

# Meeting

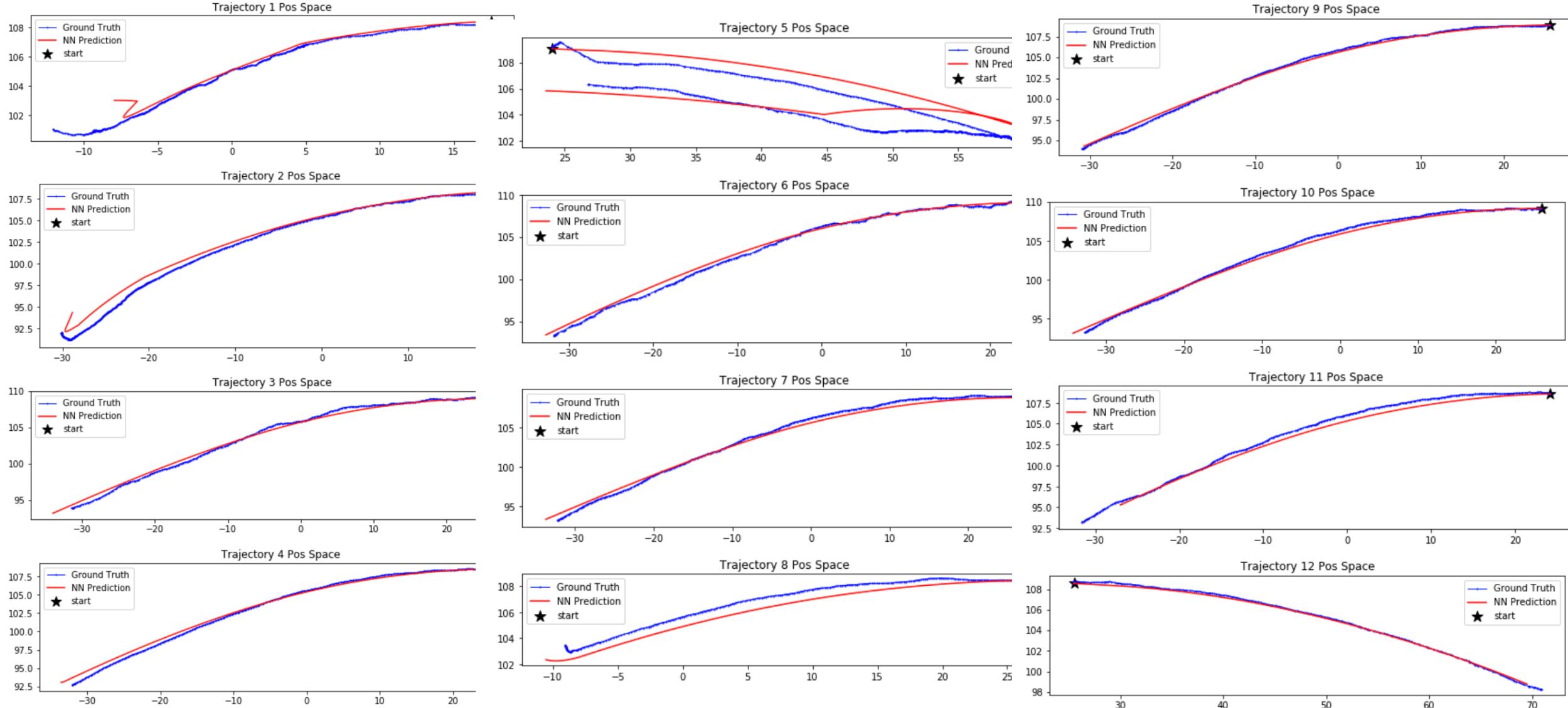
# 07/03/2020

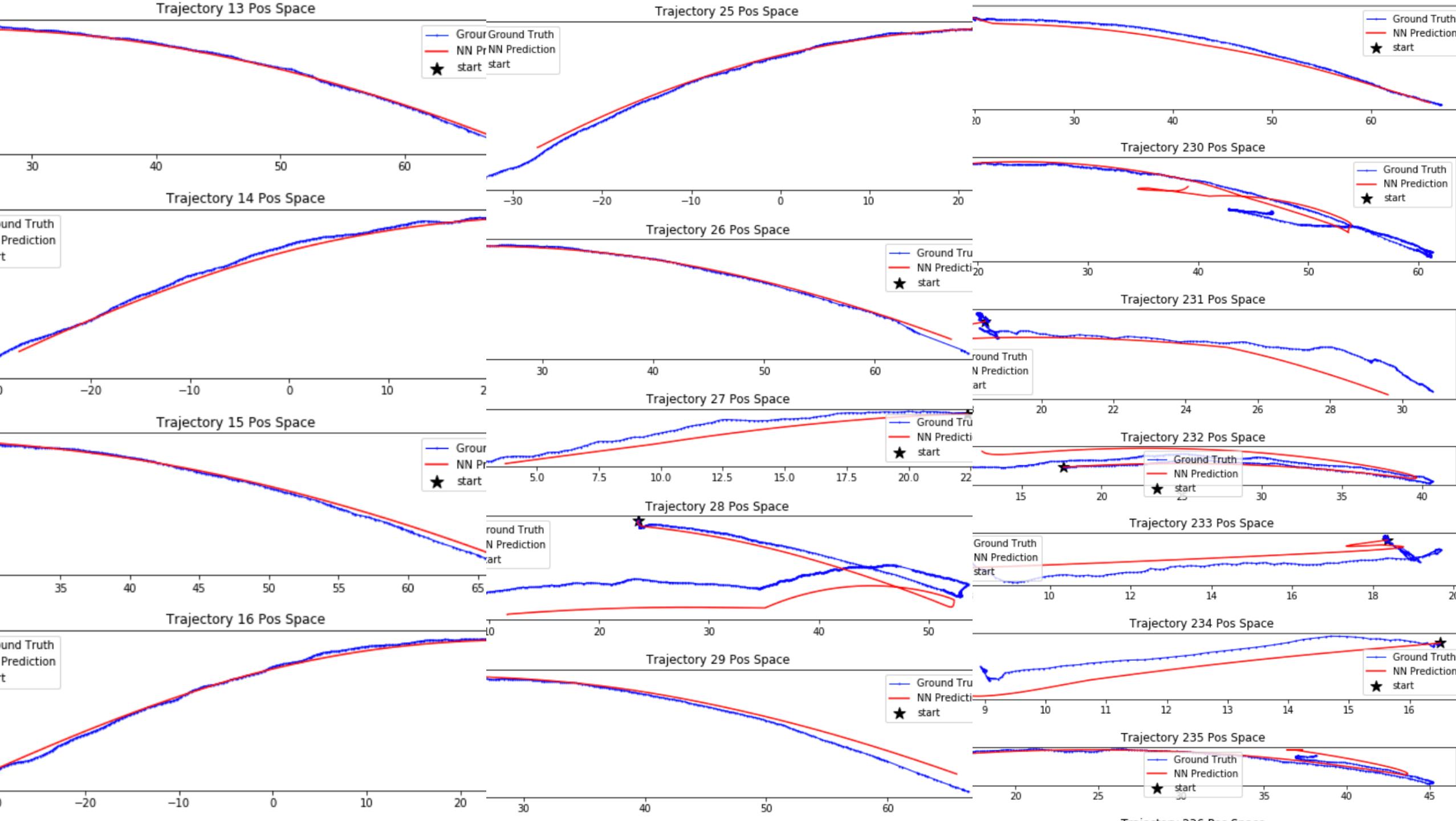
Shuo Zhang

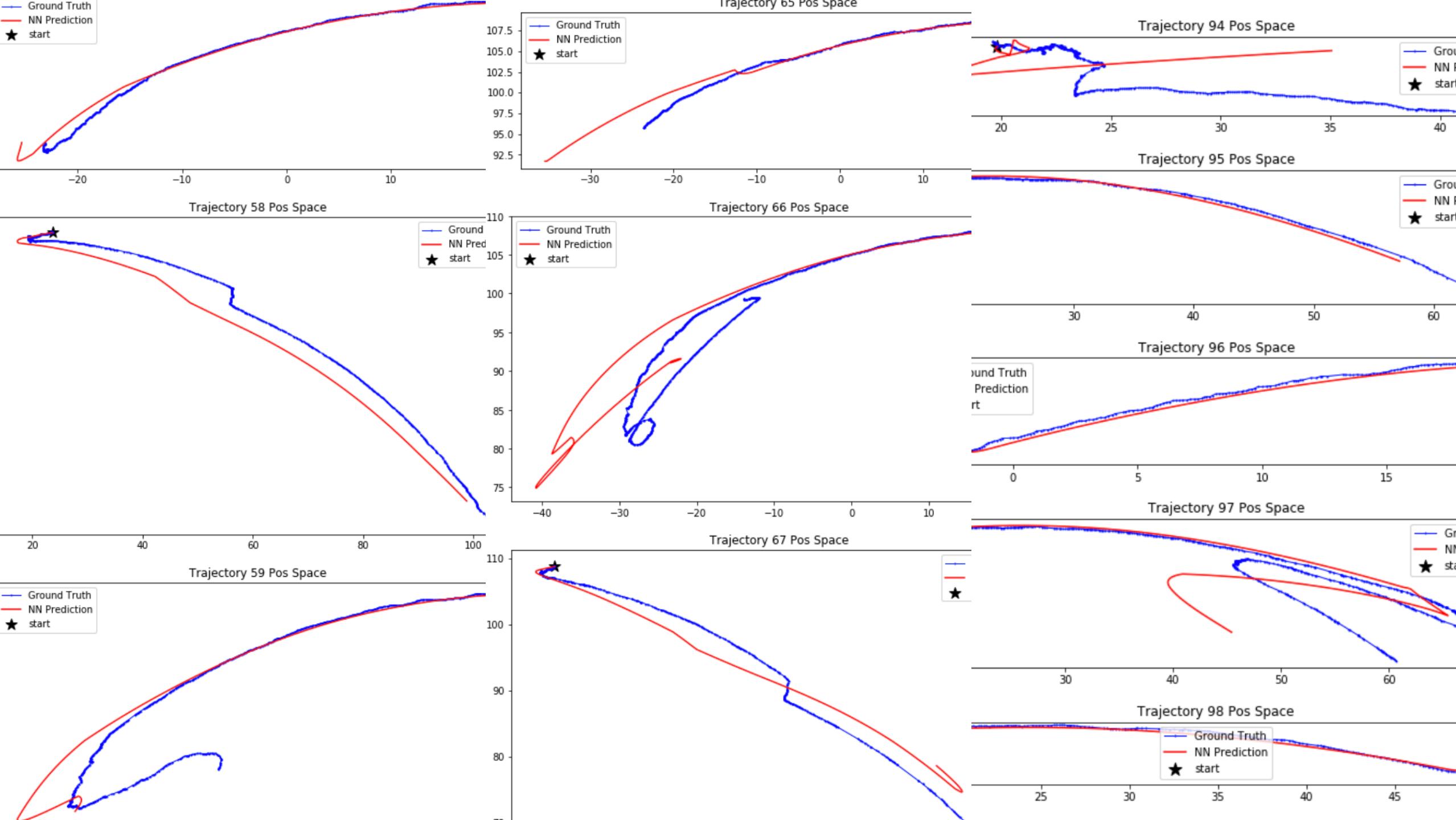
# Progress

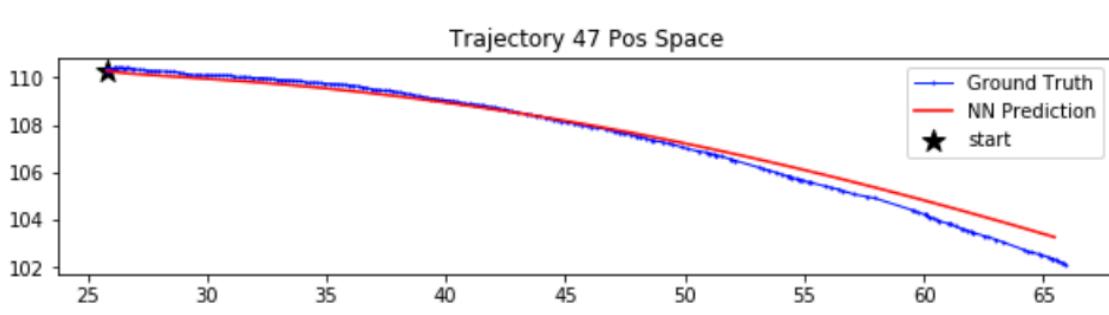
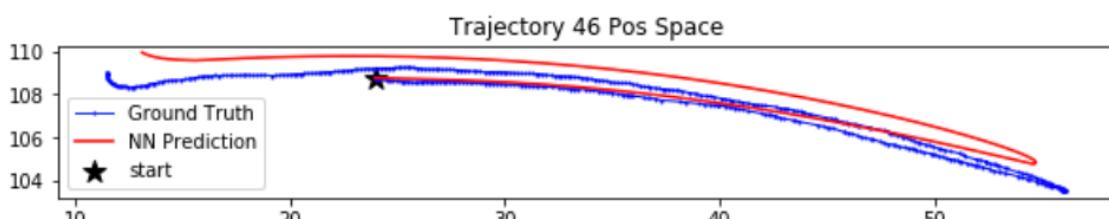
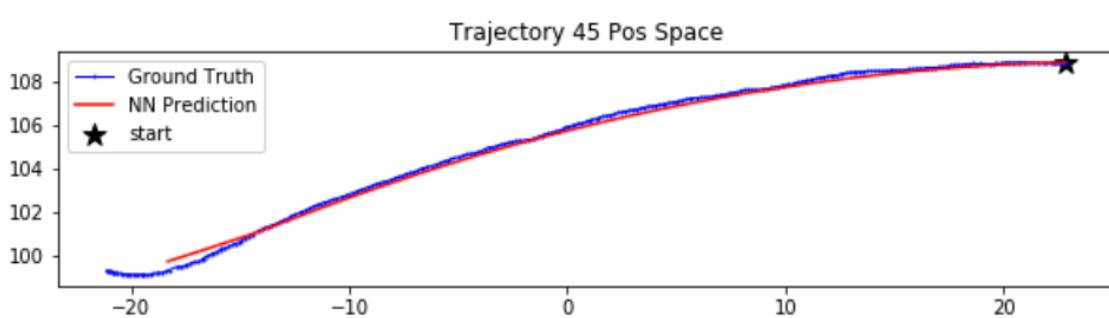
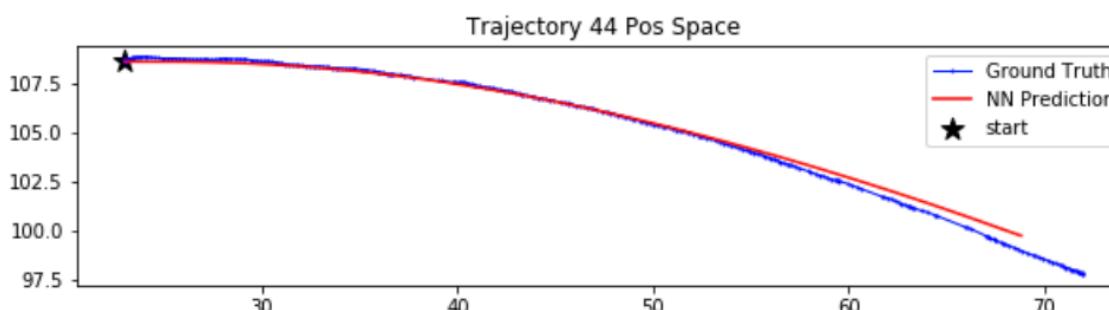
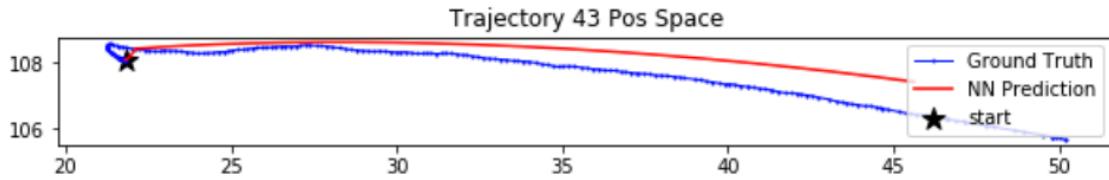
