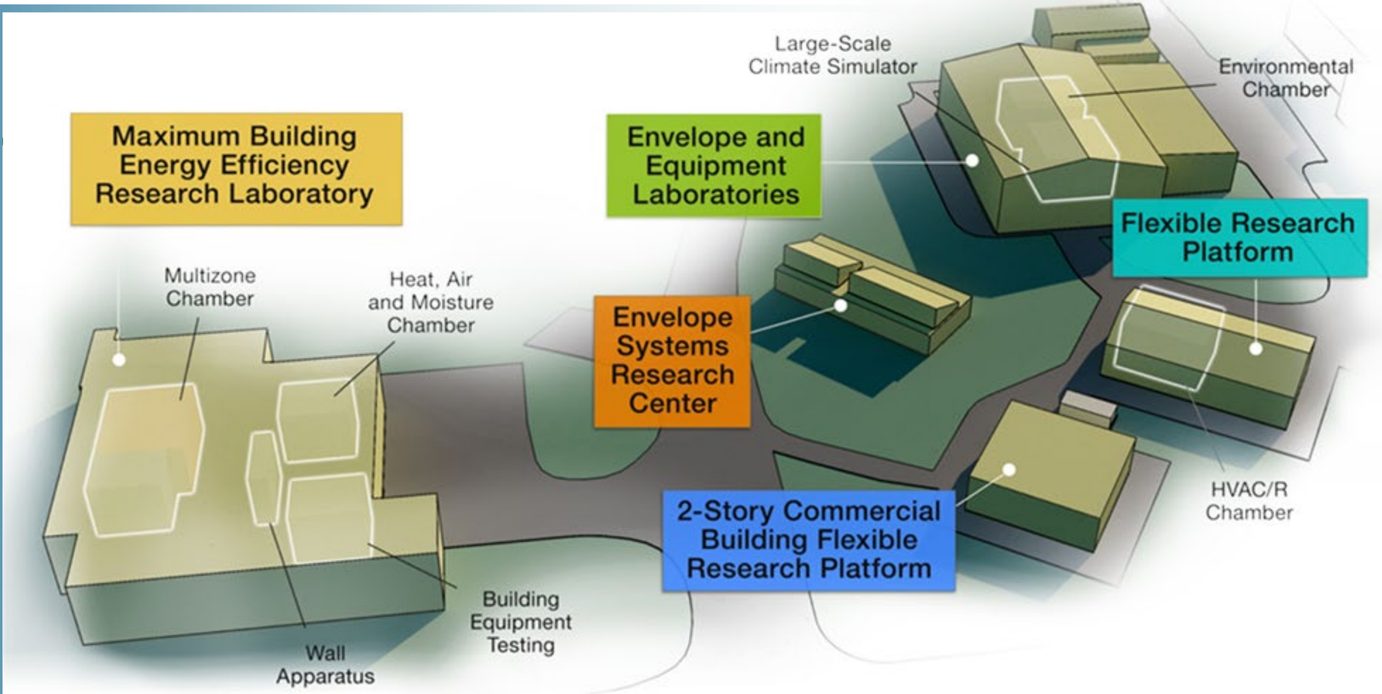
- ❑ Improve the performance of the developed algorithms for both WH and HVAC optimization
- ❑ Enhance the RL implementation with deploying RL from scratch
- ❑ Draft a technical report which contains the results of the system integration, data collection, data analysis, algorithms design, lessons learned, and field testing

Thank you

ORNL

Helia Zandi, R&D Staff

zandih@ornl.gov



ORNL's Building Technologies Research and Integration Center (BTRIC) has supported DOE BTO since 1993. BTRIC is comprised of 50,000+ ft^2 of lab facilities conducting RD&D to support the DOE mission to equitably transition America to a carbon pollution-free electricity sector by 2035 and carbon free economy by 2050.

Scientific and Economic Results

238 publications in FY20
125 industry partners
27 university partners
10 R&D 100 awards
42 active CRADAs

BTRIC is a DOE-Designated National User Facility

REFERENCE SLIDES

List of Publications

- Yan Du, Fangxing Li, Jeffrey Munk, Kuldeep Kurte, , Kadir Amasyali, Olivera Kotevska, Helia Zandi "Intelligent Multi-zone Residential HVAC Control Strategy based on Deep Reinforcement Learning" Applied Energy 2021 (*Grid (Impact Factor: 8.5)*), https://doi.org/10.1016/j.apenergy.2020.116117
- Kuldeep Kurte, Jeffrey Munk, Kadir Amasyali, Olivera Kotevska, Robert Smith, Evan McKee, Yan Du, Bori Cui, Teja Kuruganti, Helia Zandi "Evaluating the Adaptability of Reinforcement Learning based HVAC Control for Residential Houses" *MDPI Sustainability as part of the Special Issue Building and Urban Energy Prediction-Big Data Analysis and Sustainable 2020,* https://doi.org/10.3390/su12187727
- "Approximating Nash Equilibrium in Day-ahead Electricity Market Bidding with Multi-Agent Deep Reinforcement Learning", Yan Du, Fran Li, Helia Zandi, Sonny Xue, Accepted in Journal of Modern Power Systems and Clean Energy (MPCE),2021.
- Yan Du, Fangxing Li, Jeffrey Munk, Kuldeep Kurte, , Kadir Amasyali, Olivera Kotevska, Helia Zandi "Multi-task Deep Reinforcement Learning for Intelligent Multi-zone Residential HVAC Control," Electric Power Systems Research 2021, https://doi.org/10.1016/j.epsr.2020.106959
- Yan Du, Fangxing Li, Helia Zandi, Olivera Kotevska, Kuldeep Kurte, , Kadir Amasyali, Jeffrey Munk, Evan McKee "Model-Based and Data-Driven Heating, Ventilation, and Air Conditioning (HVAC) Control Strategies for Residential Demand Response (DR) Programs," under review IEEE Transactions on Industrial Informatics.
- Kadir Amasyali, Kuldeep Kurte, Jeffrey Munk, Olivera Kotevska, Robert Smith, Helia Zandi " Double Deep Q-Networks for Optimizing Electricity Cost of a Water Heater", IEEE ISGT 2021.
- Kotevska, O., Kurte, K., Munk, J., Johnston, T., McKee, E., Perumalla, K., Zandi, H. " RL-HEMS: Reinforcement Learning-based Home Energy Management System for HVAC Energy Optimization" *ASHREA Conference, 2020*.
- for HVAC Energy Control", IEEE BigData20 Industry & Government

