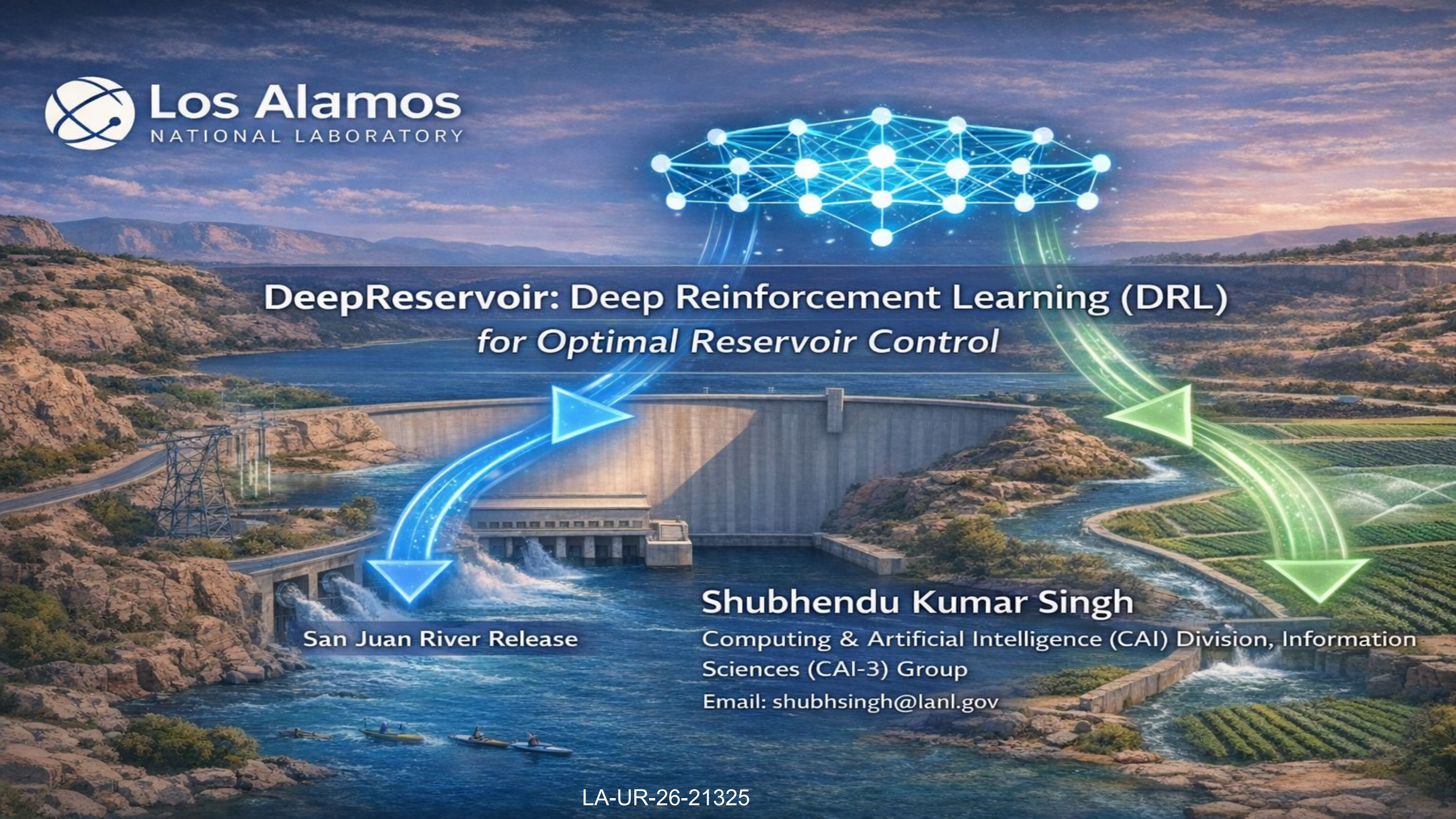
Los Alamos
NATIONAL LABORATORY

DeepReservoir: Deep Reinforcement Learning (DRL) *for Optimal Reservoir Control*



Shubhendu Kumar Singh
Computing & Artificial Intelligence (CAI) Division, Information Sciences (CAI-3) Group
Email: shubhsingh@lanl.gov

LA-UR-26-21325

Objective

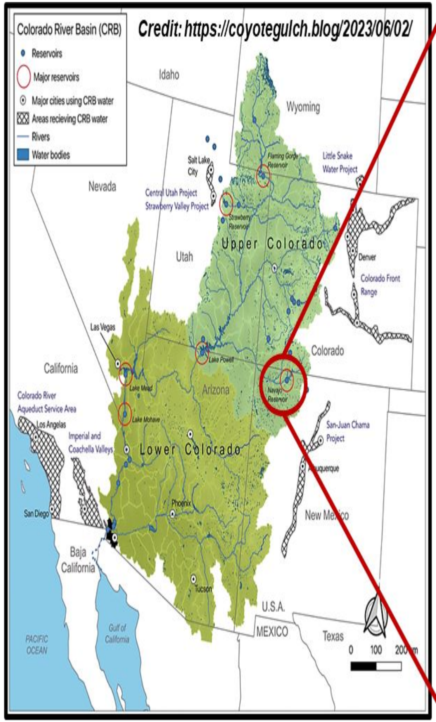
💧 **Critical Regional Asset:** Supplies water, irrigation (San Juan Basin), flood control, hydropower, and recreation across NM & CO.

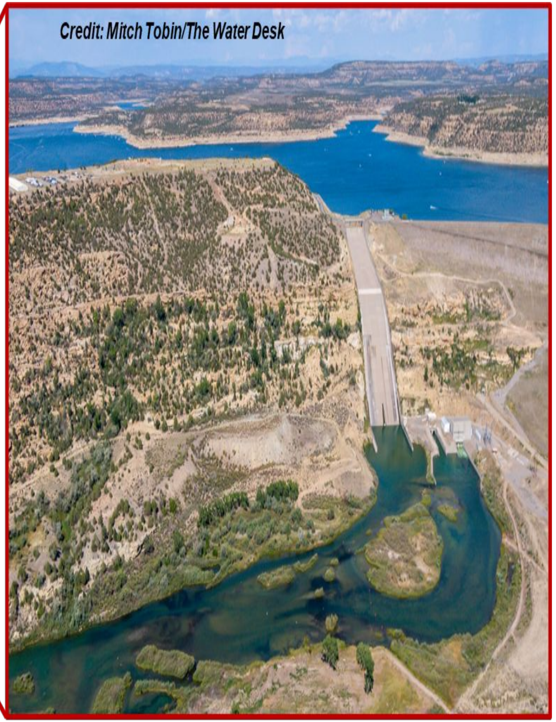
🔧 **Manual Rule-Based Decisions:** Operators follow pre-defined rule curves, which lack flexibility during extreme events.

