American Modelica Conference 2022

Reinforcement Learning based Control of Integrated Energy Systems Using OpenAI Gym

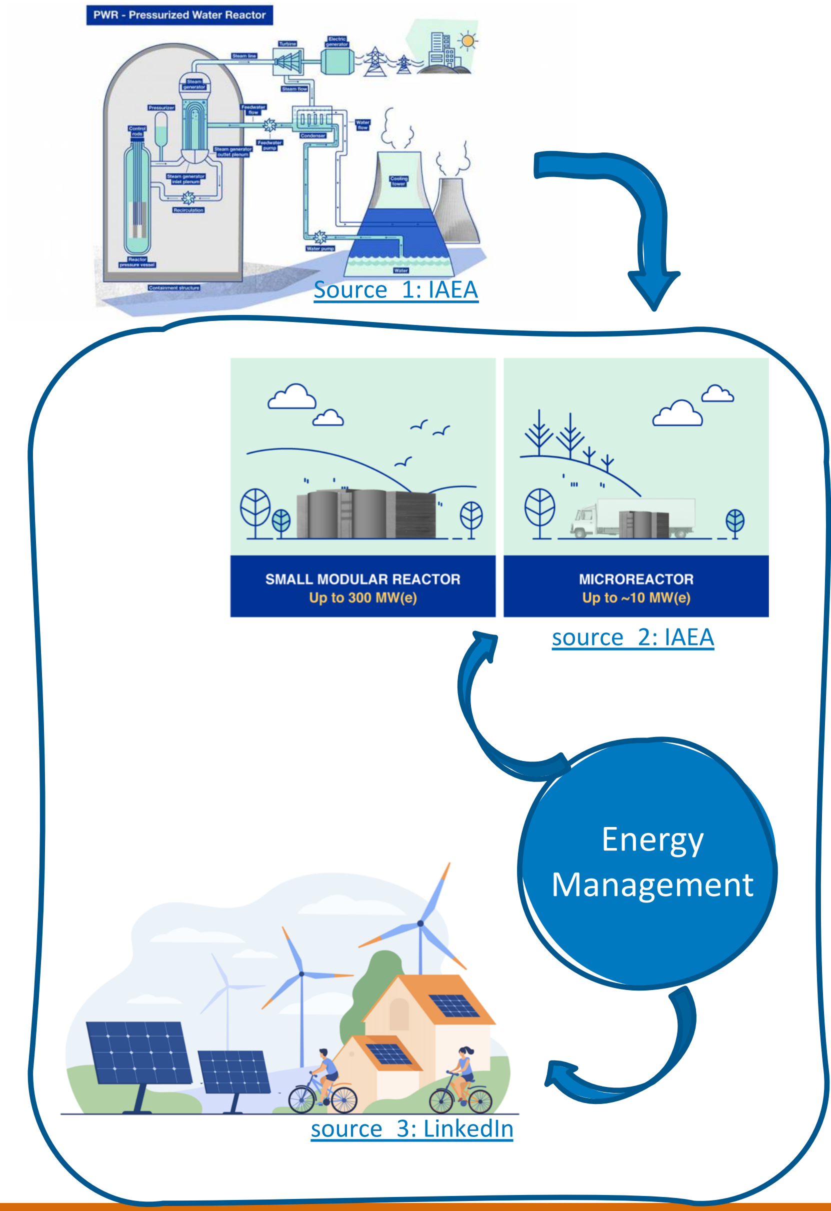
Authors: Jingbo Wang, Roshni Anna Jacob, and Jie Zhang

The University of Texas at Dallas

Presenter: Jingbo Wang

Background Introduction

- Nuclear power plants have traditionally served as centralized plants for base supply.
- The development of **advanced small modular reactors (SMRs)** in recent years has brought nuclear utilization for microgrids closer, together with other forms of green energy, such as solar and wind.
- Given the interaction of multi-domain systems, the Modelica language is an appropriate tool for developing multi-physics models.
- For this study, we used open-source packages such as **HYBRID (by INL)** and a **TRANSFORM** (by ORNL) library to create the Integrated Energy System (IES).
- What is the suitable way to control such a complex system, artificial intelligence or a traditional controller like PID, or both?



