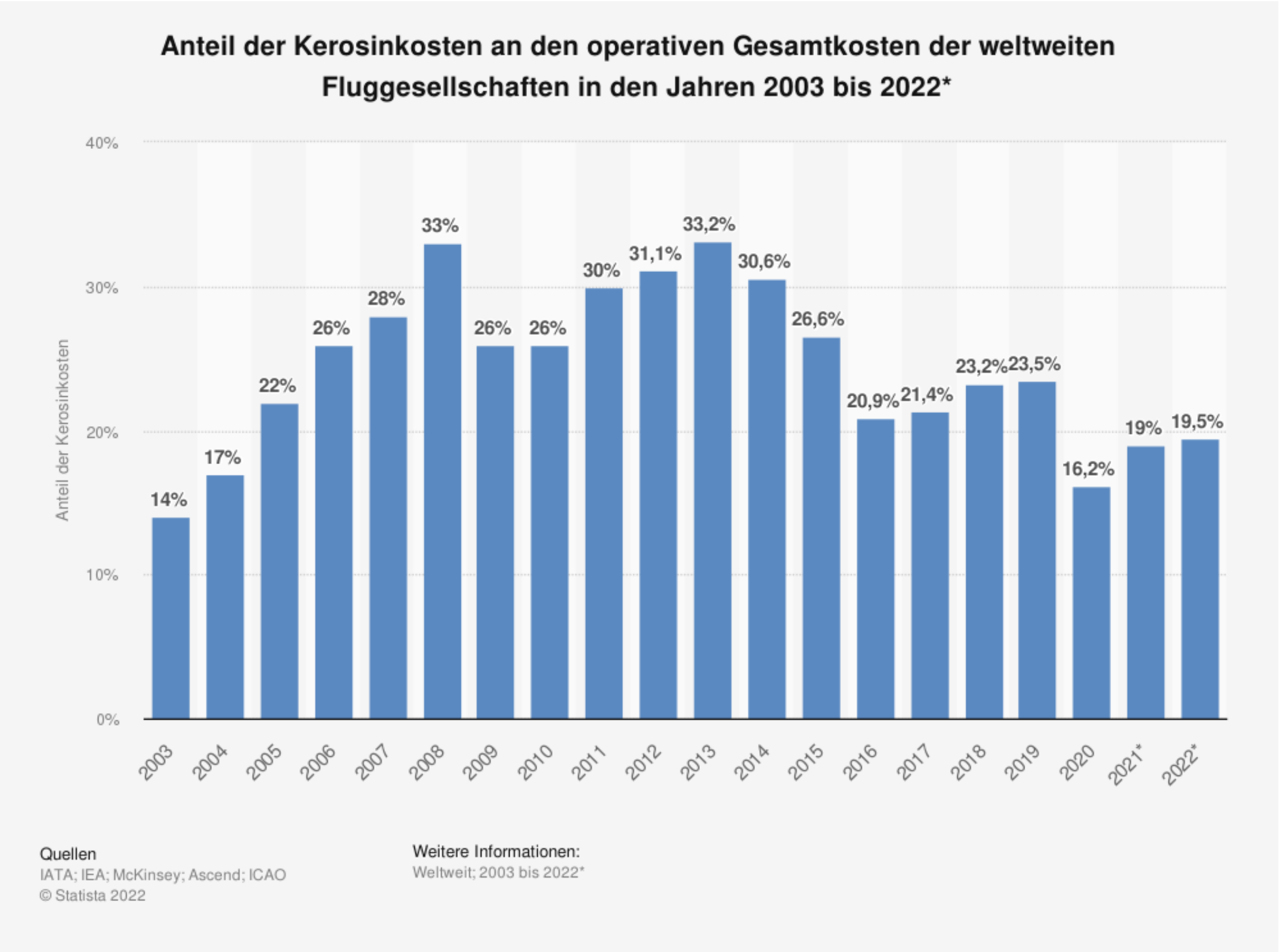


RL-Project

AirportTowerEnv

Marc Speckmann, 14.03.2022

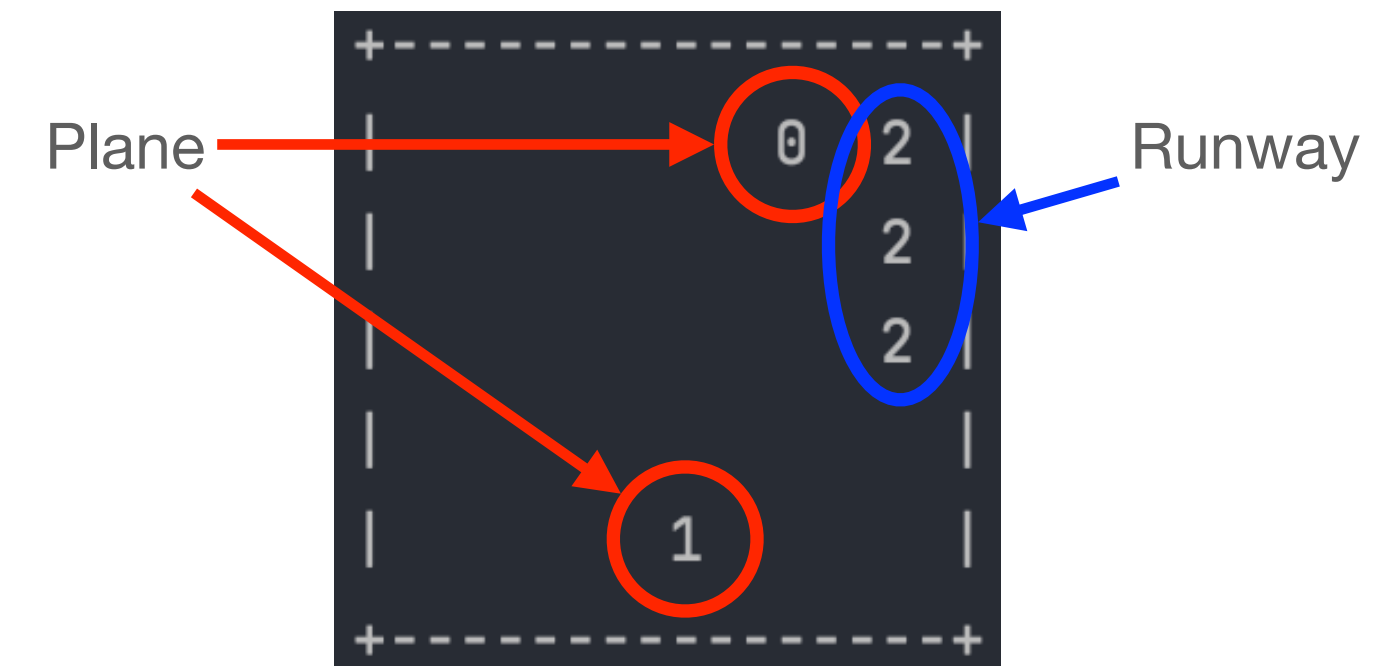
Motivation



Environment

Implementation

- OpenAi Gym Environment
- Observation space:
 - 5x5 Array
- Action space:
 - #plane x 4 directions



Environment

Overview

- Goal: Survive 200 steps
- Rewards:
 - Landed plane reward: 100
 - Plane in air penalty: -1
 - Plane on runway reward: 5
- Plane spawn rate: 0.3 on each step