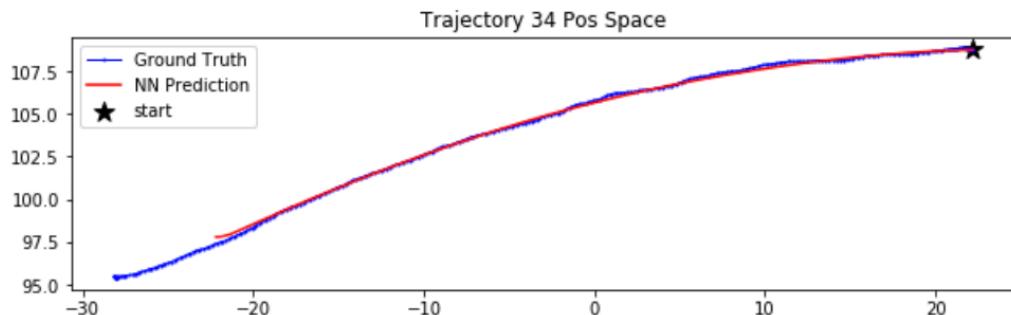
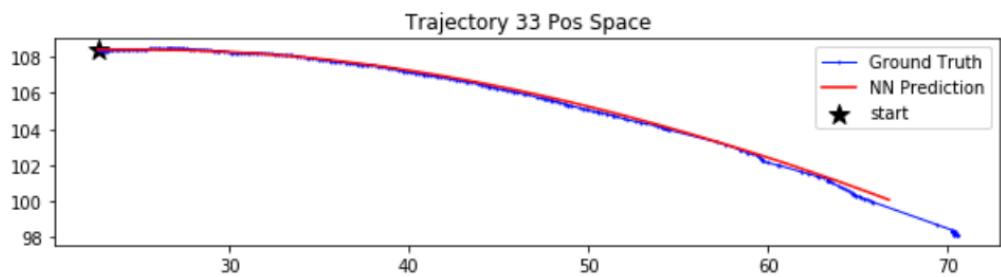
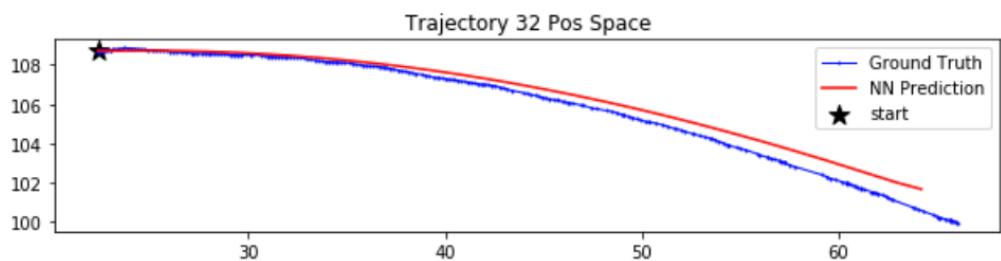
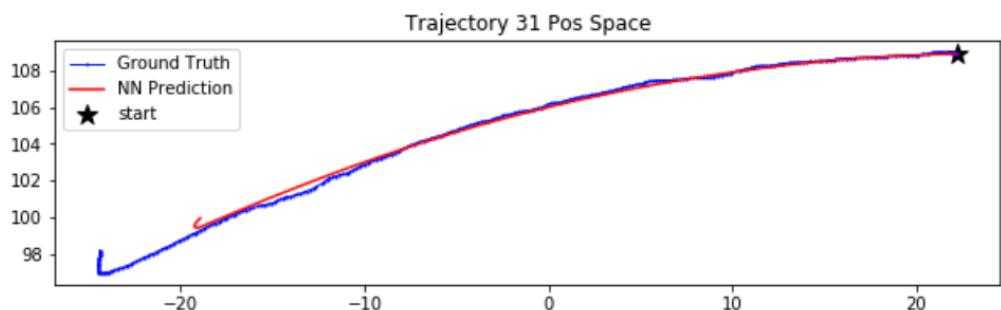
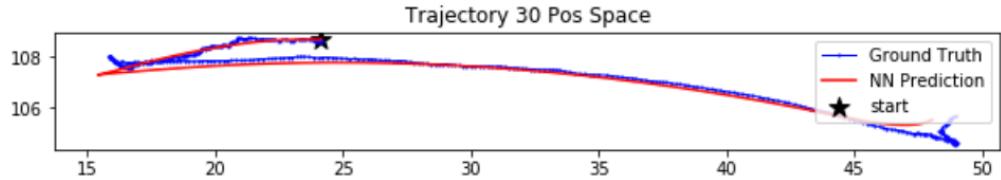
	A* Plan	A* Rollout	PPO Plan	PPO Rollout	Online Method
Acrobot	Done	Done	Done	Done	Need to discuss in more details