List of Publications

- **Evan McKee, Yan Du, Fangxing Li, Jeffrey Munk, Travis Johnston, Kuldeep Kurte, Olivera Kotevska, Kadir Amasyali, Helia Zandi "Deep Reinforcement Learning for Residential HVAC Control with Consideration of Human Occupancy"** ***IEEE Power & Energy Society (PES) General.***
- **Kuldeep Kurte, Jeffrey Munk, Kadir Amasyali, Olivera Kotevska, Robert Smith, Helia Zandi "Electricity aware deep reinforcement learning based intelligent HVAC control" ACM 1st International Workshop on Reinforcement Learning for Energ,** https://doi.org/10.1145/3427773.3427866
- **y Management in Buildings & Cities**Kadir Amasyali, Kuldeep Kurte, Jeffrey Munk, Olivera Kotevska, Robert Smith, Helia Zandi "Double Deep Q-Networks for Optimizing Electricity Cost of a Water Heater" ACM 1st International Workshop on Reinforcement Learning for Energy Management in Buildings & Cities
- Xiao Kou, Mohammad Olama, Helia Zandi, Chenang Liu, Saiid Kassaee, Brennan Smith, Ahmad Abu-Heiba, Ayyoub M. Momen, "Bi-Level Optimization for Electricity Transaction in Smart Community WitYang **Xiao Kou, Fangxing Li, Jin Dong, Michael Starke, Jeffrey Munk, Yaosuo Xue, Mohammed Olama, Helia Zandi, "A Scalable and Distributed Algorithm for Managing Residential Demand Response Programs using Alternating Direction Method of Multipliers (ADMM)"** ***IEEE Transactions on Smart Grid (Impact Factor: 10.48).*** *DOI:* [*10.1109/TSG.2020.2995923*](https://doi.org/10.1109/TSG.2020.2995923)*,*
- **Xiao Kou, Fangxing Li, Jin Dong, Yang Chen, Mohammed Olama, Helia Zandi "A Comprehensive Scheduling Framework using SP-ADMM for Residential Demand Response with Weather and Consumer Uncertainties," IEEE Transactions on Power Systems 2020** ***Grid (Impact Factor: 6),*** [10.1109/TPWRS.2020.3029272](https://doi.org/10.1109/TPWRS.2020.3029272)**.** Cheh Modular Pump Hydro Storage", *ASME conference, 2020.*
- Olivera Kotevska, Jeffrey Munk, Kuldeep Kurte, Kadir Amasyali, Robert Smith, Helia Zandi "Methodology for Interpretable Reinforcement Learning Model

Project Budget

Project Budget: $700K(FY19), $700K(FY20), $700K(FY21)
Variances: N/A
Cost to Date: $1,824K
Additional Funding: None

Budget History					
FY2019– FY 2020 (past)		FY 2021 (current)		FY 2022 (planned)	
DOE	Cost-share	DOE	Cost-share	DOE	Cost-share
$1,400K	$0	$700K	$0	$0	$0

Project Plan and Schedule

Project Schedule		
Project Start: 10/2018		Completed Work
Projected End: 9/30/2021		Active Task (in progress work)
		Milestone/Deliverable (Originally Planned) use for missed
		Milestone/Deliverable (Actual) use when met on time

	FY2019				FY2020				FY2021			
Task	Q1 (Oct-Dec)	Q2 (Jan-Mar)	Q3 (Apr-Jun)	Q4 (Jul-Sep)	Q1 (Oct-Dec)	Q2 (Jan-Mar)	Q3 (Apr-Jun)	Q4 (Jul-Sep)	Q1 (Oct-Dec)	Q2 (Jan-Mar)	Q3 (Apr-Jun)	Q4 (Jul-Sep)
Past Work												
Q1 Milestone: Draft software architecture specification for the load management system												
Q2 Milestone: Develop field evaluation plan												
Q3 Milestone: Implement software application for load management system												
Q4 Milestone: Define and formulate the initial RL algorithm												
Q1 Milestone: Initial field validation at Yarnell Station												
Q2 Milestone: Try various RL strategies and algorithm improvement												
Q3 Milestone: Draft a data collection & analysis report												
Q4 Milestone: Scalability testing												
Q1 Milestone: Improve RL algoirthm performance & its cost svaing and comfort in simulation												
Q2 Milestone: Field test the improved methodology at Yarnell station house												
Q3 Milestone: Draft a data collection an danalysis report												
Current/Future Work												
Q4 Milestone: Improve RL performance and draft a technical report												