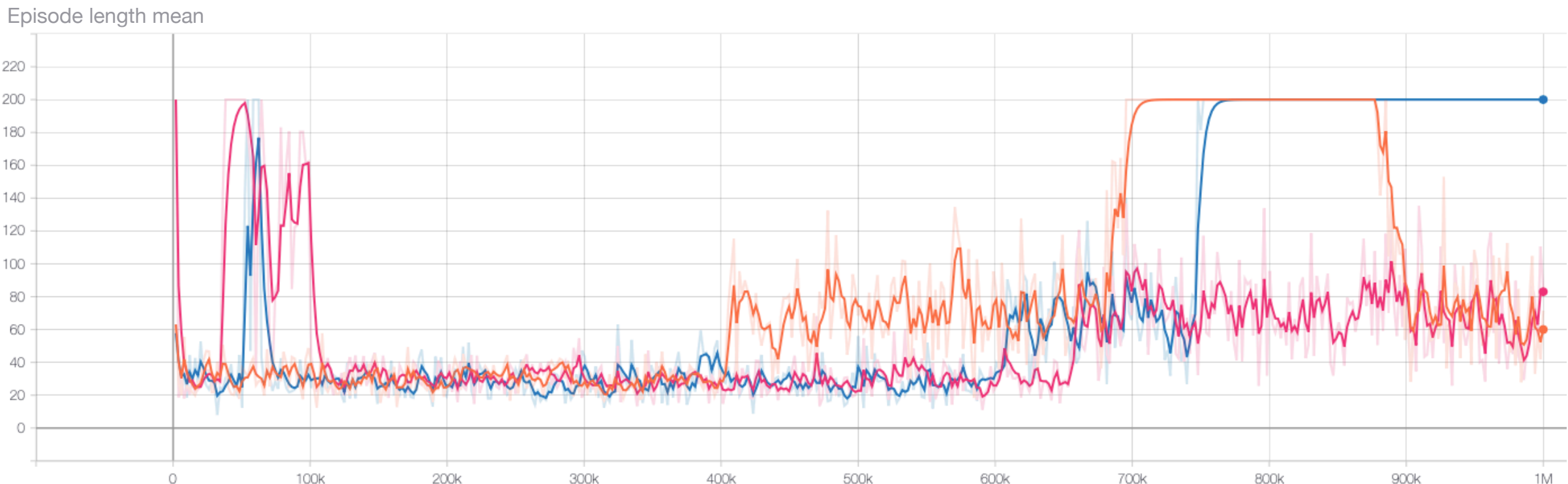
Training

Settings

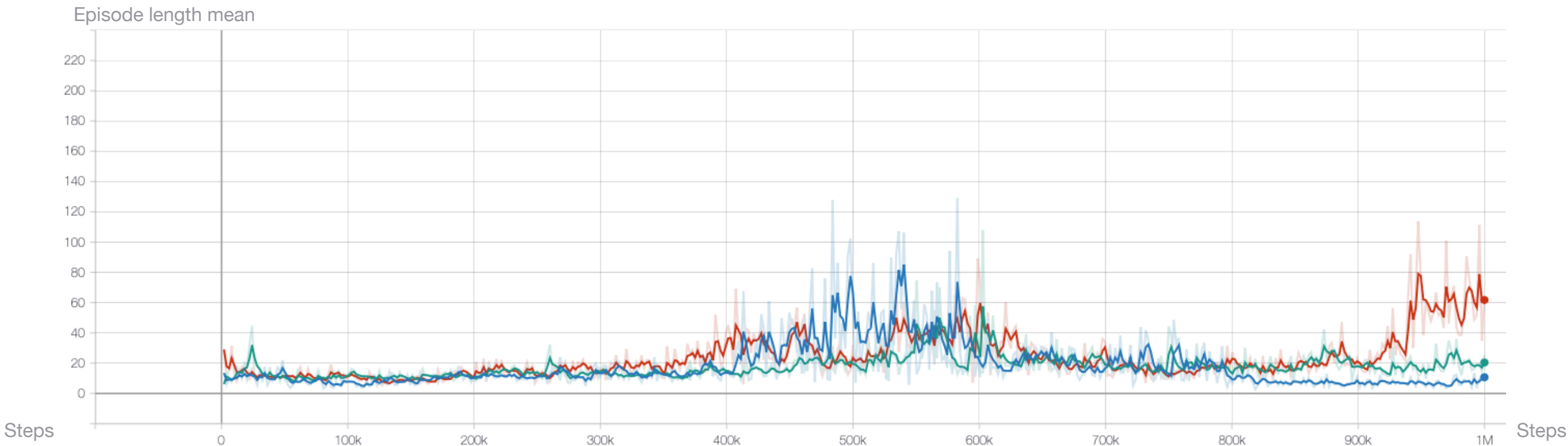
- Python(3.8.12), RLlib(1.9.2), Tune(1.9.2), Tensorflow(2.7.0), Gym(0.21.0), Numpy(1.22.2)
- Training on Cluster of University (Slurm, 8 CPU (Skylake), 8 GB RAM, per Agent)
- Double Duelling DQN with priority replay
- 3 different random Seeds per trial
- 1 Mio. Steps \approx 1000 Iteration, 10 Episodes every second iteration for Evaluation
- EpsilonGreedy (1.00 \rightarrow 0.02, 0 \rightarrow 950,000)
- Exact settings in GitHub repository