Reacher	Done	Done	Done	Done	Need to discuss in more details
Gazebo Hand (normal obstacles + horseshoe )	Done	Done	Done	Done	Need to discuss in more details
Real Had (no obstacles)	Done	Wait for Mridul until the marker issues fixed	In Progress (Hyperparameter Search)	Wait for Mridul until the marker issues fixed	Need to discuss in more details

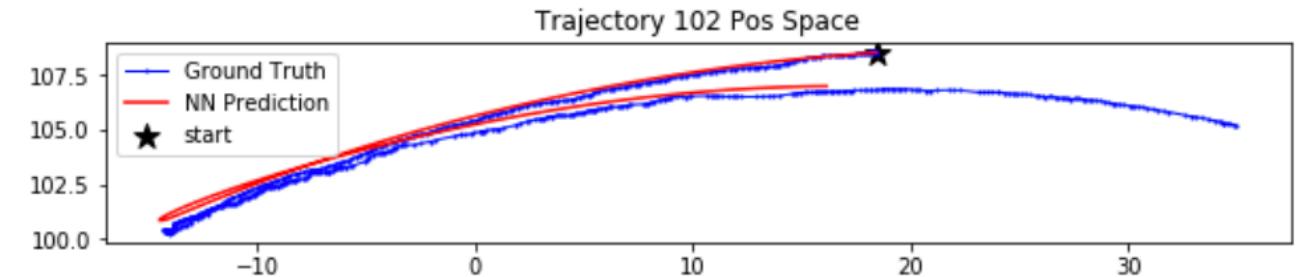
