Airport Runway Configuration Management with Offline Model-free Reinforcement Learning

Milad Memarzadeh, Tejas Puranik, Krishna Kalyanam, and Wes Ryan
NASA Aeronautics Research Institute, NASA Ames Research Center

SCITECH Forum, Jan 23-27, 2023

Decision support for the air traffic control

- Runway configuration management (RCM) is a challenging task, and it affects the efficiency of the National Airspace System (NAS) and airport surface operations significantly.

- Many factors such as traffic load, wind conditions or other environmental factors, safety measures and regulations and noise abatement procedures might affect the choice of the configuration.

- Current process is manual from the perspective that humans consider a range of characteristics and determine the "best" configuration based on experience and expert knowledge. This usually results in sub-optimal decision-making.

- Automated approaches based on Artificial Intelligence (AI) and Machine Learning (ML) can be used to improve the quality of the decision-making process by assisting the controllers.

Our solution: offline model-free RL

Offline RL combines reinforcement learning with data-driven machine learning. The reasons as to why we adopt this methodology are:

- Removes the need for an online interaction with the operational/simulation environment.
- Utilizes the vast amount of historical data to learn a good policy and build a powerful decision-support engine for the controllers.

<u>Challenges</u>:

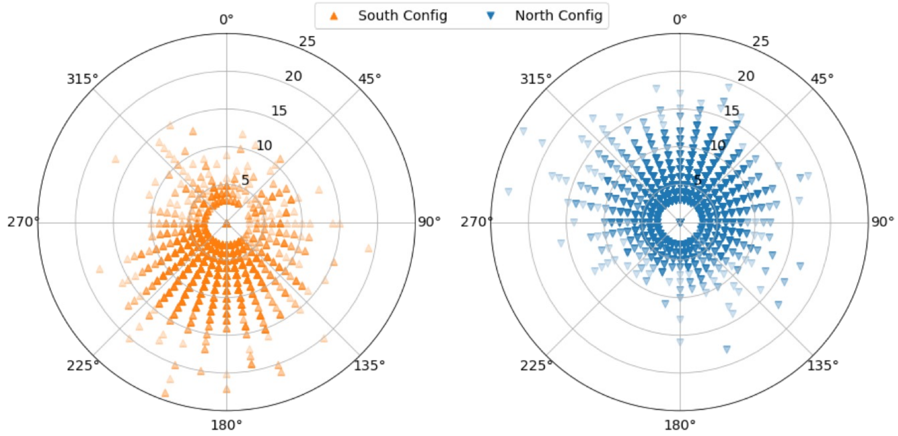
- The offline nature of the algorithm does not allow any exploration.
- Distributional shift, where agent's learned policy deviates significantly from the behavior policy.
- Being overly optimistic about Out-Of-Distribution (OOD) data in operational setting.

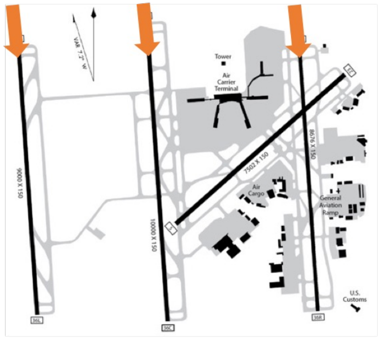
We deploy Conservative Q-Learning (CQL) [9] to provide a solution to the RCM:

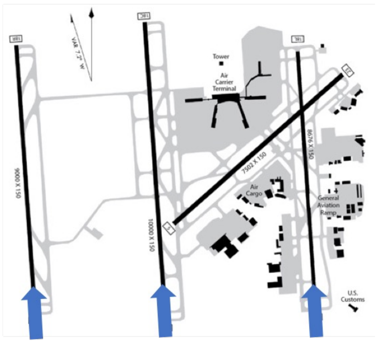
- CQL uses a simple mechanism to regularize the estimates for the OOD data to prevent unsafe behavior
- It also prevents excessive distribution shift.

[9] Kumar, et al., *Arxiv*, 2020.

Runway diagram at Charlotte Douglas International Airport (CLT)







CLT is our first proof-of-concept airport:

- It has two major configurations: North flow and South flow.
- There is a strong correlation between the wind conditions and the choice of runway configuration.