SingleAgent Environment

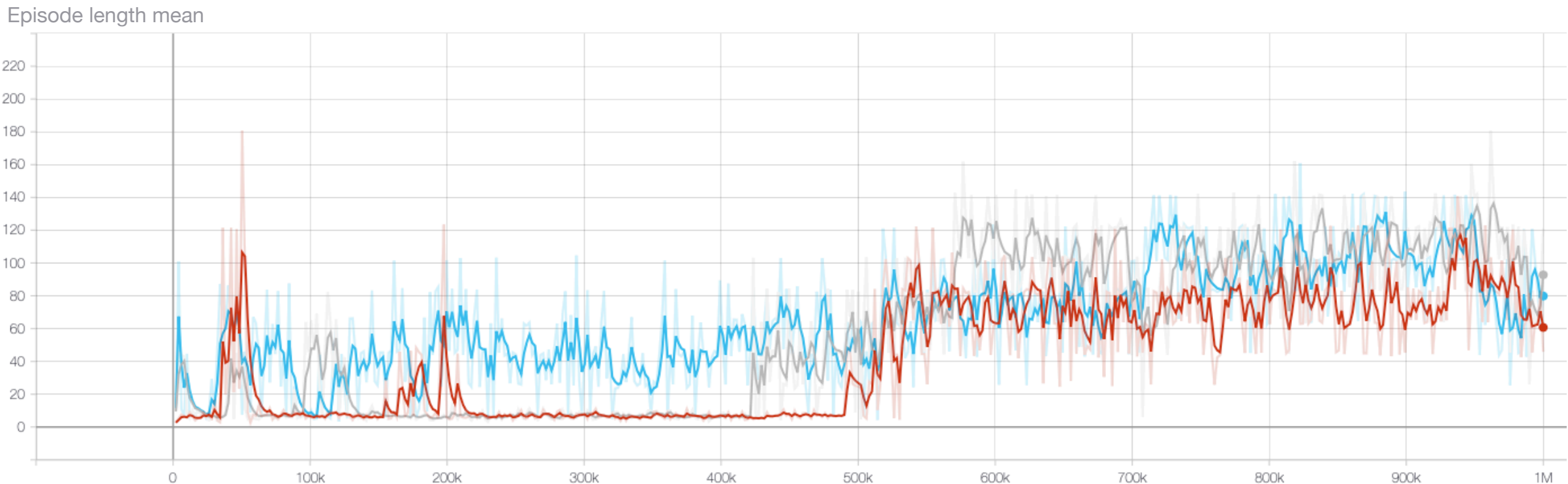
Evaluation Results



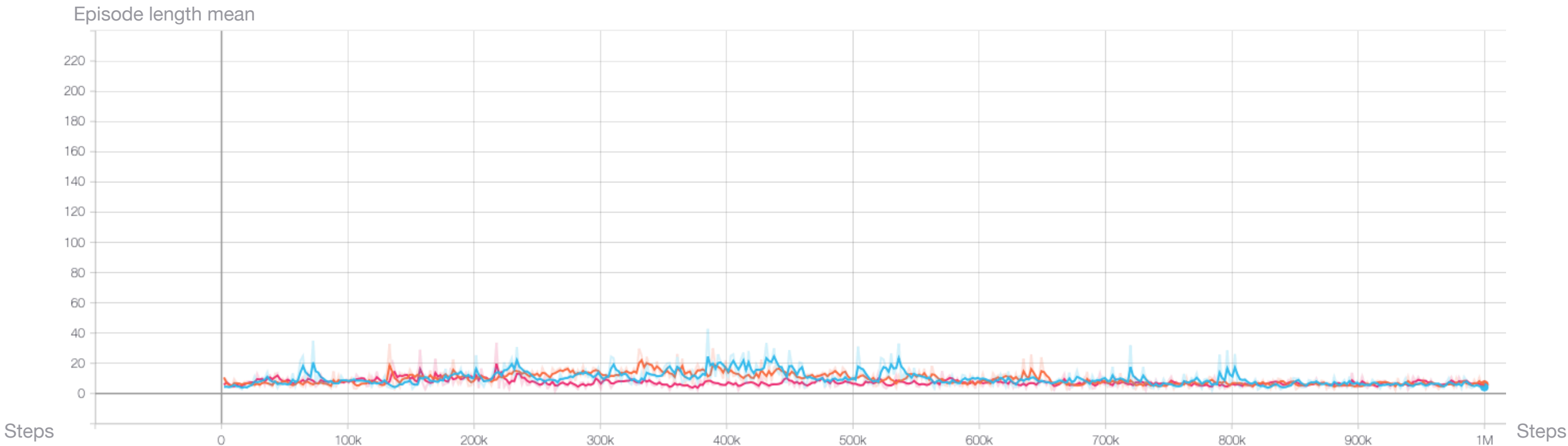