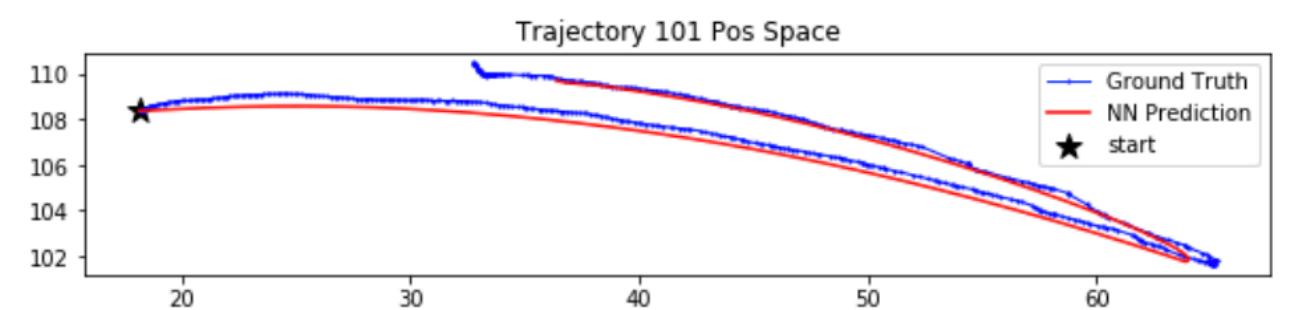
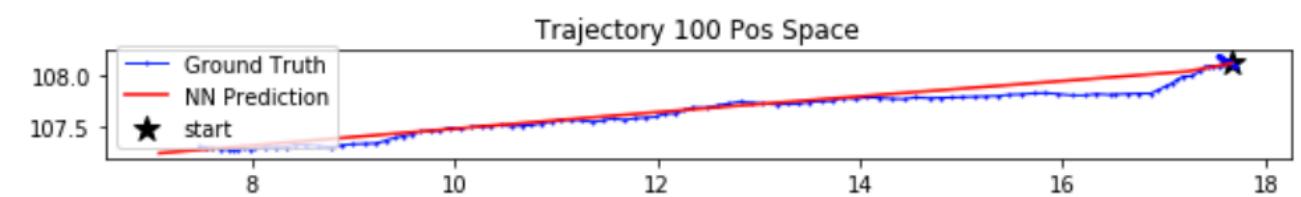
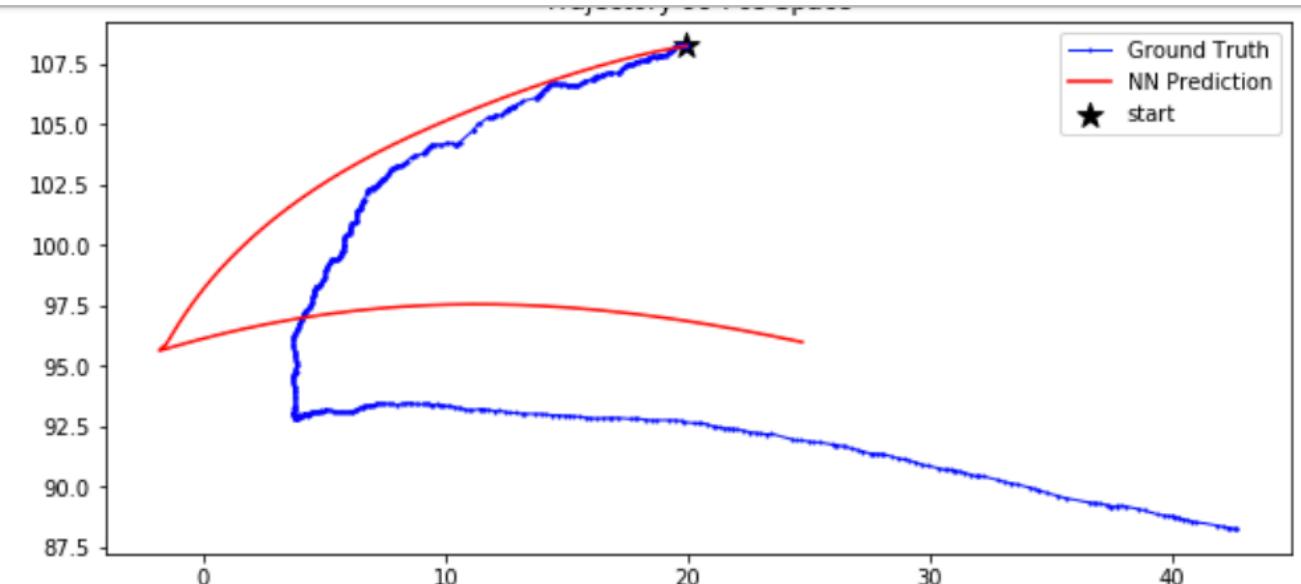
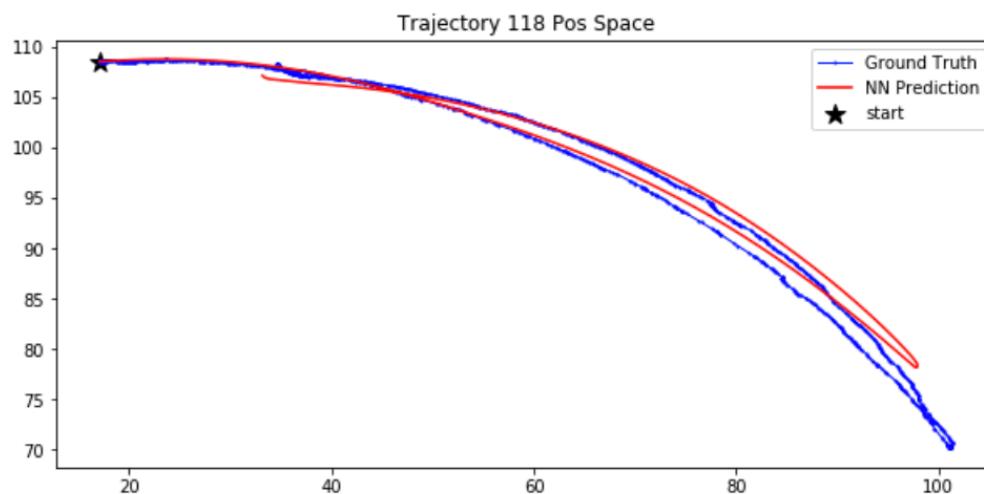
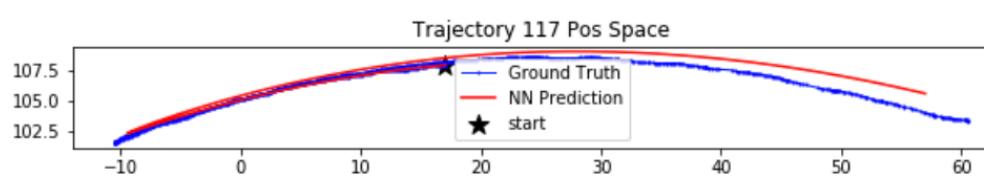
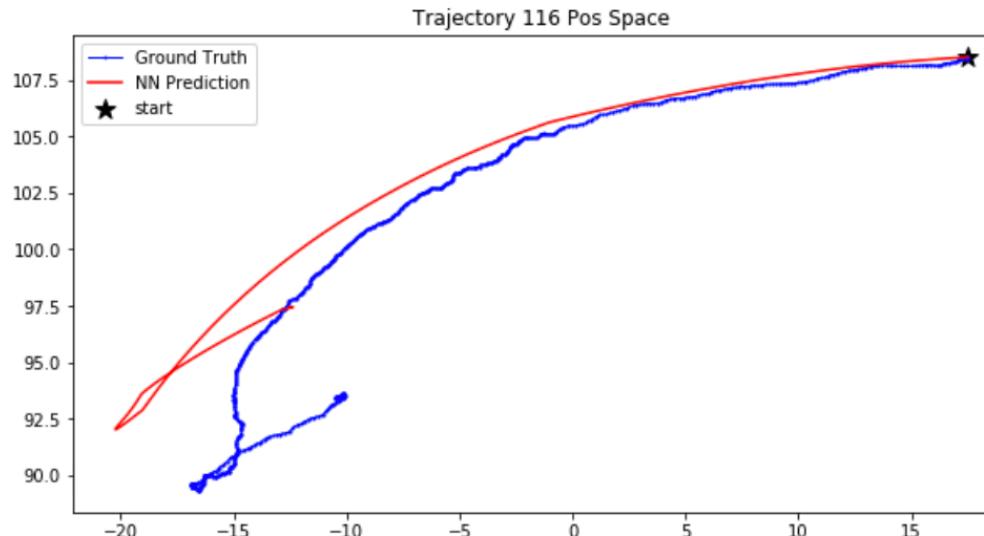
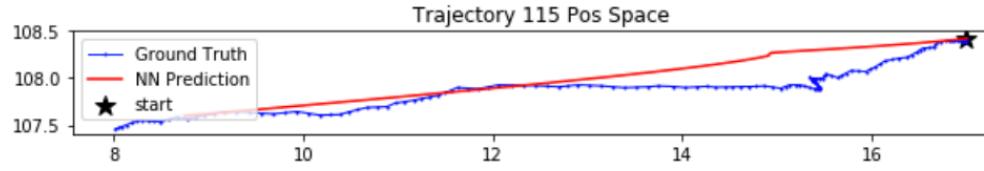
# Real Hand Dynamic Model (missing z values are calculated)

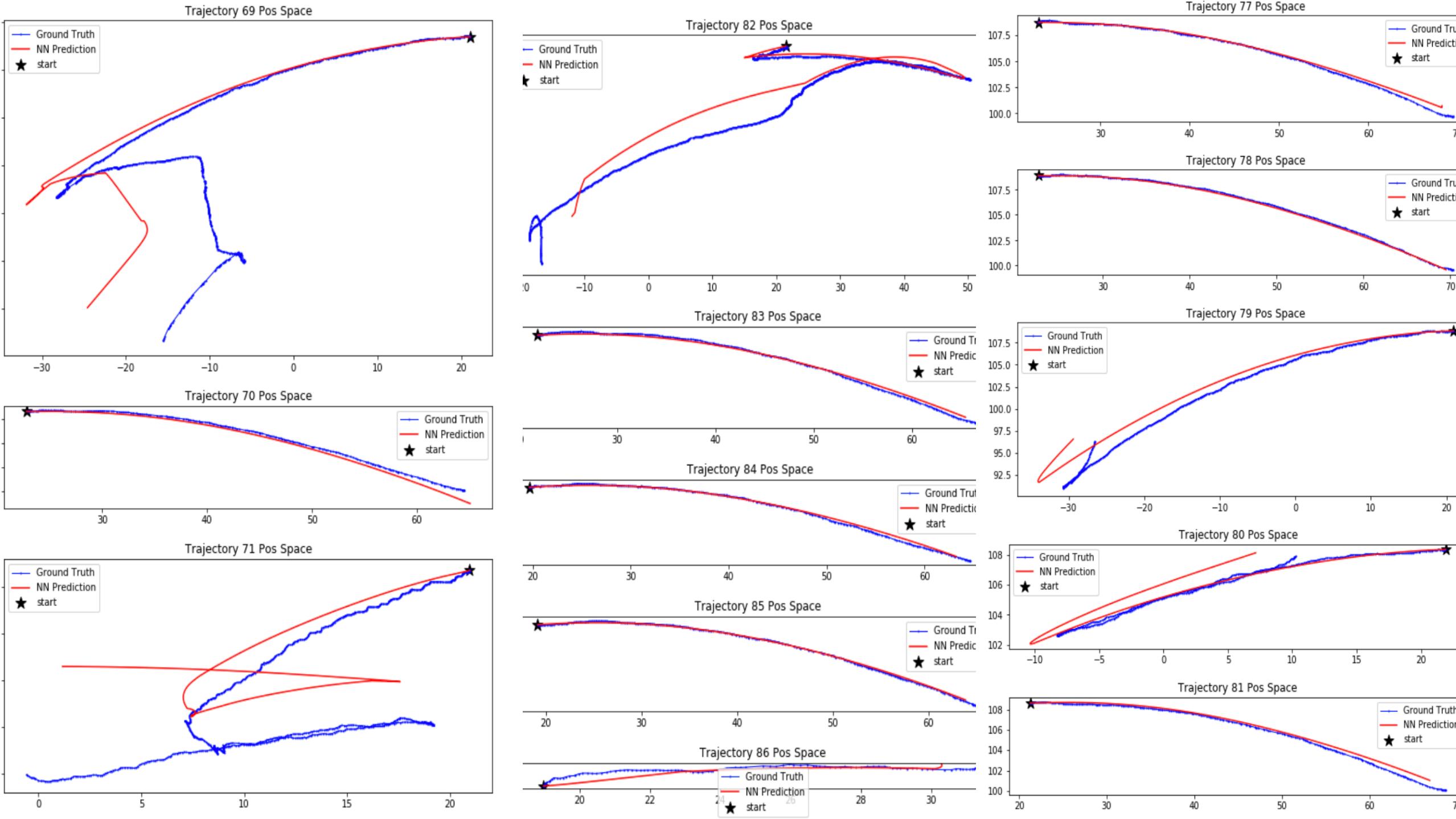