🌦️ **Uncertainty in Hydrologic Conditions:** Inflow is influenced by snowmelt, rainfall, and upstream conditions — making precise planning difficult.

⚖️ **Balancing Conflicting Demands:** Must simultaneously manage flood control, water supply, ecology, and recreation — often with trade-offs.

📉 **Risk of Sub-Optimal Operations:** Static or reactive decisions can lead to overflow, low storage, or failure to meet downstream needs.





.

Is reinforcement learning a viable tool for delivering realistic, data-guided, optimal solutions and exploring alternative management strategies?

Why Reinforcement Learning??

🧠 **RL enables adaptive control**: Learns optimal release strategies through interaction, improving resilience and efficiency.

🌐 **Scalable framework**: RL can generalize to similar reservoir systems under changing climate and hydrologic conditions.

✅ **Proven effective:** RL has been crucial to a number of cutting-edge technologies

Deep Reinforcement Learning (DRL) for Reservoir Control

🌍 **Virtual Environment**

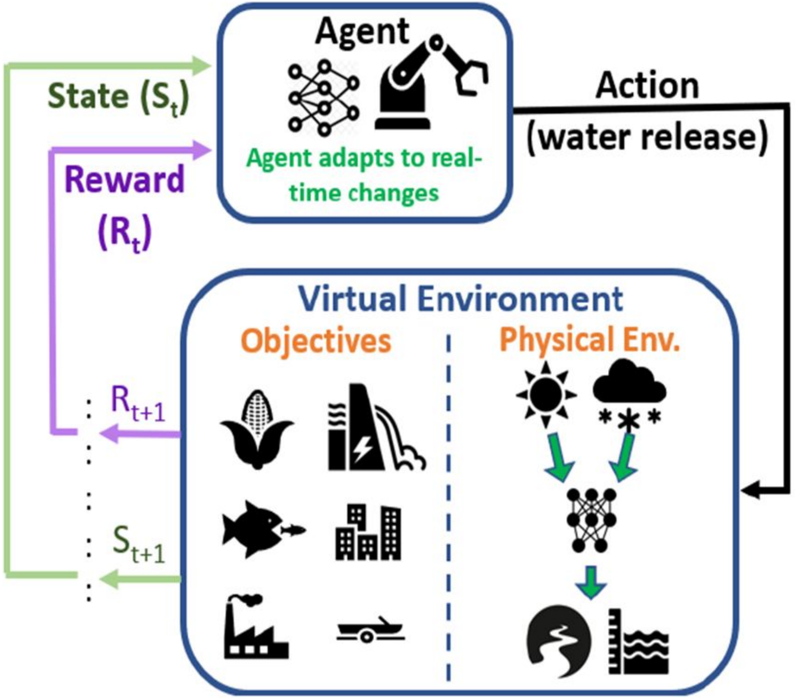
- Simulates water release management for Navajo Reservoir
- Captures reservoir dynamics: inflow, storage, evaporation, and releases
- Includes the management objectives

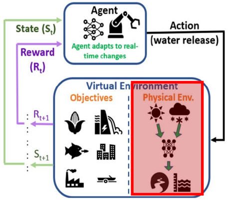
📊 **State (Observation Space)**

- **Storage**: Current volume of water in the reservoir
- **Evaporation**: Estimated water loss due to evaporation
- **Inflow**: Daily inflow into the reservoir

🤖 **Agent (Actor - PPO)**

- **Action**: Determines daily water release based on observed state
- **Objective**: Maximize efficiency while satisfying operational constraints
- **Algorithm**: *Proximal Policy Optimization (PPO)* for stable learning



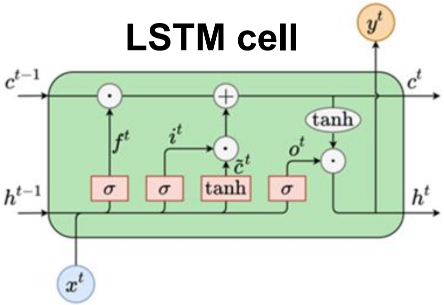


Physical Environment- Streamflow Model

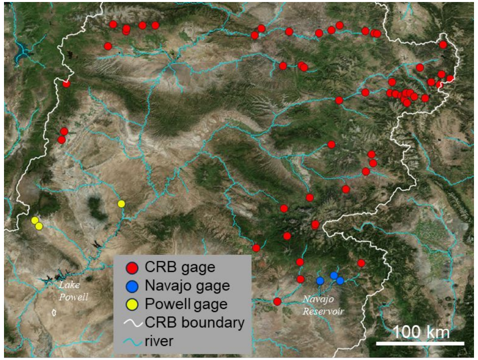
- **LSTM** model to estimate streamflow was trained with the following data:
  - *precipitation, wind speed, max temp., min temp.*
  - *snow water equivalent, soil moisture*

- **LSTM** models outperform process-based models in modeling streamflow in ungauged basins utilizing remote sensed data