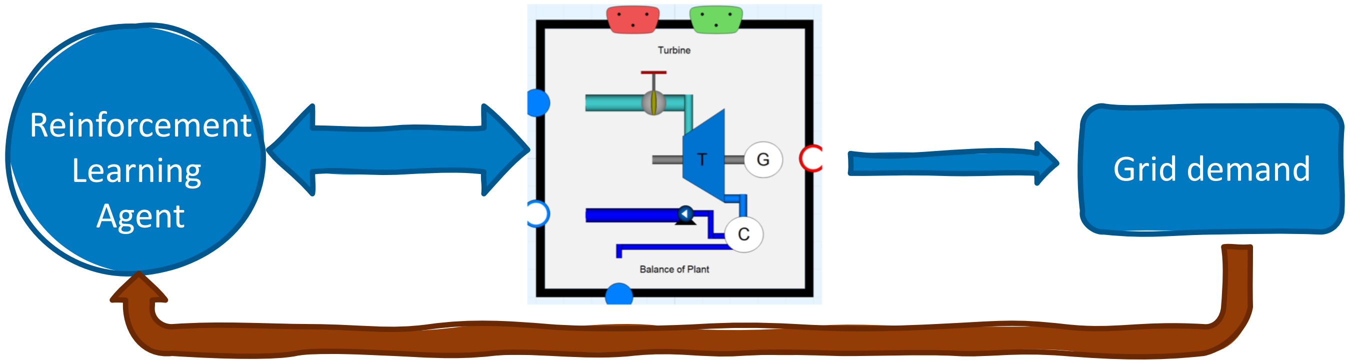
Motivation and Objective

Motivation:

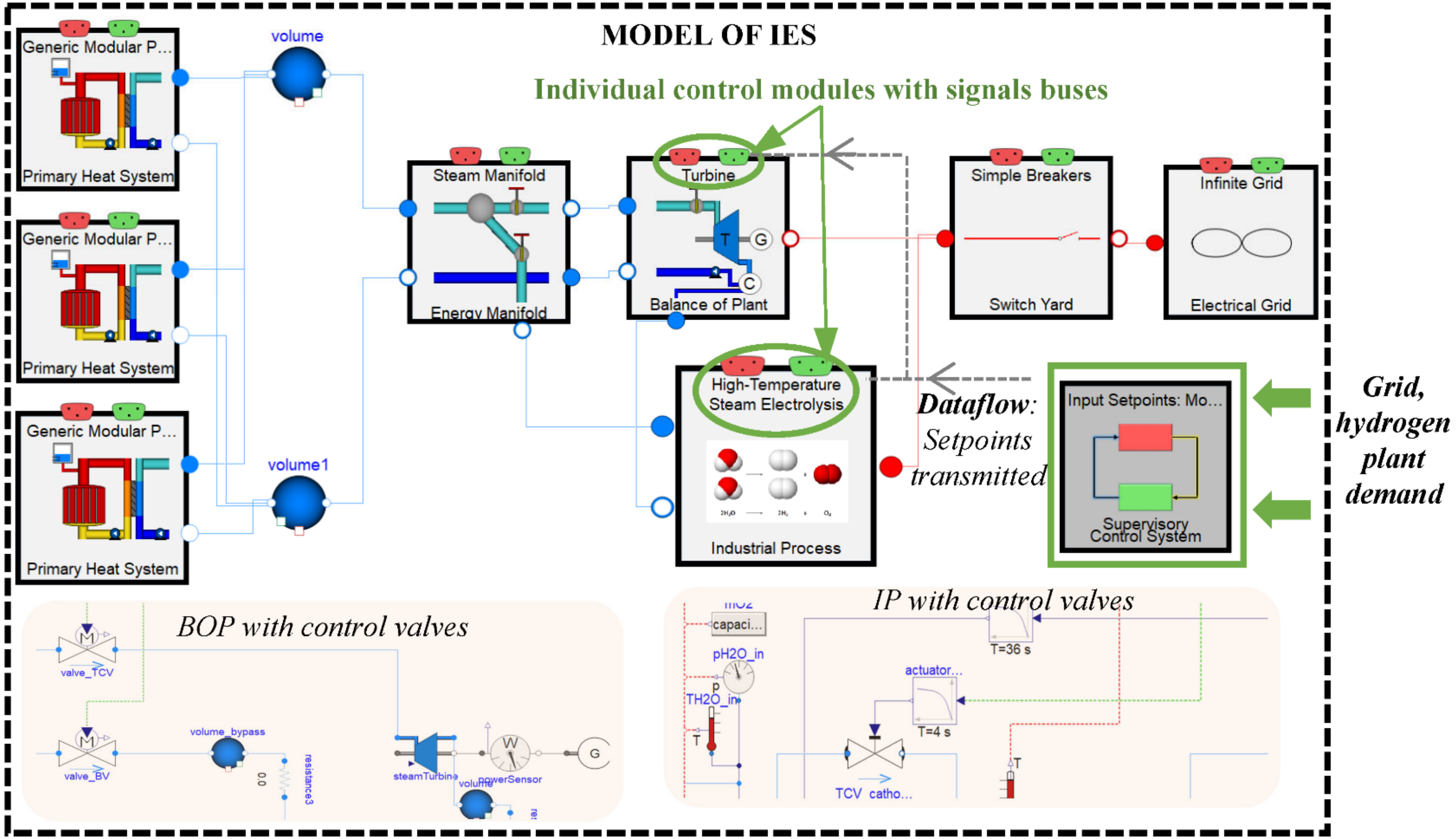
- When a variety of renewable energy systems combine to form a complex integrated energy system (IES), the benefit of reinforcement learning could be leveraged to control the IES with associated uncertainties.