1 Plane, 1 Runway



2 Planes, 1 Runway



1 Plane, 2 Runways

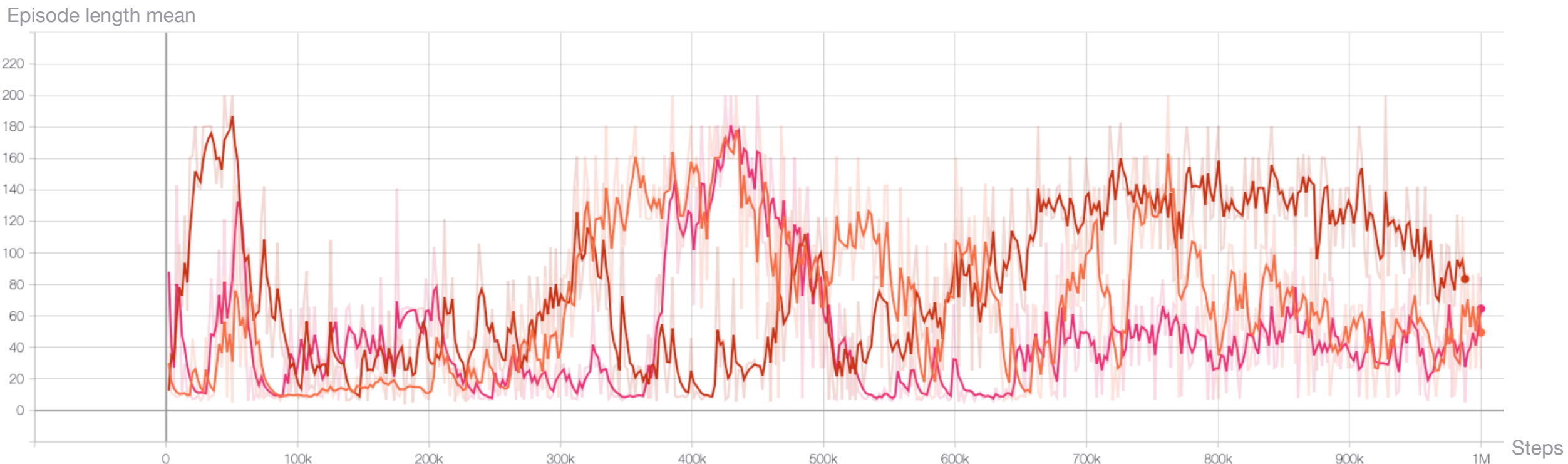