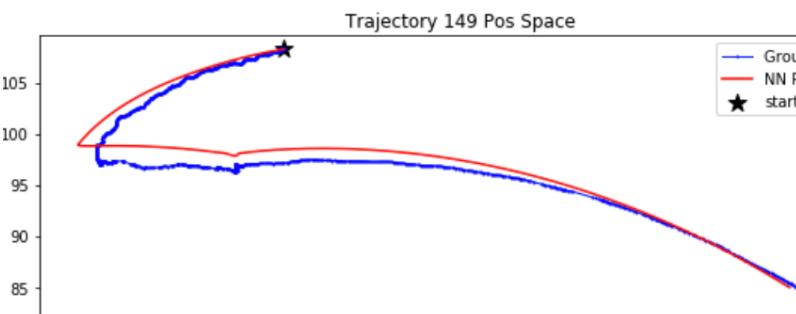
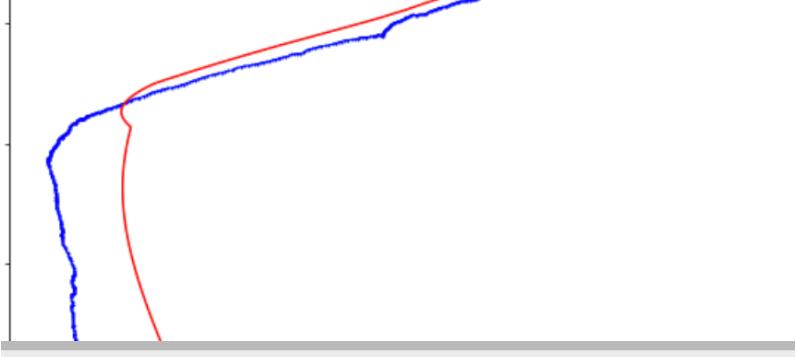
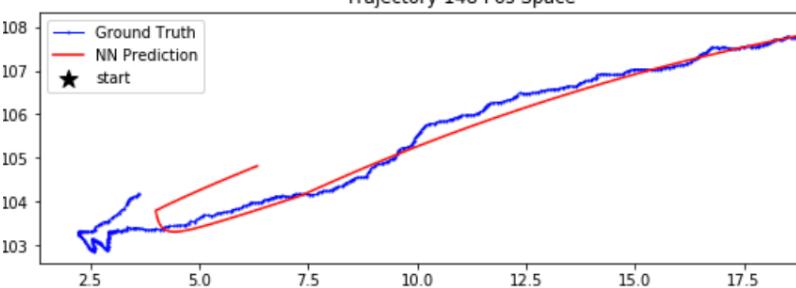
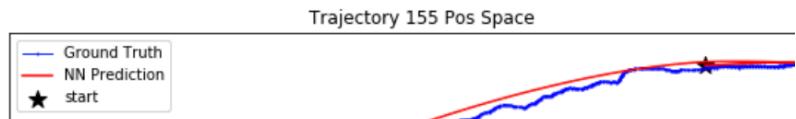
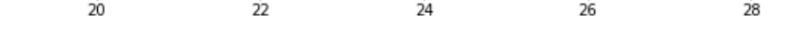
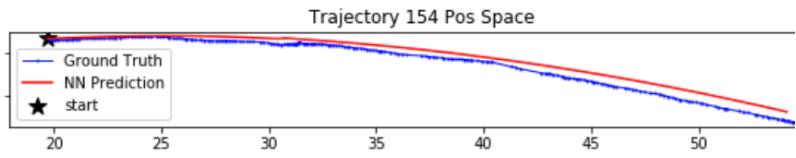
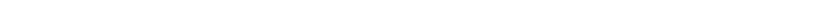
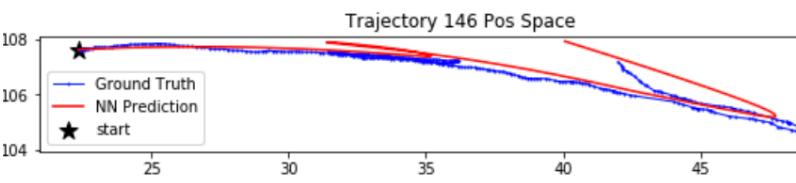