Objective:

- Develop an artificial intelligence module (**reinforcement learning** agent) in conjunction with a physics-based model.
- Control the valves **to meet electricity generation and pressure balancing.**



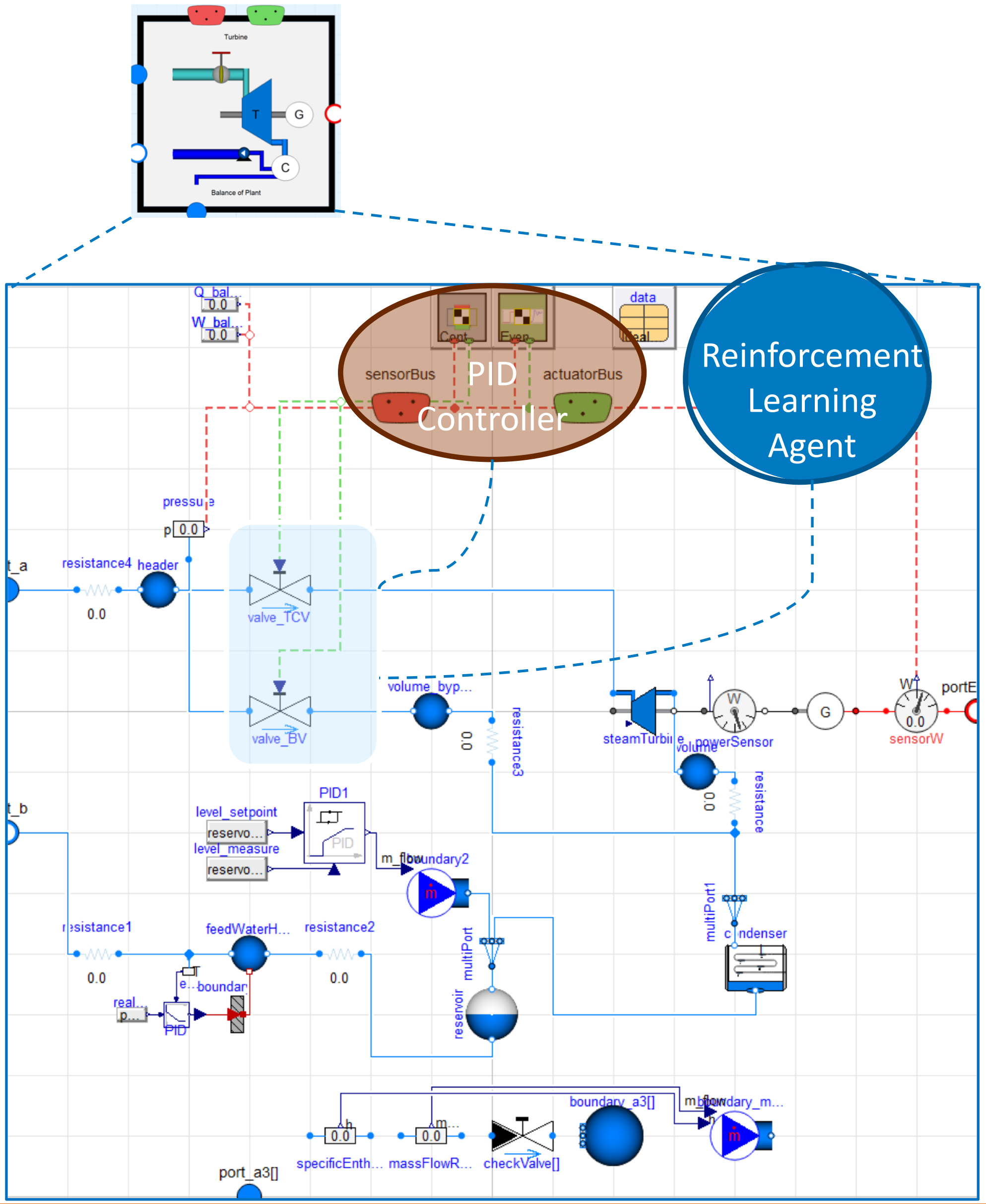
Environment: Modelica based Integrated Energy Systems