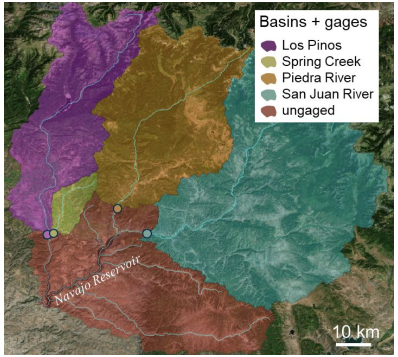
LSTM cell



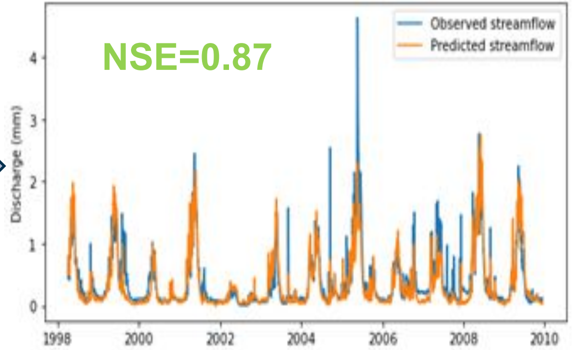
Training

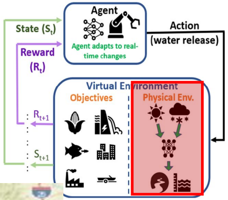


Fine Tuning



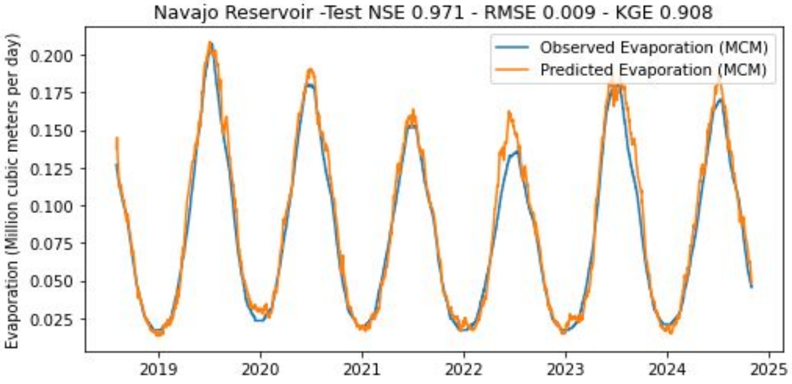
Testing- Reservoir Inflow

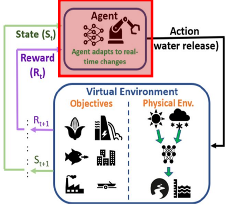




Physical Environment- Evaporation Model

- LSTM model to estimate evaporation was trained using 11 nearby reservoirs with the following data:
  - In-situ measurements – *storage volume, elevation, total release, input flow*
  - ERA5 meteorological inputs – *median temperature, mean wind speed, total precipitation*

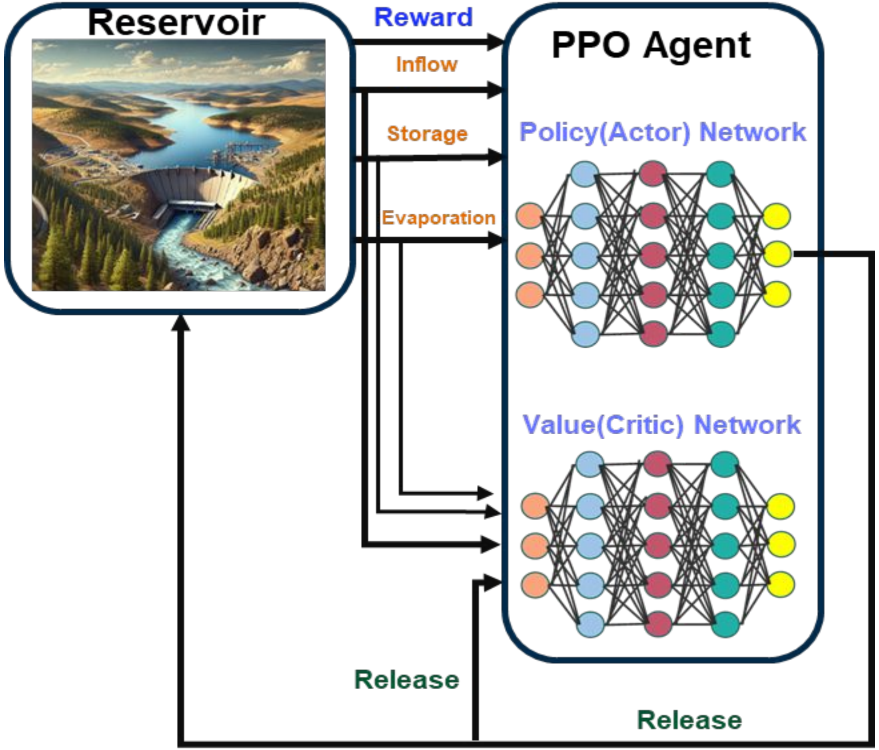




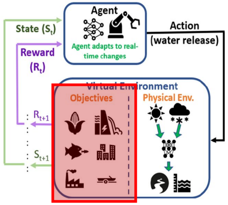
Agent - Proximal Policy Optimization (PPO)

💧 Why PPO for Navajo Reservoir?

- Handles **continuous action space**: realistic release decisions
- Takes all relevant inputs (inflow, storage, evaporation, forecasts) to provide adaptive and **data-driven control**
- Learns **anticipatory release patterns**, balancing short-term demands and long-term reservoir health



Reward Functions to Achieve Objectives

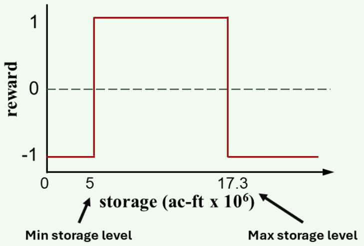


How to best-design reward functions in an RL context?

Objective:
Release water while maintain the storage with min and max storage limits

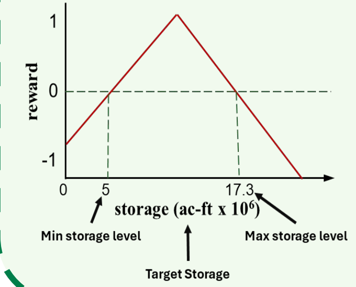
simplest approach

Reward Function:
+1 if the storage within the desired range else -1



more reward for being near mid-storage

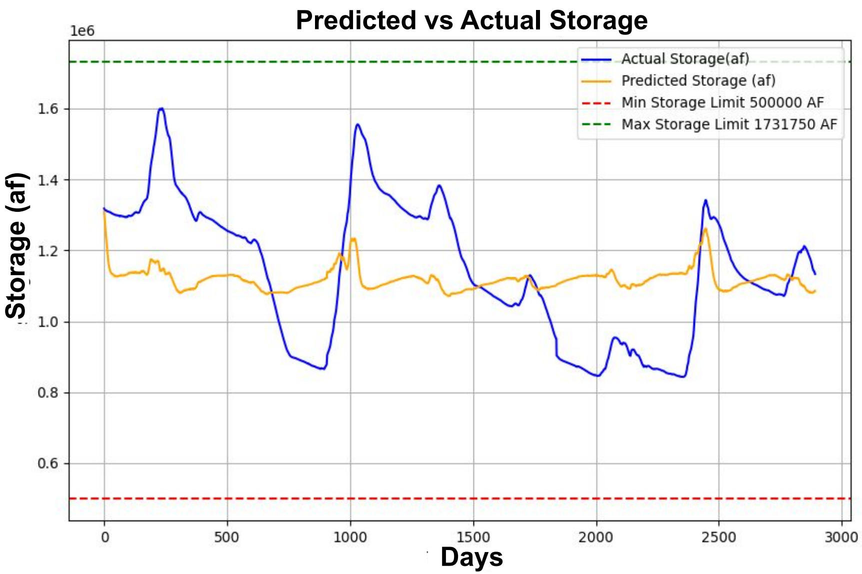
Reward Function:
+1 if the storage equal to target else less than 1