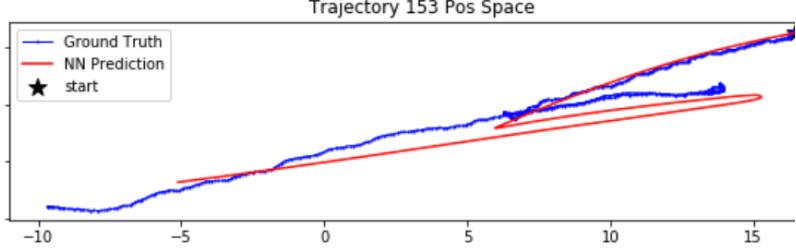
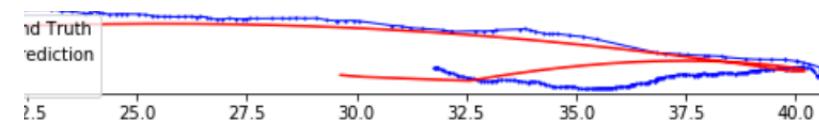
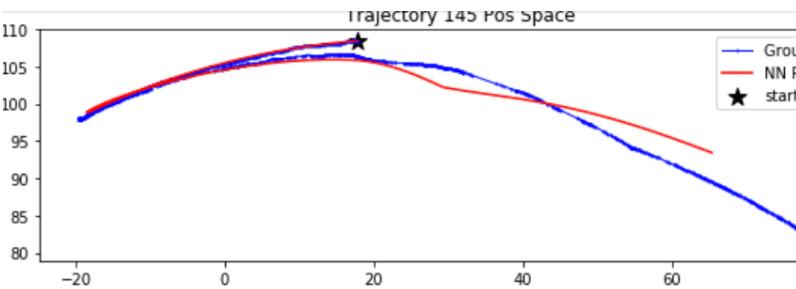
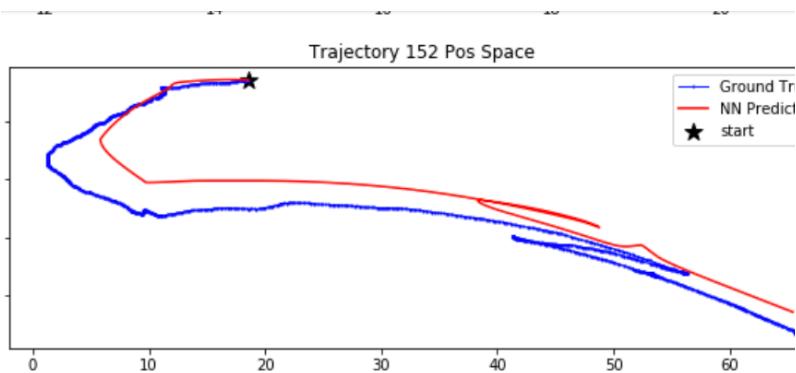


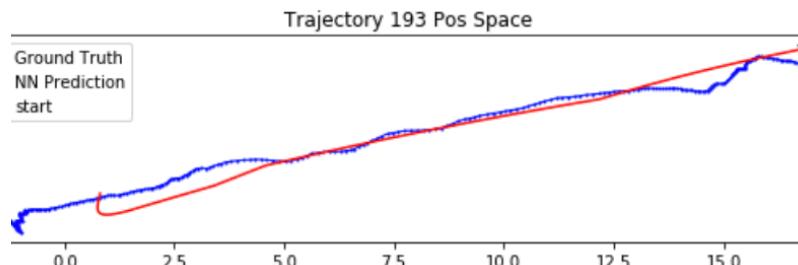
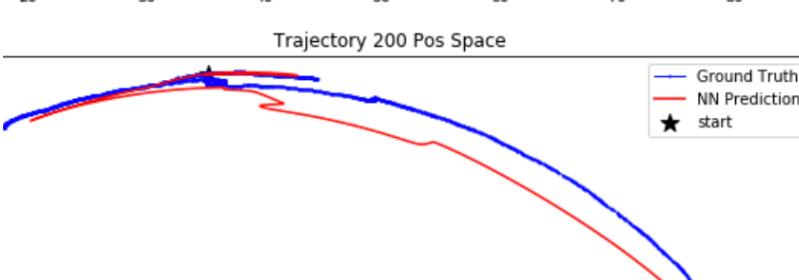
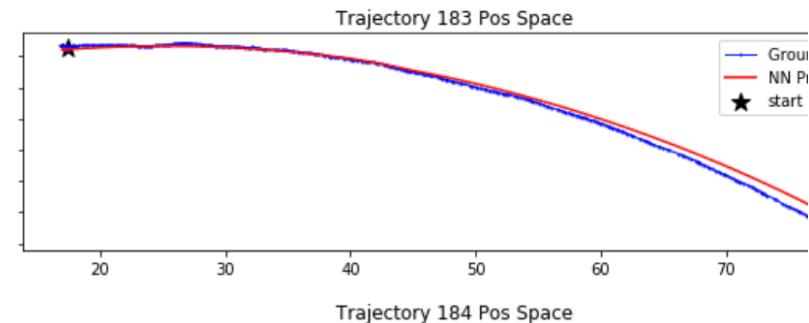
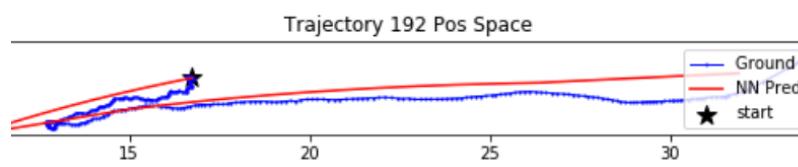
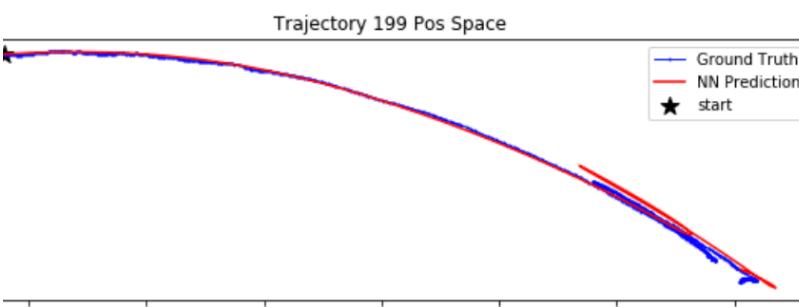