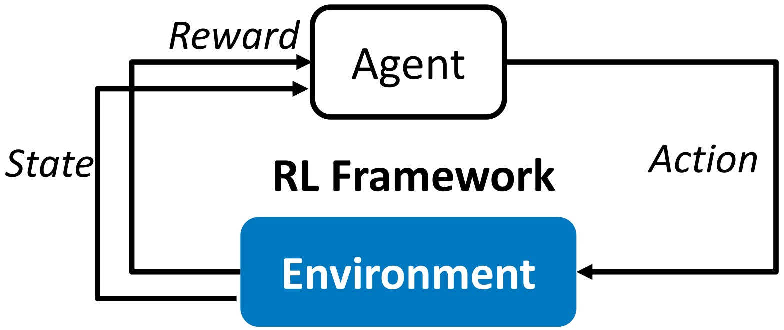
- Conventional PID controllers are used in these models.
- The RL control mechanism is autonomous, and can be deployed online for real-time control of the IES.

Experiment on the Balance of Plant

- **Model environment:** Dymola 2022
- **Model to invest with RL algorithm:** Balance of Plant (From the HYBRID library developed by Idaho National Lab)
- **Model Details:**
  - An ideal turbine model, a condenser a feedwater system for reheating.
  - Several **valves** that allow steam to flow to the turbine or as a bypass to the condenser.
    - Original control method: **PID controller** to meet electricity generation and pressure balancing.
    - Investigated Method: **RL algorithm** to control valves
  - From a modeling standpoint, test the viability of co-simulation with an artificial agent and a complex physics model, as well as comprehend the effort of training the agent with such a system.
  - From the control point of view, to compare the PID performance and intelligent agent.