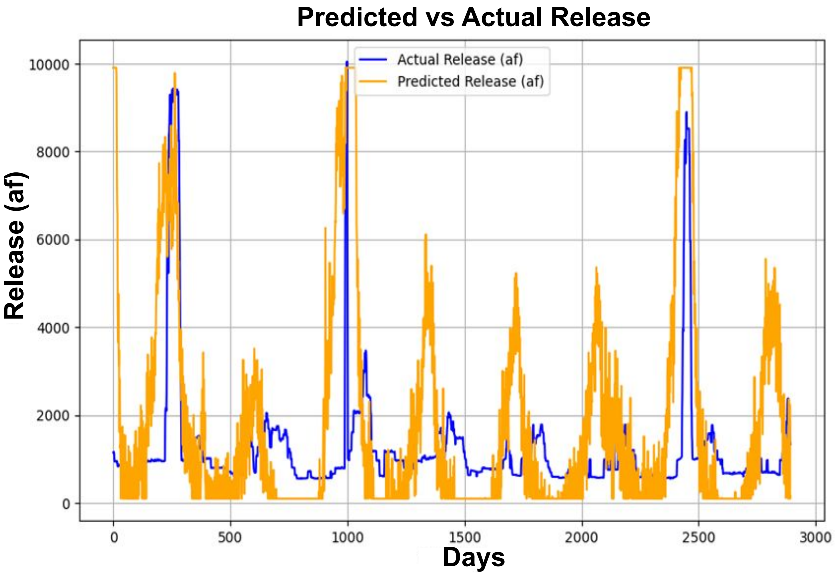


A Simple Objective

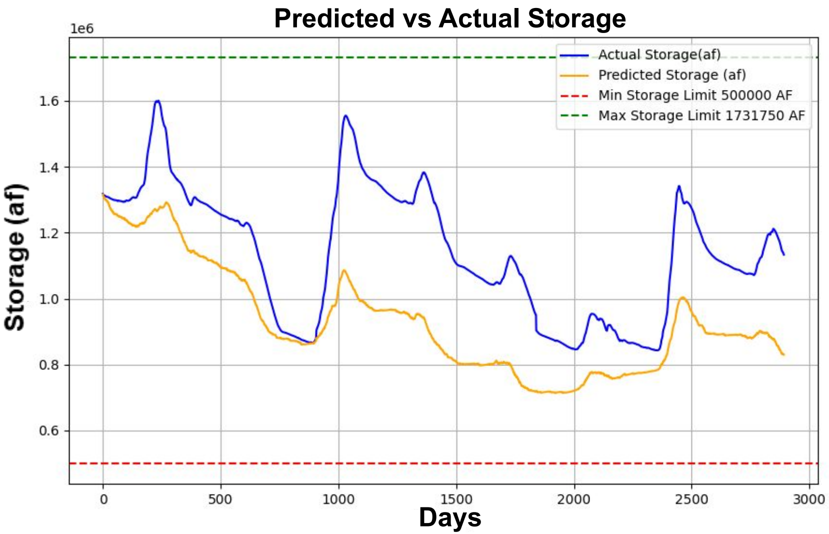
Reward :

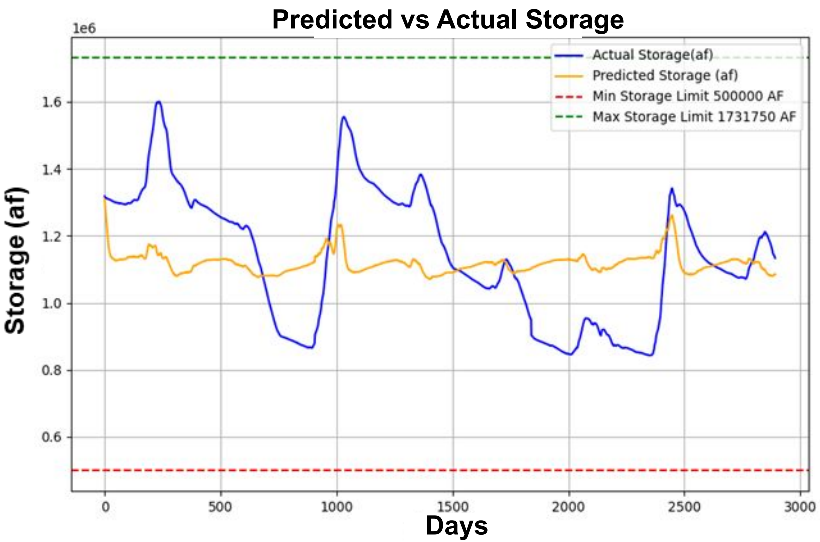
- Min-max limits with maximum reward for the midpoint of the limits



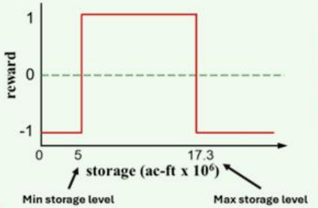


Effect of Reward Function Design

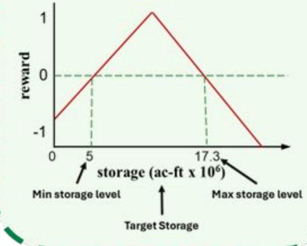




Reward Function:
+1 if the storage within the desired range else -1



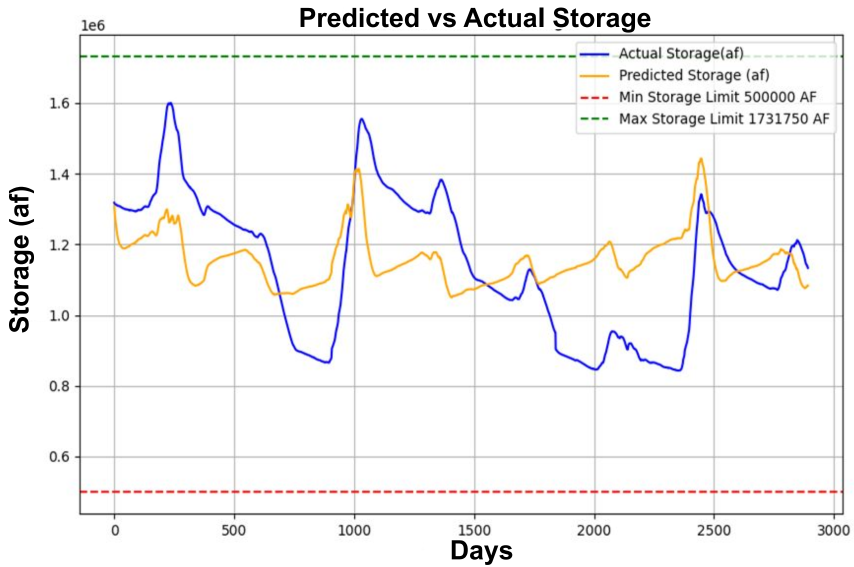
Reward Function:
+1 if the storage equal to target else less than 1