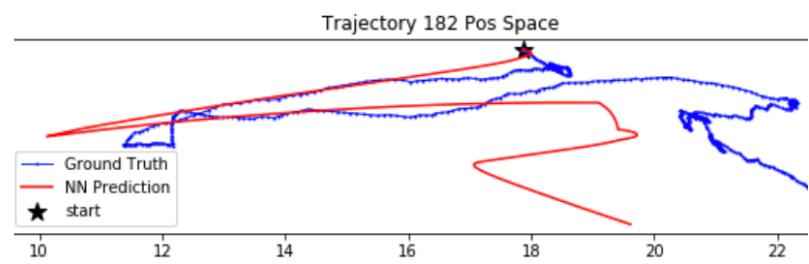
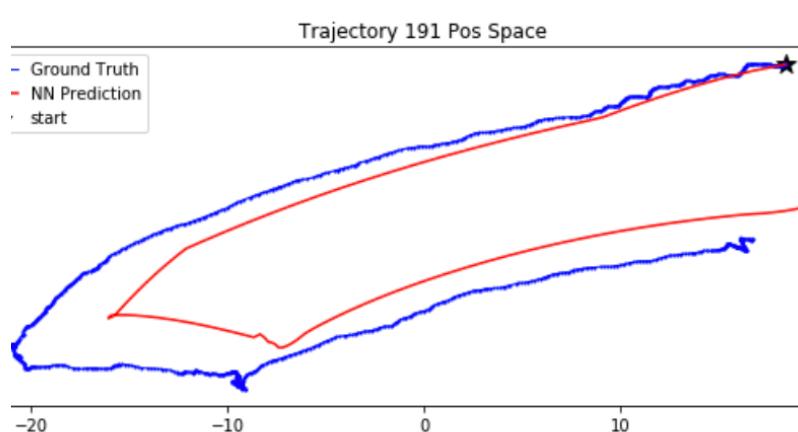
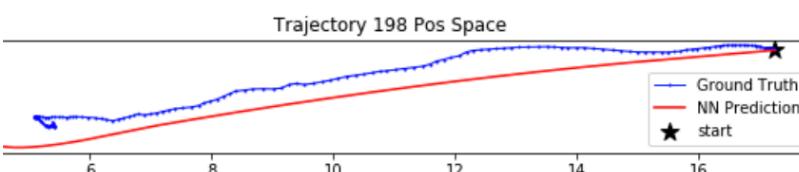
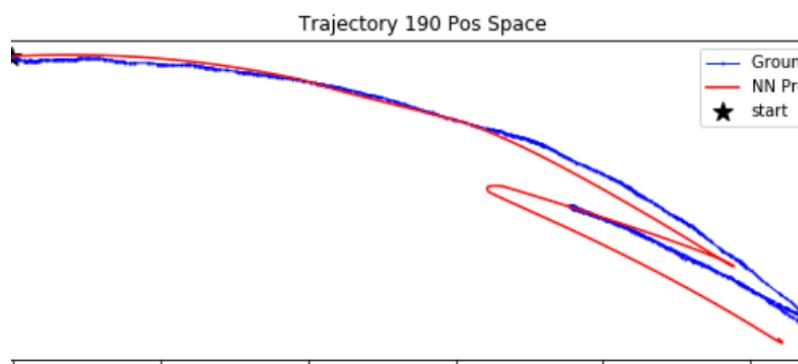
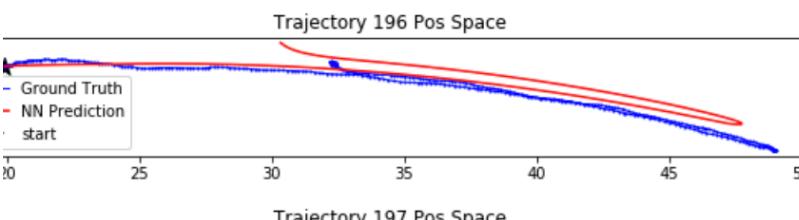
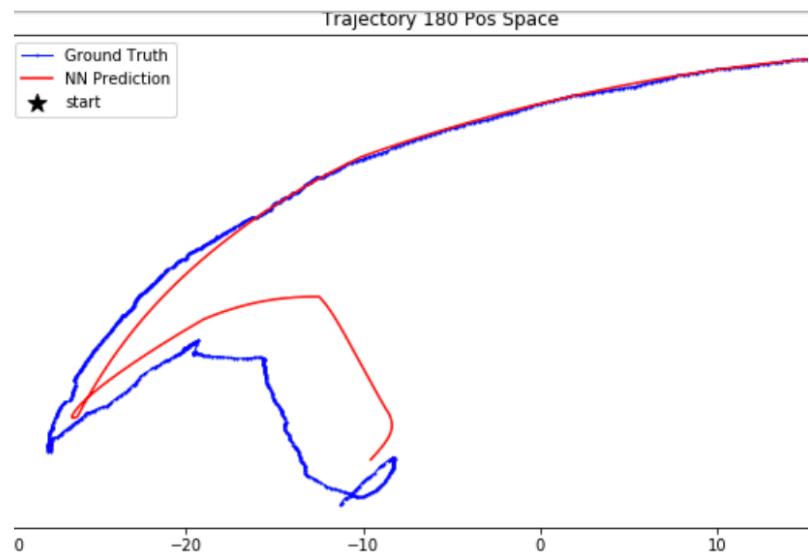


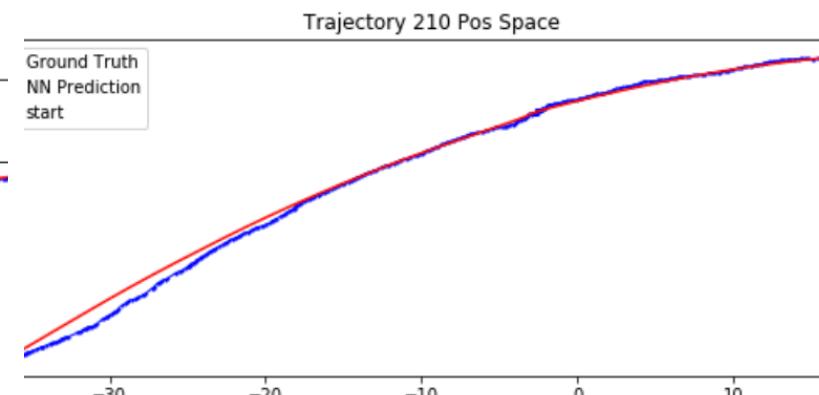
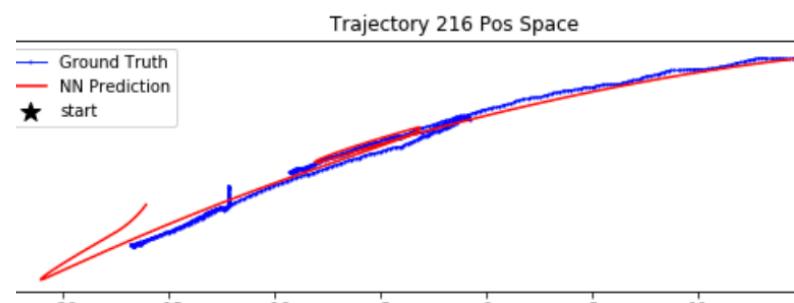