OpenAI and Modelica Environment Interface

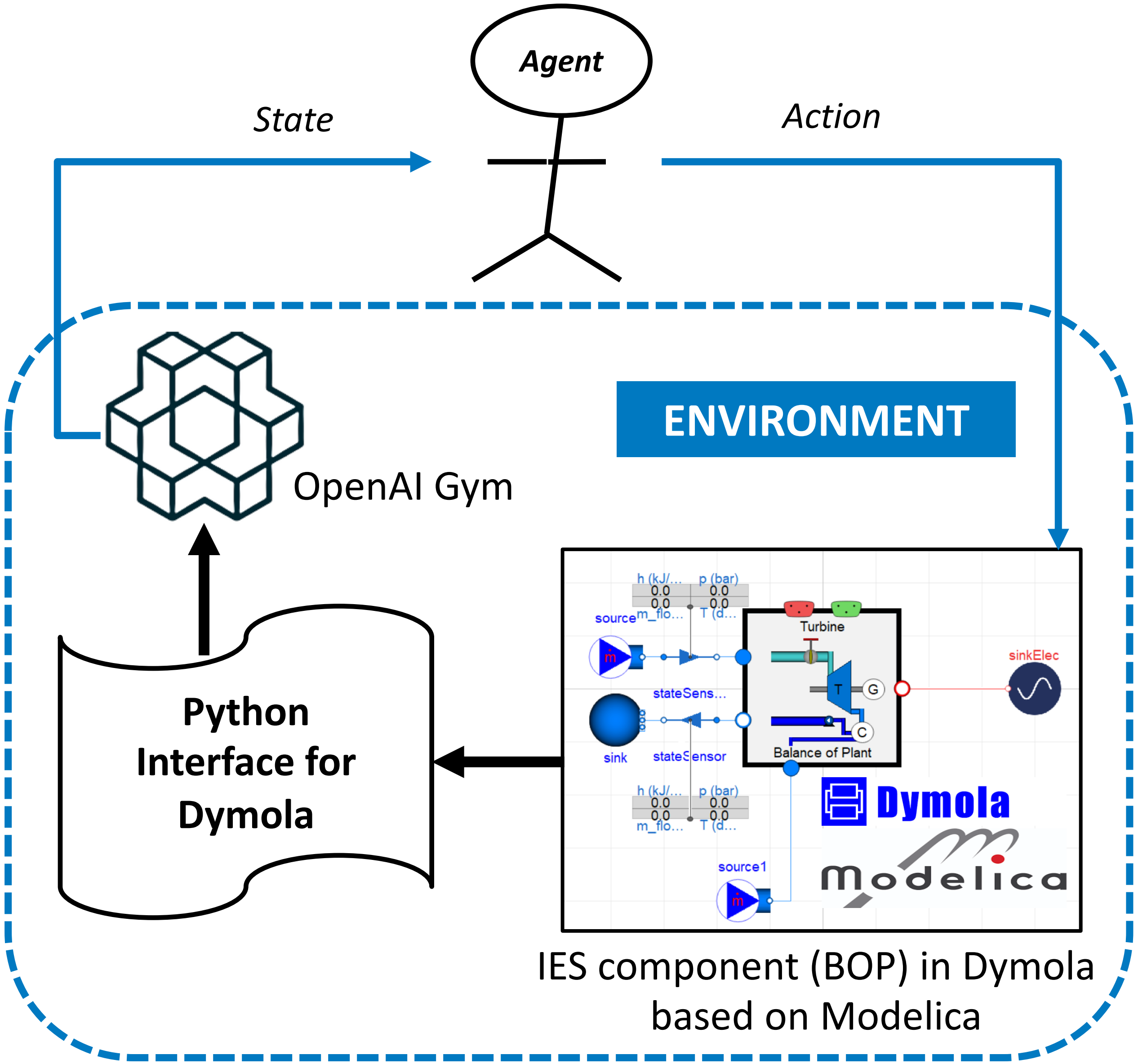


<u>State</u>:
Power produced by the BOP.
$s = \frac{P_{BOP}}{P_{BOP}^{max}}$

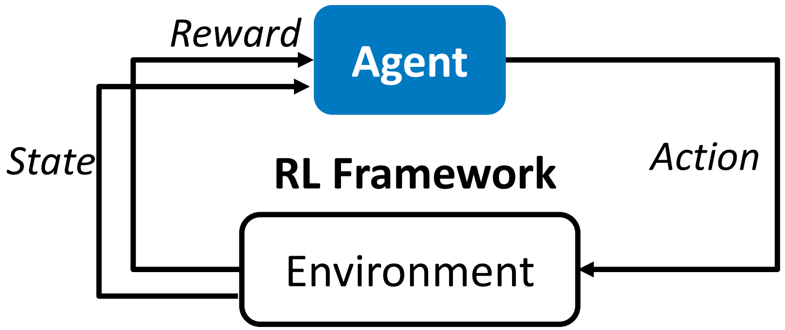
<u>Action</u>:
Turbine control valve (TCV) and
Bypass valve (BV) positions (0 to 1).