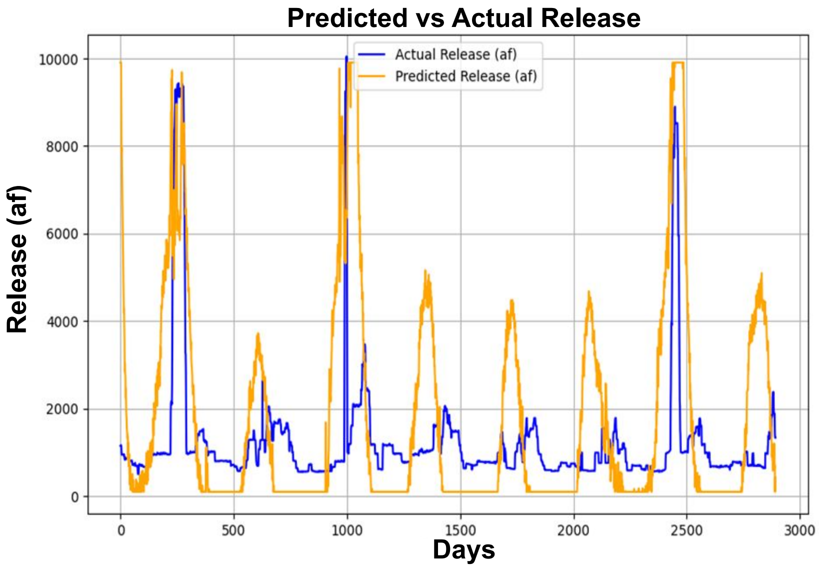


Enforcing Smoothness

Reward :

- Min-max limits with maximum reward for the midpoint of the limits
- Penalty for abrupt release

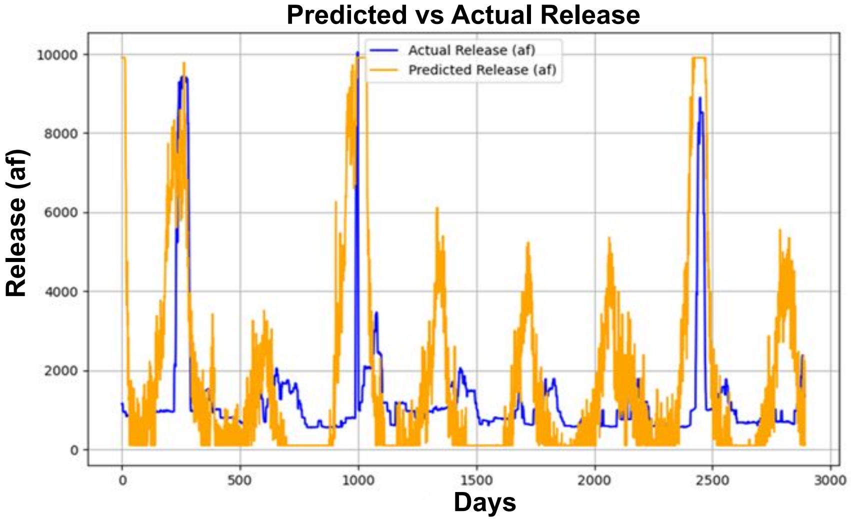




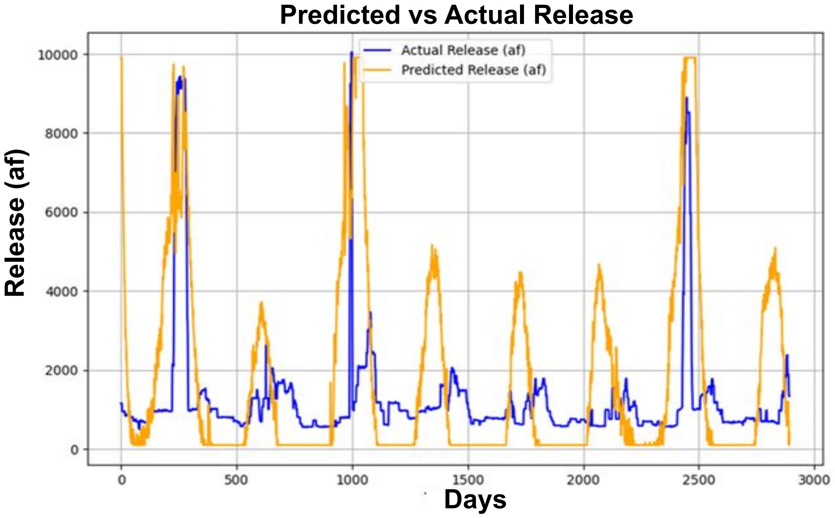
Quantifying Smoothness

$$Smoothness\ Metric = 1 - \frac{\sum_{k=kc}^{N-1} P_k}{\sum_{k=0}^{N-1} P_k}$$

Where:
P_k represents the energy at each frequency component
K_c is the cut-off frequency index
S ε [0.1], where higher S indicates smoother release



