# DMD-Augmented Deep Reinforcement Learning for Fault-Tolerant Multi-Zone HVAC Control

Aman Makwana[1,2] Joseph Yamé[1] Tushar Jain[2]

[1]Université de Lorraine, CNRS, CRAN, Nancy, France
[2]Indian Institute of Technology (IIT), Mandi, India

3ème Congrès Annuel de la SAGIP
Mulhouse, 21–23 mai 2025

# Outline

# Introduction & Problem Relevance

- Heating, Ventilation, and Air Conditioning (HVAC) systems are crucial infrastructure components that account for approximately 40% of building energy and are prone to experience faults like sensor drifts and actuator degradation.
- Traditional fault-tolerant control relies on precise system models, which is challenging for large, coupled multi-zone HVAC installations.
- Recent studies indicate that combining data-driven system identification methods with reinforcement learning controller enhances control performance. [2,3,4]
- Example: Combining data-driven modeling (DMD) with deep RL has improved control tasks (e.g. stabilized vortex shedding with 8% drag reduction).
- This motivated our hybrid approach: we integrated real-time DMD-based system identification with a Deep Q-Network controller for robust, fault-tolerant multi-zone HVAC control.

# Problem Statement & Objectives

- **Problem:** Maintain desired temperature setpoints in each zone of a multi-zone HVAC system under sensor and actuator faults. Thermal coupling between zones and measurement errors (e.g. a biased temperature sensor) make control difficult.
- **Objectives:**
  - Develop a hybrid control framework that combines DMD-based system identification with a deep Q-Network (DQN) controller.
  - Detect faults online by monitoring prediction residuals between the DMD model and actual sensor readings.
  - Ensure robust temperature regulation: learn optimal control policies with the DQN, and activate a safe PID fallback mode if a significant fault is detected.

# Multi-zone HVAC System Model

- Zone thermal dynamics:

$$\frac{dT_i}{dt} = -\frac{T_i - T_{amb,i}}{\tau} + \sum_{j \neq i} c_{ij}(T_j - T_i) + \gamma_i u_i, \quad (1)$$

- where: $T_i(\mathrm{t})$ is the temperature in zone $i$ at time $t$. $T_{amb,i}$ is the ambient temperature. $\tau$ is the thermal time constant. $c_{ij}$ represents inter-zone coupling coefficients. $\gamma_i$ is the control gain, modified by actuator efficiency. $u_i(t) \in \{-1, 0, 1\}$ represents discrete control actions.

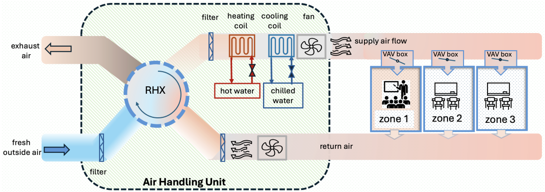


Figure 1: Schematic of the multi-zone HVAC system

# The Dynamic Mode Decomposition Module

- **Dynamic Mode Decomposition (DMD)** is a data-driven method used to analyze the dynamics of complex systems.[1]
- DMD breaks down the evolution of a dynamic system into modes that are each associated with a certain frequency and growth rate.
- It uses snapshots of the system at different time points and applies techniques like SVD to extract spatial and temporal modes.
- The basic idea is to approximate the system's behavior as a sum of eigenmodes, similar to spectral decomposition, but applied to time-series data.
- Let $x_k = \mathrm{col}(T_1(t_k), T_2(t_k), T_3(t_k))$ where $t_k = kh$, $h$ being the sampling period. So here, what the DMD actually does is to approximate the Leading eigen decomposition of the best fit linear operator (the $\Omega$ matrix) that advances $x_k$ to $x_{k+1}$ $[x_{k+1} \approx \Omega x_k]$.
- By decomposing the data into Dominant Spatial Coherent Modes, we can more easily understand the dominant dynamics of the system and predict future behavior.

# The Dynamic Mode Decomposition Module

- We define the DMD residual as:

$$r = \|\Omega x_k - x_{k+1}\|_2. \tag{2}$$

- Under the nominal conditions, the residual $r$ remains small and nearly constant. A sudden change around the mean value of $r$ could indicate a system fault. If that change is significant and exceeds a predefined threshold then a PID fall-back mechanism is triggered that takes over the entire control of the system.

# The Deep Q-Network FTC Controller

- The FTC controller is a Deep Q-Network(DQN) controller [6] that learns the control policies based on the state estimations from DMD. The state vector of DQN includes the zone temperatures, and the action space includes all possible joint actions. The training follows standard DQN procedures with $\epsilon$-greedy exploration and experience replay. The reward function $R_t$ is designed to penalize deviations from the desired setpoint and energy consumption:

$$R_t = -\left( \sum_{k=1}^{N} \beta_1 \|x_k - T_{sp}\|_1 + \sum_{k=1}^{N} \beta_2 \|a_k - a_{k-1}\|_1 \right) \tag{3}$$

  where : $T_{sp}$ is the set point vector of the three zones and the action at time $k$ is defined as $a_k = \mathrm{col}(u_1(t_k), u_2(t_k), u_3(t_k))$ , $\beta_1$ is the weight for the temperature deviation penalty, and $\beta_2$ is the weight for the energy consumption penalty.