<u>Reward</u>:
$e = |s - \frac{P_{BOP\_demand}}{P_{BOP}^{max}}|$; minimize error.
$r = \frac{1}{e + \delta}$

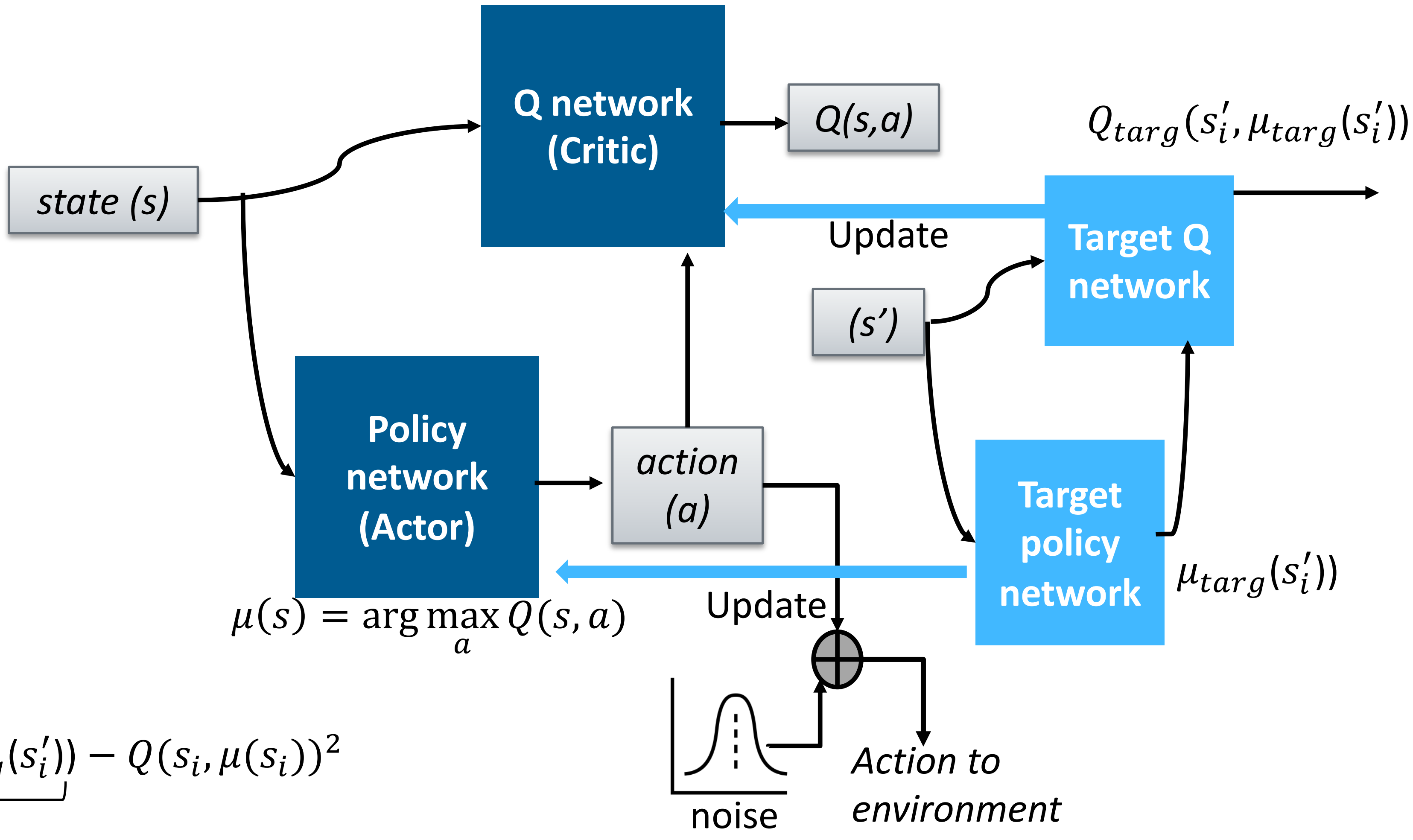
Episode
varying
power
profiles



Learning Framework



Deep Deterministic Policy Gradient (DDPG) algorithm