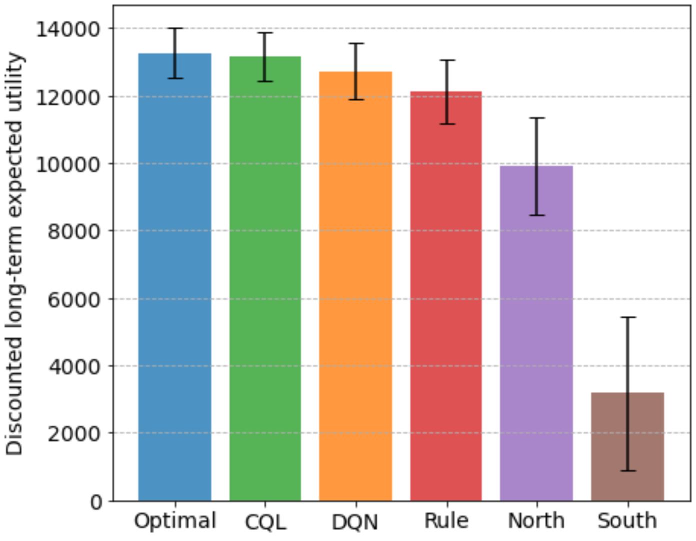
Results: forward simulation

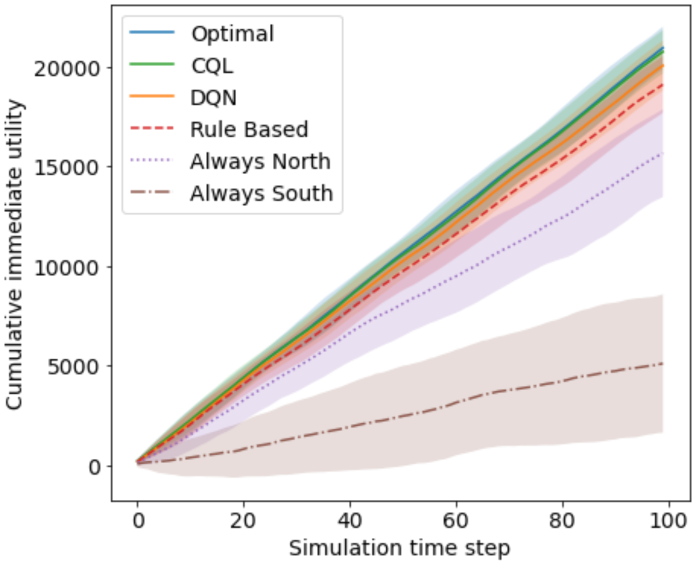
Here we compare performance of different policies compared to the optimal policy based on 100 independent forward simulations and 50,000 episodes of training:

$$V^{\pi} = \sum_{t=0}^{T} \gamma^t u(s_t, \pi(s_t))$$

Expected net present value of utility under policy π

utility at time t according to policy π

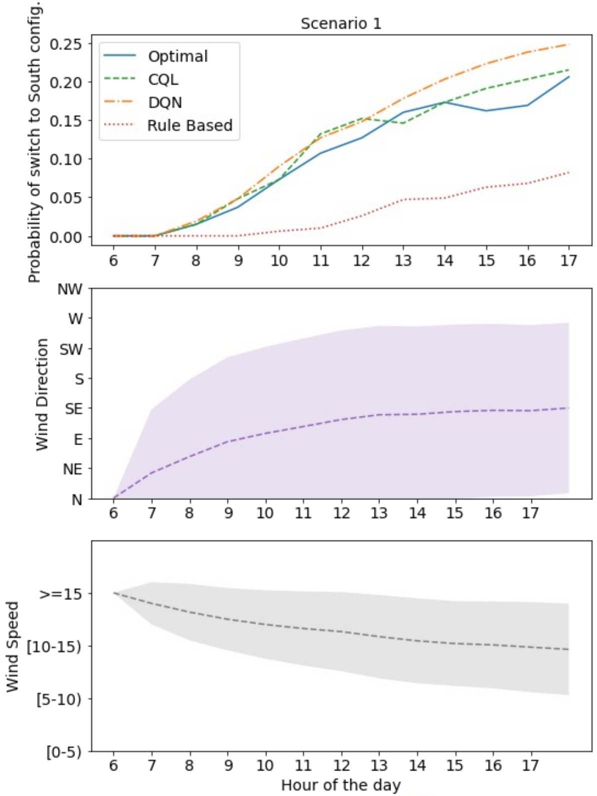




Forecast in the operational setting

In this scenario, we are at the current state as follows,

- Wind is blowing from North with intensity level 3 (>15 knots).
- It is 6am.
- The optimal policy right now is to select North configuration.
- Y-axis shows the probability of switch to South configuration.



Concluding remarks and next steps

We deployed a state-of-the-art offline model-free RL methodology, Conservative Q-Learning (CQL), to address the runway configuration management.

We showed that CQL performs better than more traditional offline RL approaches as well as simplified rule-based policies.

Lastly, we illustrated the performance of CQL in an operational setting for forecasting the decision-making quality in RCM for up to 12 hours in the future.

Next steps: validate the CQL with the RCM problem at CLT and test its generalizability to more complex airports across the National Airspace System.

Corresponding author: Milad Memarzadeh, milad.memarzadeh@nasa.gov