- As a low-dimensional system representation given by DMD is used by the DQN controller to learn the control policies, a significant reduction is observed in the computational complexity and training time, while the dominant dynamics of the original system model are still retained.

# Simulated Experimental Setup

- The entire framework was implemented in Python using PyTorch. Simulations were conducted over 120 episodes, with faults (a $+2^{\circ}$C sensor bias and 60% actuator degradation) introduced at Episode 70. Metrics show rapid training performance and stable DMD residuals. After fault introduction, the DQN module adapts to the sudden dynamical changes, resulting in zero PID fallbacks, thus demonstrating the FTC performance.

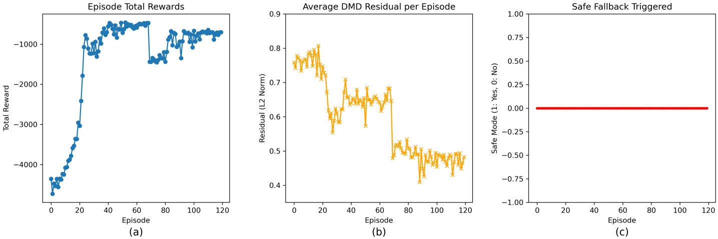


Figure 2: Performance metrics across training episodes. (a) Total reward; (b) Average DMD residual; (c) Safe Fallback Triggered.

# Results and Conclusion

- Results demonstrate pre (Ep 66) and post fault injection (Ep 101) stable control with true temperature trajectories converging to the set point (i.e. 11°C) and sensor reading plot showing robust control even after +2°C sensor bias.

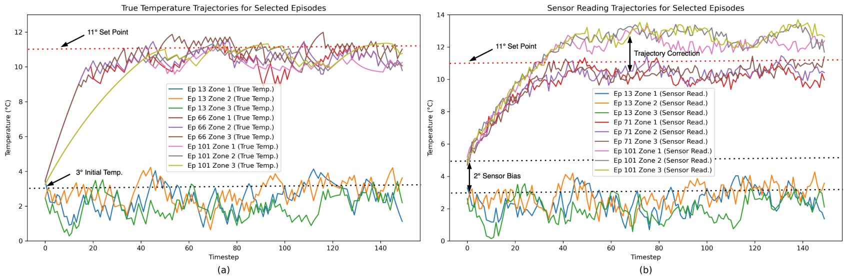


Figure 3: Temperature trajectories for selected episodes. (a) True Zone Temperature; (b) Noisy Sensor Readings.

# Results and Conclusion

- **Conclusion:** We presented a hybrid DMD-Augmented-DRL framework for Fault-Tolerant Control of Multi-Zone HVAC systems.The DMD module extracts a low-dimensional representation of system dynamics from true temperature trajectories, enabling prompt fault detection via prediction residuals. The DQN then utilizes these enhanced state estimates to learn optimal control policies, with a PID fallback activated when residuals exceed a threshold. This approach reduces computational complexity while ensuring robust control under sensor and actuator faults.

- **Acknowledgment:** The authors would like to thank the ORION program for its contribution to the funding of A. Makwana's research internship. This work has benefited from a government grant managed by the Agence Nationale de la Recherche with the reference ANR-20-SFRI-0009.

# References


[1] P. J. Schmid (2010). Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656 : 5–28.

[2] V. S. Donge, B. Lian, F. L. Lewis, & A. Davoudi (2023). Accelerated reinforcement learning via dynamic mode decomposition. *IEEE Transactions on Control of Network Systems*, 10(4) : 2022–2034.

[3] T. Sakamoto, & K. Okabayashi(2024). Optimization of fluid control laws through deep reinforcement learning using dynamic mode decomposition as the environment. *AIP Advances*, 14(11).

[4] Q. Sang, Y. Tian, Q. Jin, & J. Yu (2021). A path planning strategy for marine vehicles based on deep reinforcement learning and data-driven dynamic flow fields prediction. *"In 6th International Conference. on Automation, Control and Robotics Engineering (CACRE)"*, (pp. 466–471).

[5] O. Ait-Essi, J. Yamé, H. Jamouli, & F. Hamelin (2024). Q-Learning Based LQR Occupant-Centric Control of Non-Residential Buildings. *"In 21st International Conference on Informatics in Control, Automation and Robotics (ICINCO)"*.

[6] R. S. Sutton and A. G. Barto (2018). *Reinforcement Learning: An Introduction (2nd ed.)*, MIT Press.

[7] T. Jain, J. Yamé, & D. Sauter(2018). *Active Fault-Tolerant Control Systems*, Cham, Switzerland: Springer.

# Thank You!

*Questions or Suggestions?*

**Aman Makwana**
M.Tech Research, IIT Mandi
Orion Research Intern, Université de Lorraine
aman.work.cs@gmail.com