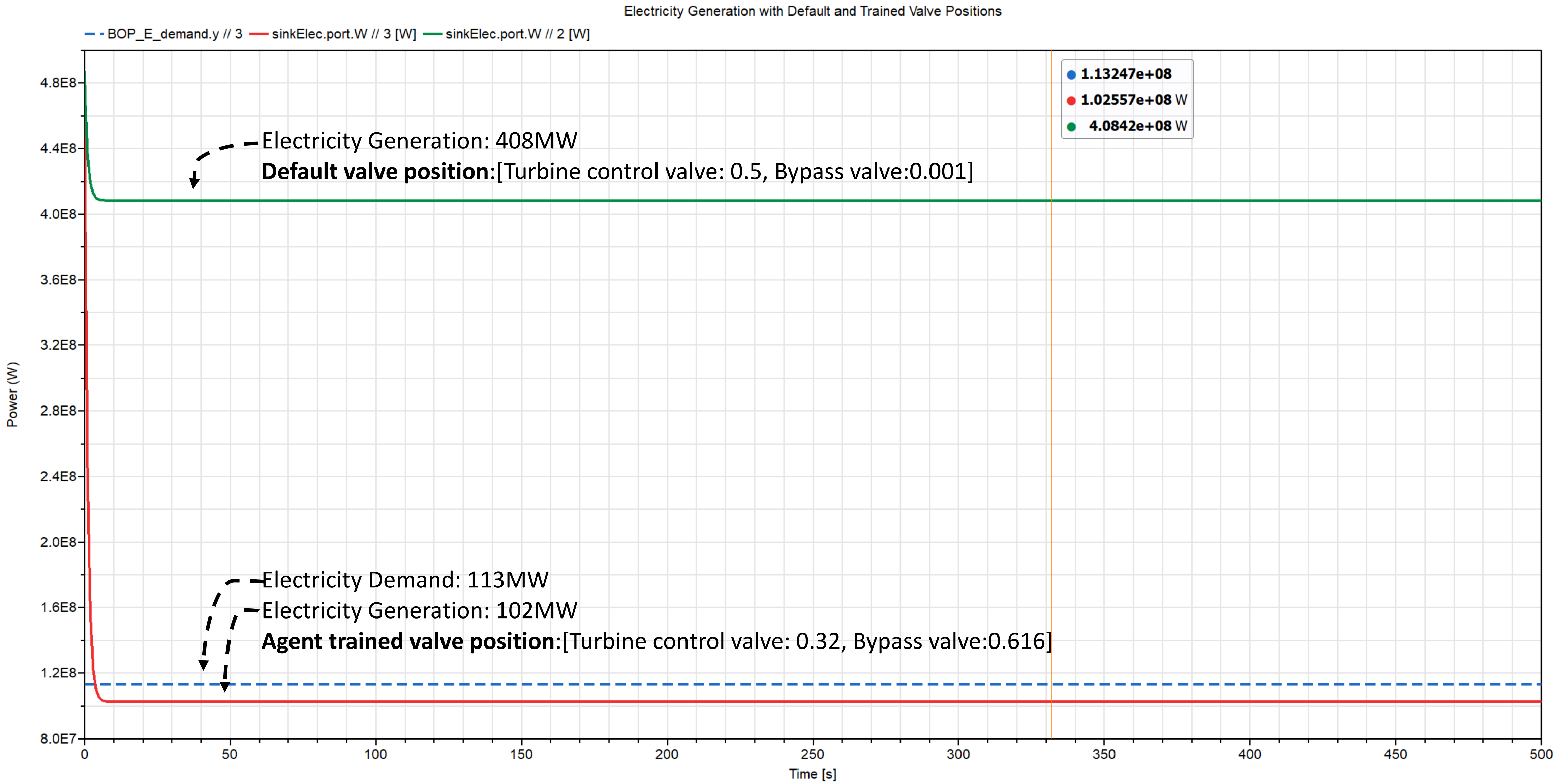
- Four feedforward neural networks are used.
- The target networks are delayed networks compared to main networks.
- The weights of targets are updated periodically based on the main networks.