2 Planes, 2 Runways

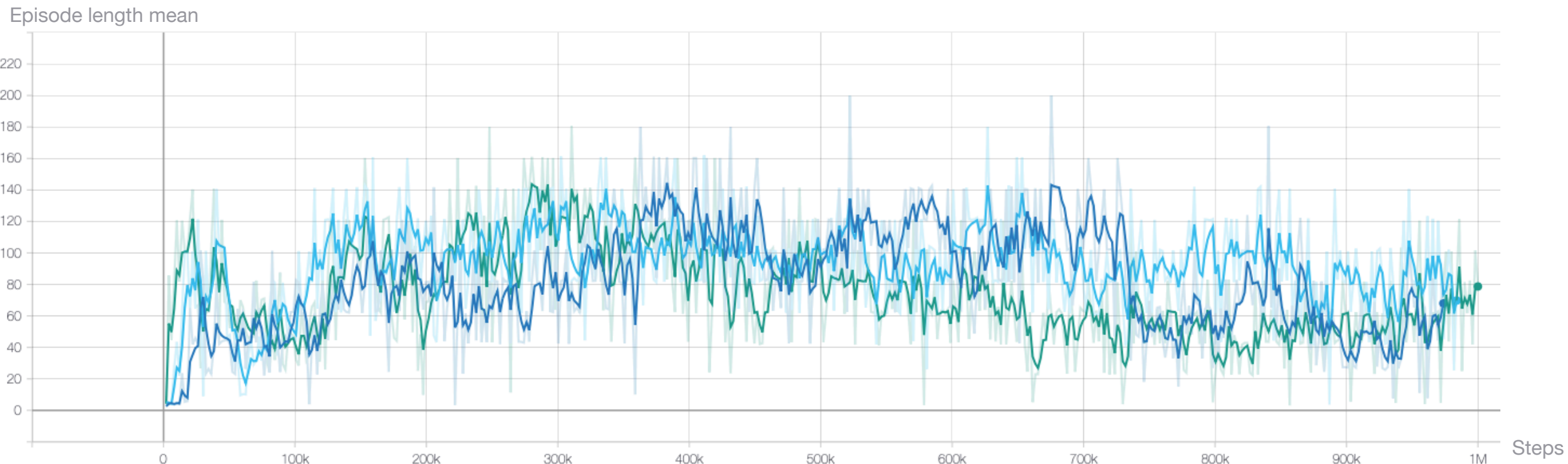
Seeds: [24088626, 30953886, 20735918], Eval_env=[666]

