## Objectives:

- Independently defined setpoint schedule for tank 1 and 3, with different values

- Minimize tracking error over a 10-hour horizon

Generalized Policy Search

- Direct policy search over neural network parameter space using random sampling + local refinement

- Evaluates N=32 trajectories in parallel using batch simulations
- Trained for 1000 epochs
- Initial rewards around **-2000**
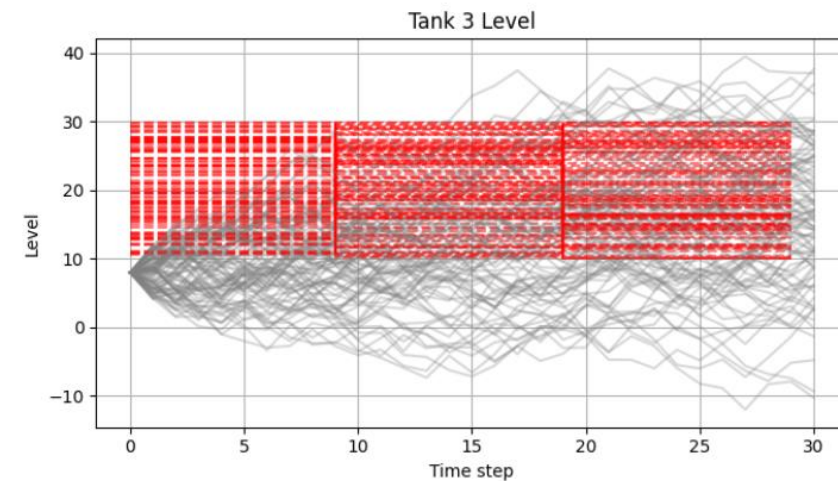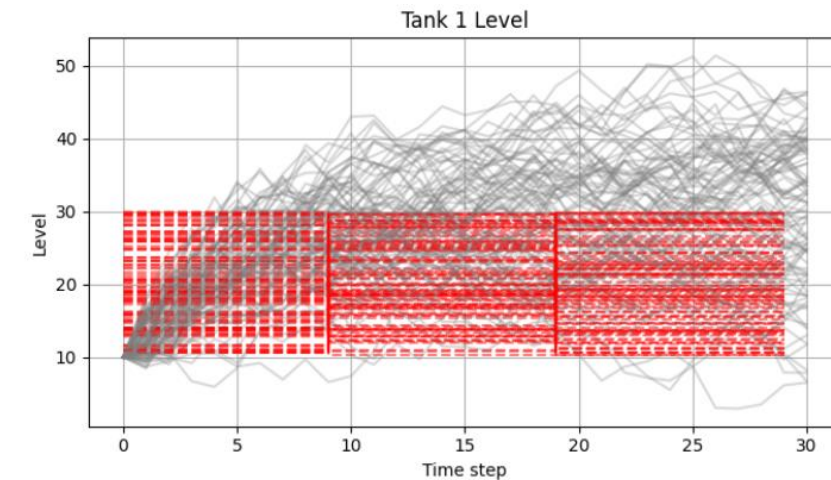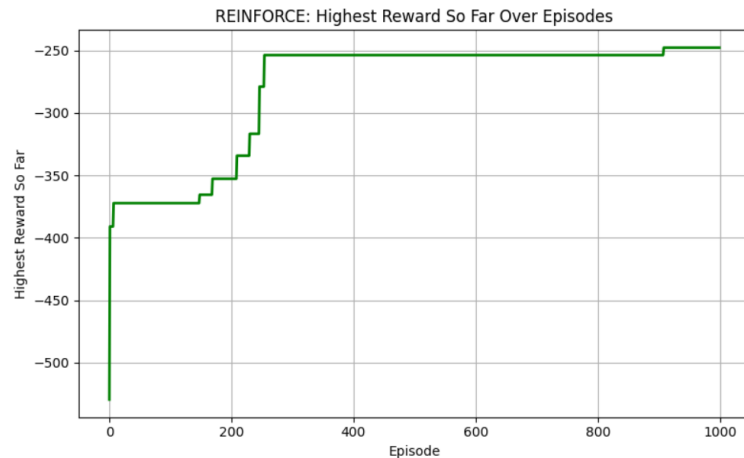
Proximal policy optimization

- With PyTorch
- Trained for 300–1000 epochs
- Initial rewards around **-10000**