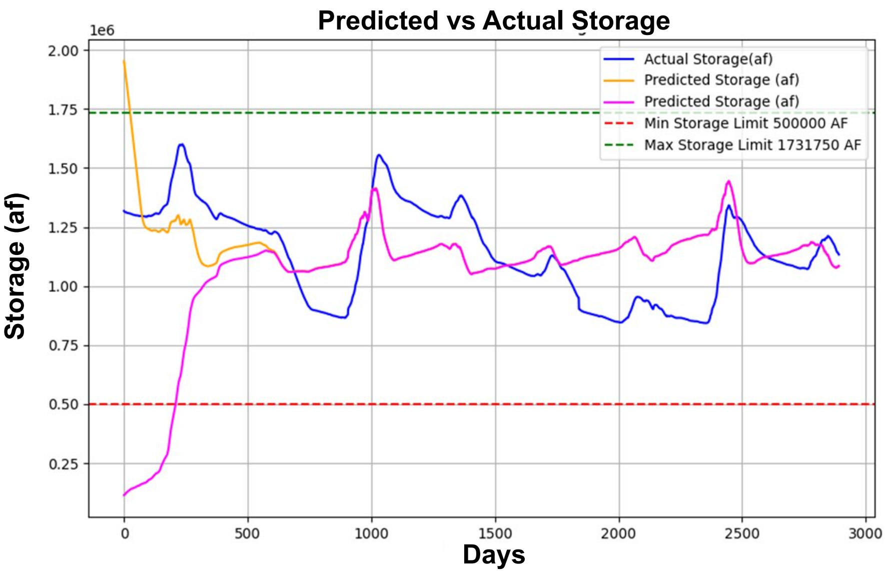
Smoothness Metric : 0.5308

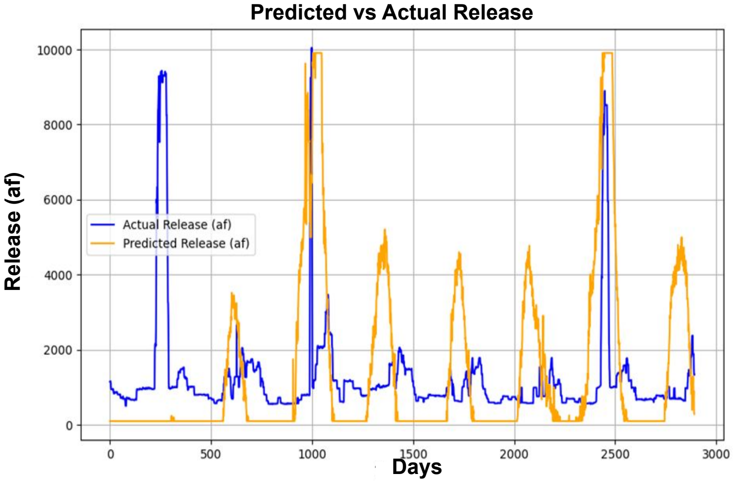


Smoothness Metric : 0.7762

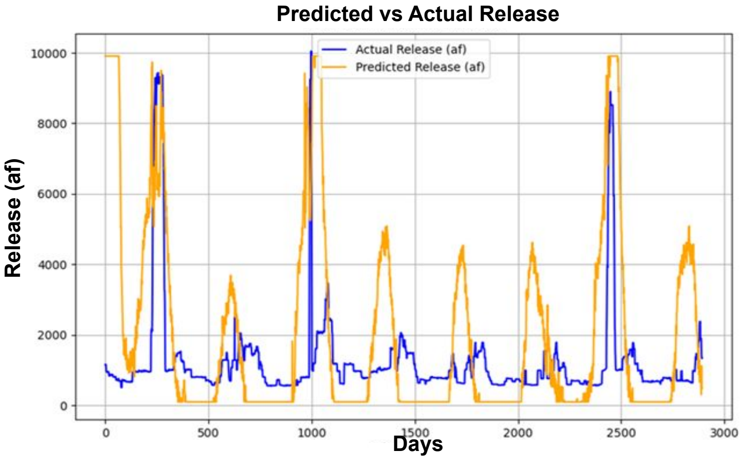
Stress Testing

Test 1: Initial storage 50 percent below/above the minimum/maximum storage limit





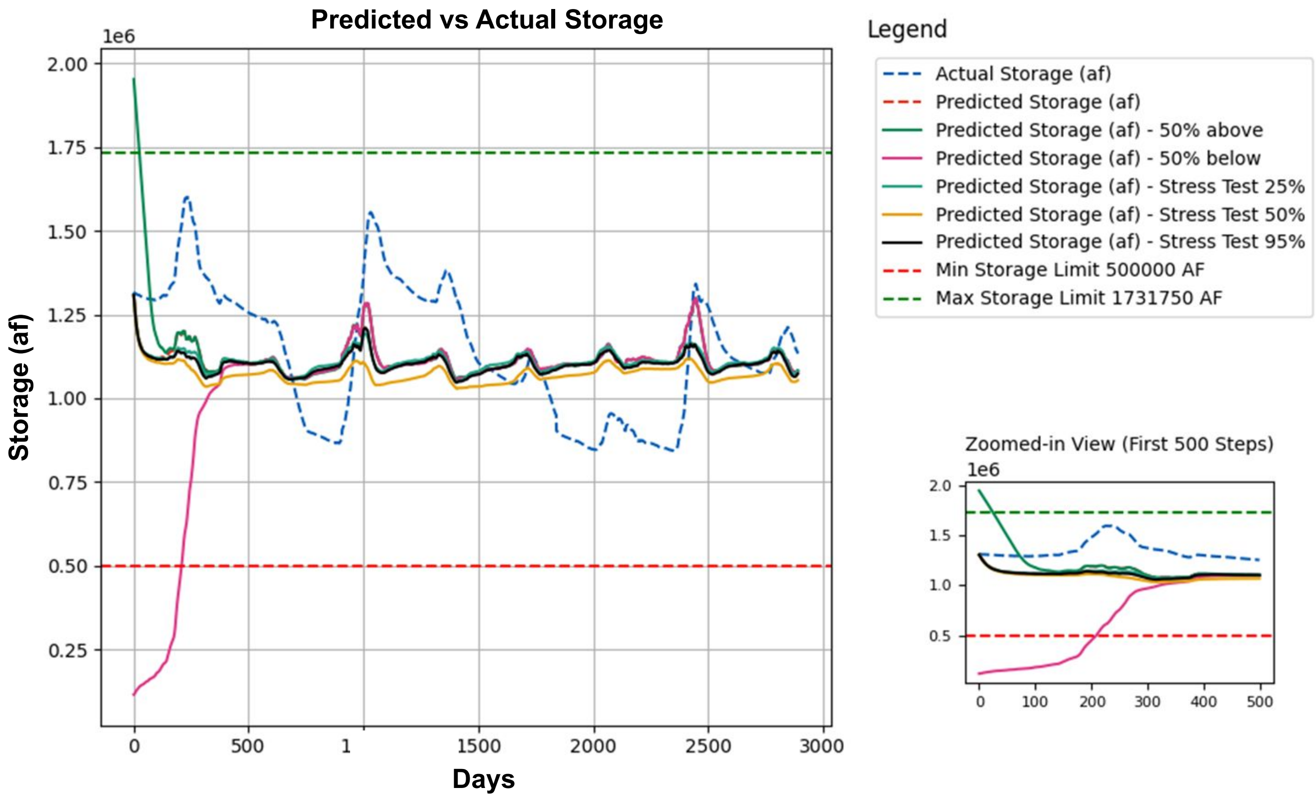
Release Pattern- 50% Below Min Storage

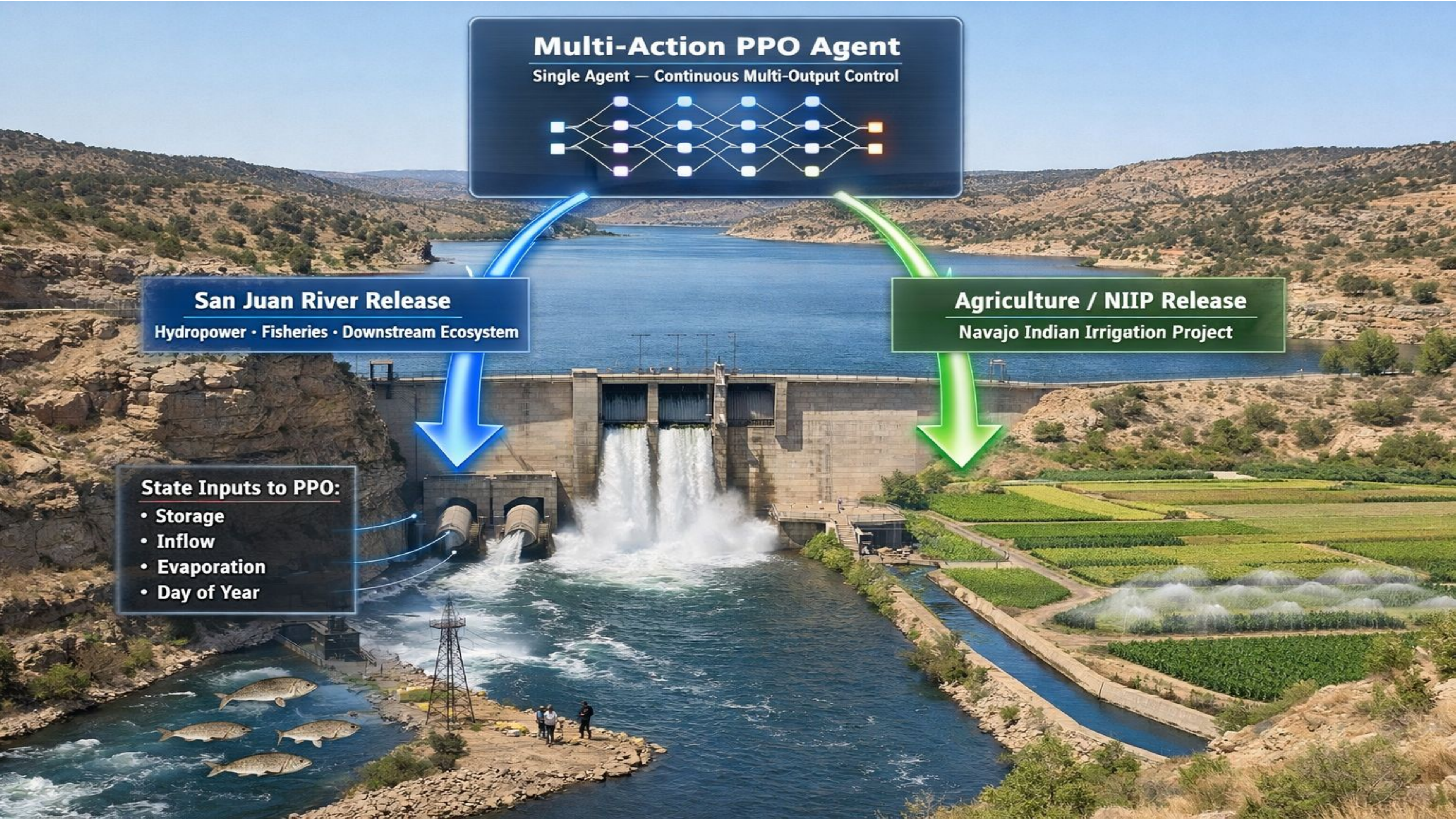


Release Pattern- 50% Above Max Storage

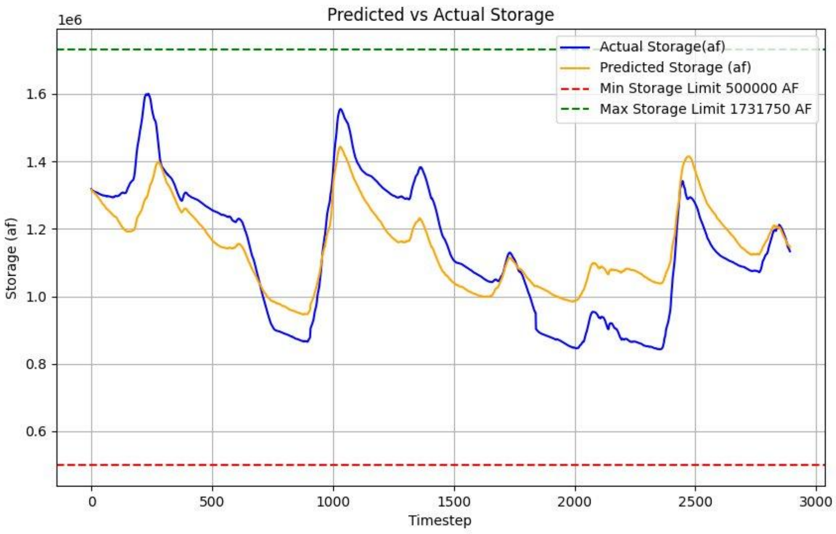
Stress Testing

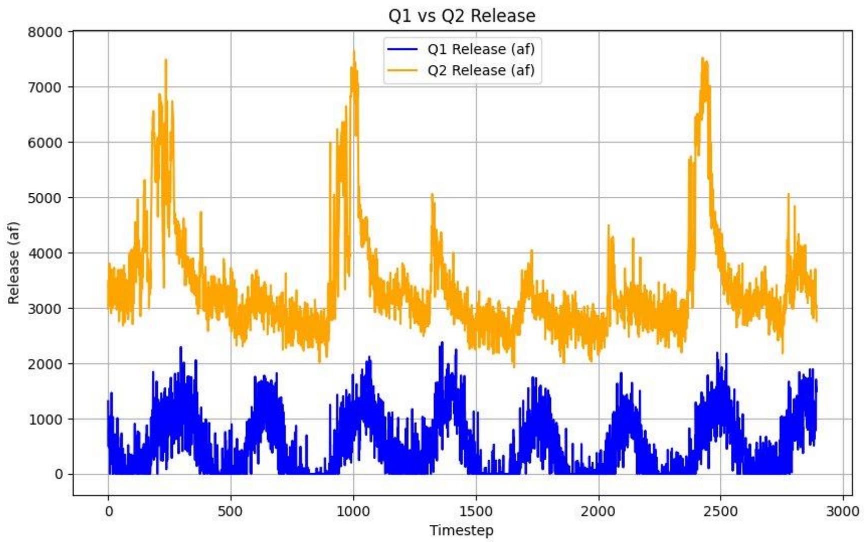
Test 2: Scaling inflows by 25%, 50%, and 90% (less available water into system)



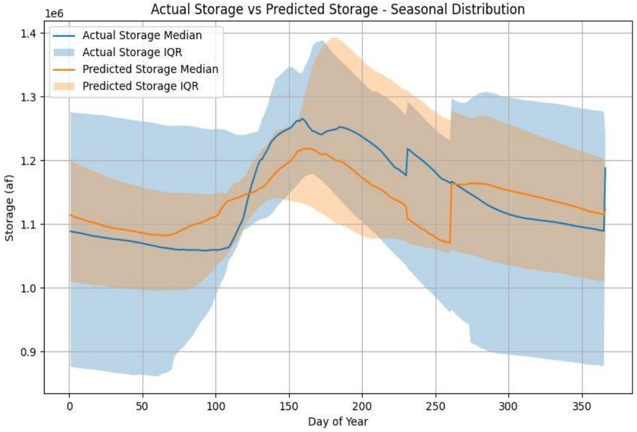

Multi-Action PPO Agent
Single Agent — Continuous Multi-Output Control
San Juan River Release