MultiAgent Environment

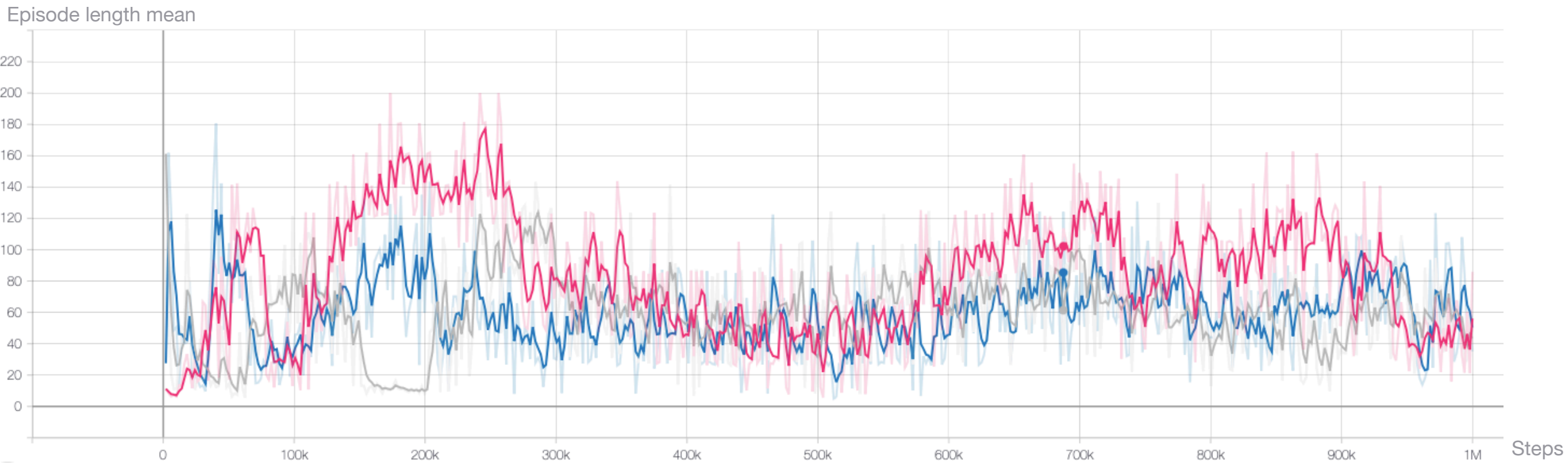
Evaluation Results



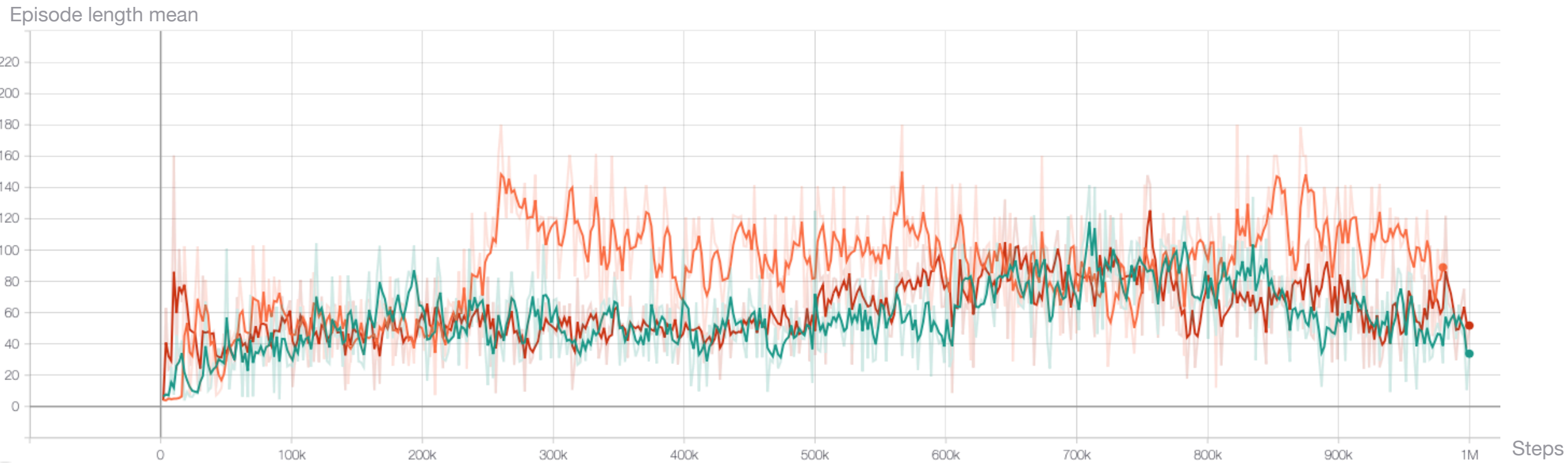
2 Planes, 1 Runway, 2 Agents with same policy



2 Planes, 2 Runways, 2 Agents with same policy



2 Planes, 1 Runway, 2 Agents with different policy



2 Planes, 2 Runways, 2 Agents with different policy

Seeds: [24088626, 30953886, 20735918], Eval_env=[666]

Outlook

- Implementation of real flight rules.
- Support rewards/penalties for abide rules
- Bigger Environment
- Different RL algorithm & Hyperparameter optimisation
- Curriculum learning

Thanks for the attention



GitHub