<u>Loss functions :</u>

$$J_\mu = \frac{1}{N} \sum_{i=1}^{N} Q(s_i, \mu(s_i))$$

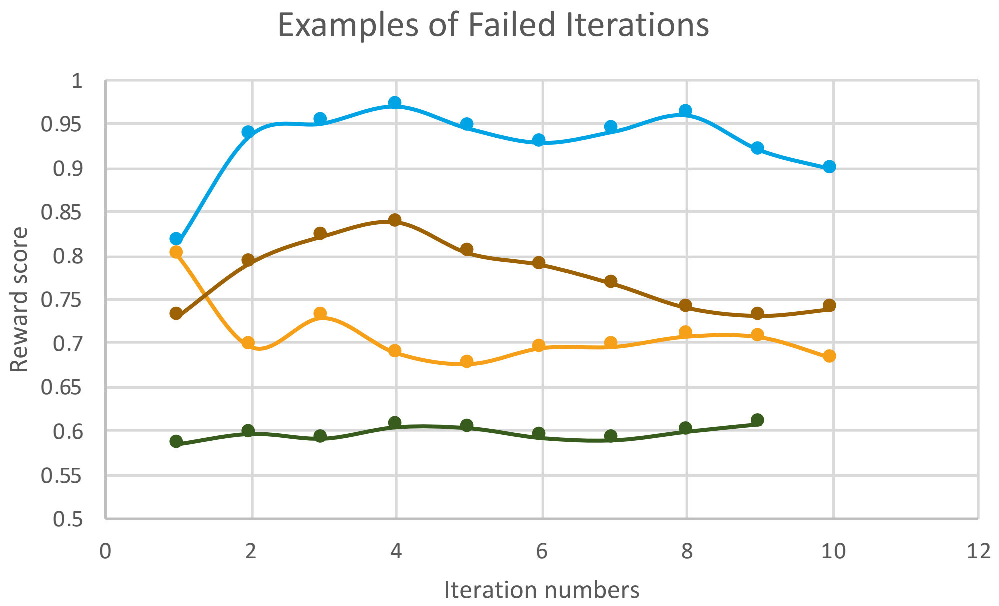
$$J_Q = \frac{1}{N} \sum_{i=1}^{N} \underbrace{r_i + \gamma(1-d)\, Q_{targ}(s_i', \mu_{targ}(s_i'))}_{\text{Bellman equation}} - Q(s_i, \mu(s_i))^2$$

Electricity Generation with Default and Trained Valve Positions



Reinforcement Learning Agent Training Progress (On-going)

- **The agent generates normalized electricity demand while learning how to control valves to meet the power demand.**
- **Performed training with multi step multi episodes**



- **Multiples episodes indicated general trend during iterations, but not for every episode.**

Conclusion and Challenges

- **The RL agent training process revealed a trend in which the agent can learn both turbine control and bypass valve positions based on electricity demand.**
- **There is still work to be done before the agent is fully trained.**
  - During training, the agent became stuck at the boundary conditions.
  - The working range of the action item does not completely match the working zone of the agent. By fixing this issue, the agent can be trained more sufficiently.

Funded by


multidisciplinary
design
optimization
linear
system
failure
solar
data
cost
engineering
deformation
statistical
battery
energy
experiments
fidelity
optimum
surrogate
learning
reliability
evaluation
methodology
decomposition
renewable
mathematical
stochastic
algorithm
bounds
illustration
exploration
minimization
approximation
objective
complex
equivalence
uncertainty
computing
criteria
metamodeling
wind
inequality
nonlinear
integer
displacement
probabilistic
mining
bayesian
gradient
heuristic
storage
discrete
machine
power
recursive
structural
loads
value