Hydropower • Fisheries • Downstream Ecosystem
Agriculture / NIIP Release
Navajo Indian Irrigation Project
State Inputs to PPO:
• Storage
• Inflow
• Evaporation
• Day of Year

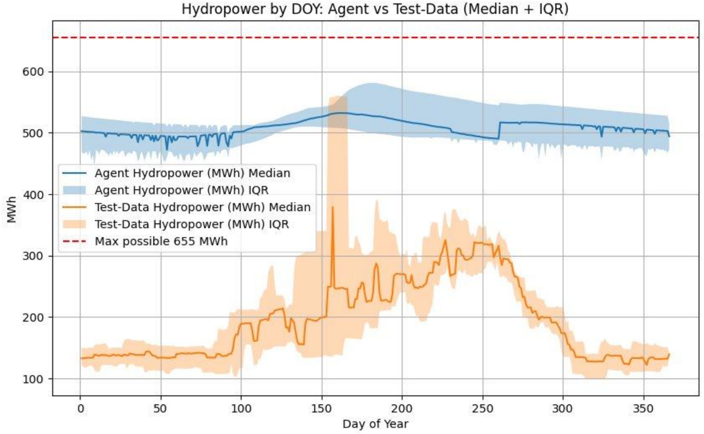
PPO Test Data Performance – Multi-Objective Reservoir Control

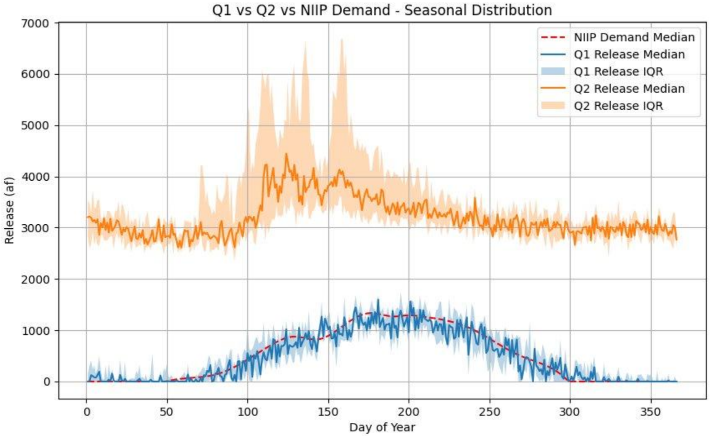




PPO Test Data Performance – Multi-Objective Reservoir Control



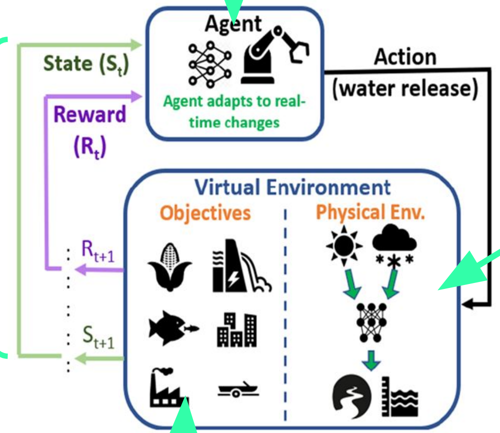




Conclusions & Future Work

PPO Agent able to comply with basic constraints; flexible to multi-objectives, and resilient against out-of-sample stress tests

The framework shows promise, providing smooth and stable storage outcomes with flexibility to expand



Validated LSTM models for inflow and evaporation, allowing experimentation with alternative weather scenarios

Objectives' functional forms impact control in intuitive ways

⏳ Next Steps

- Using forecasts to guide decision making
- Alternative weather scenarios (extreme drought and flooding)
- More complicated objective sets:
  - Honoring Endangered Species Act (release restrictions)
  - Hydropower generation
  - Water supply (municipal, agricultural, and water treaties.
  - Recreational constraints

Our Team

Jon Schwenk
(jschwenk@lanl.gov)

Cristina Garcia Cardona
(cgarciac@lanl.gov)

Rajiv Ranasinghe
(ranasinghe@lanl.gov)

Shubhendu Singh
(shubhsingh@lanl.gov)

Katrina E. Bennett
(kbennett@lanl.gov)

Susan Behery
(sbehery@usbr.gov)

Thank You!


INFLOW
INFLOW
INFLOW
EVERLAIN
EVARATION
STORAGE
STORAGE
STORAGE
RELEASE
CFS