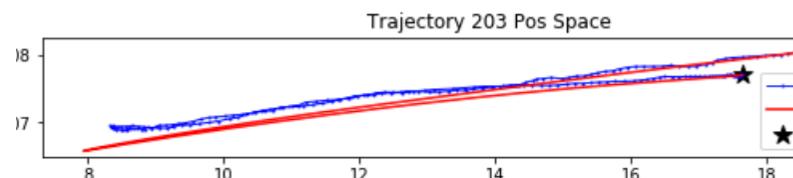
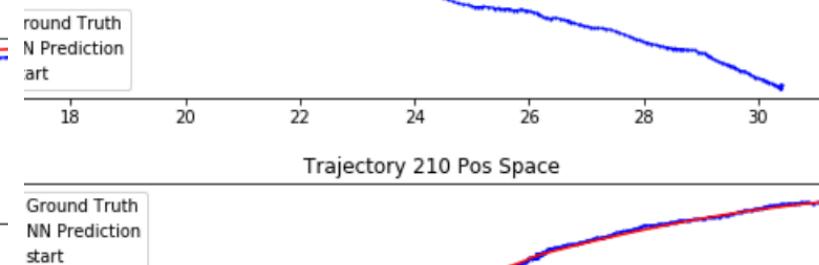
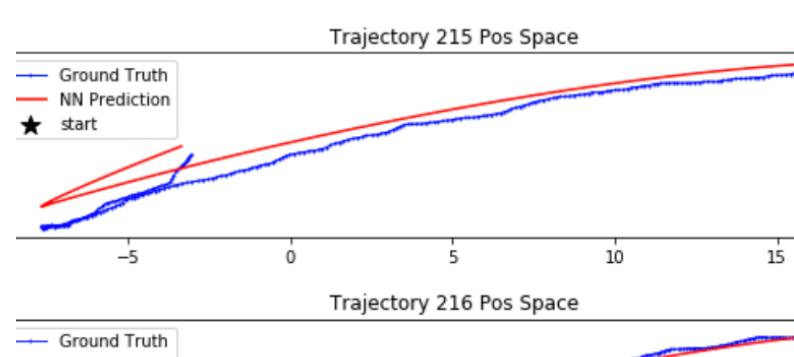
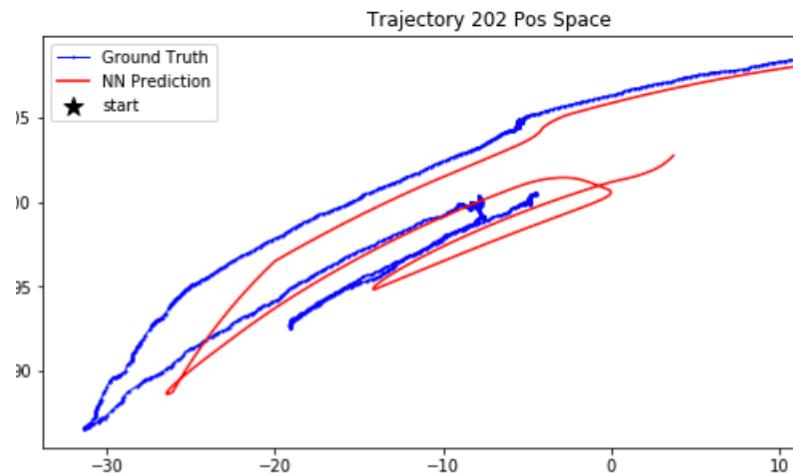
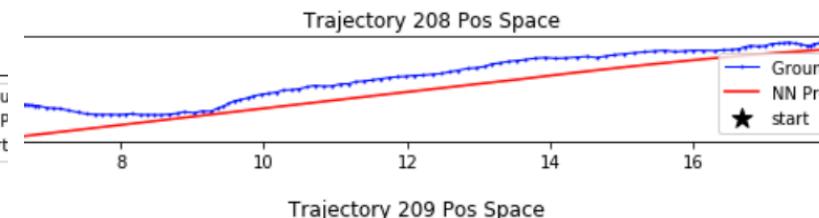
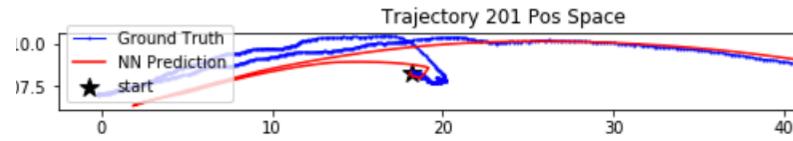
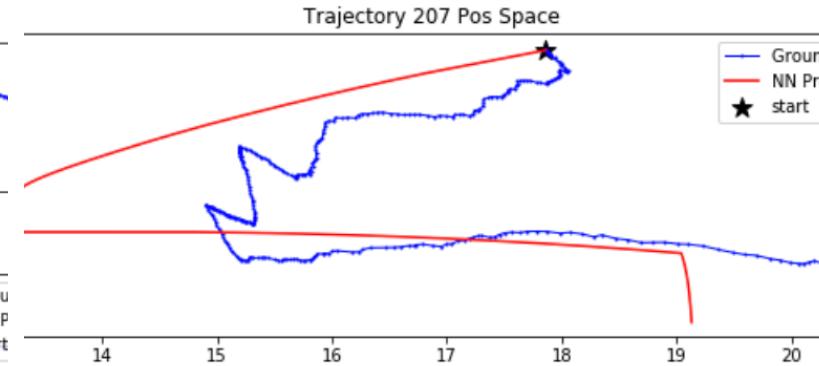
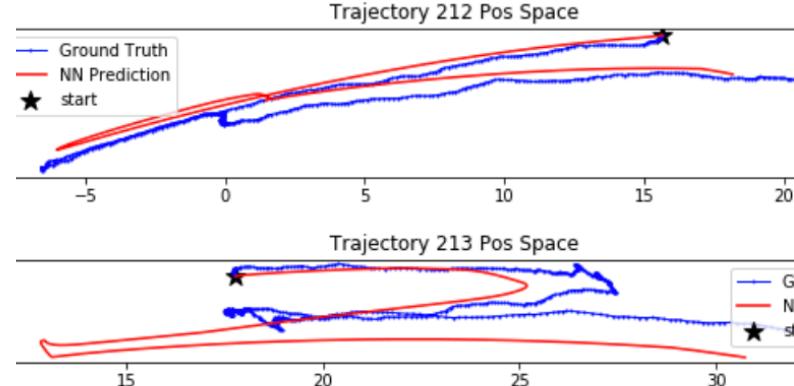
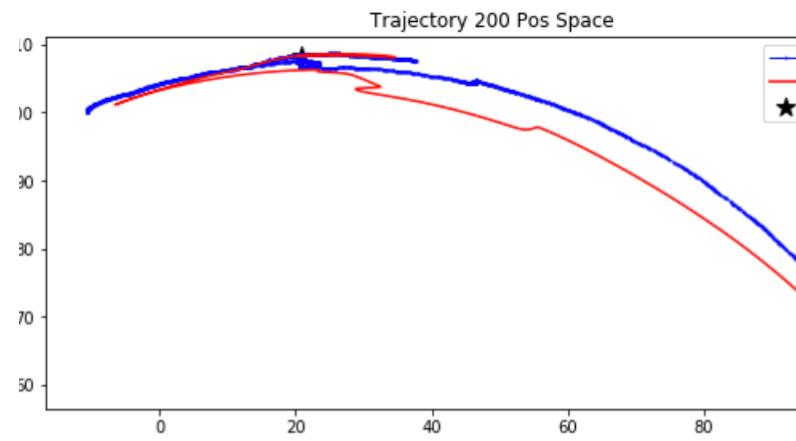


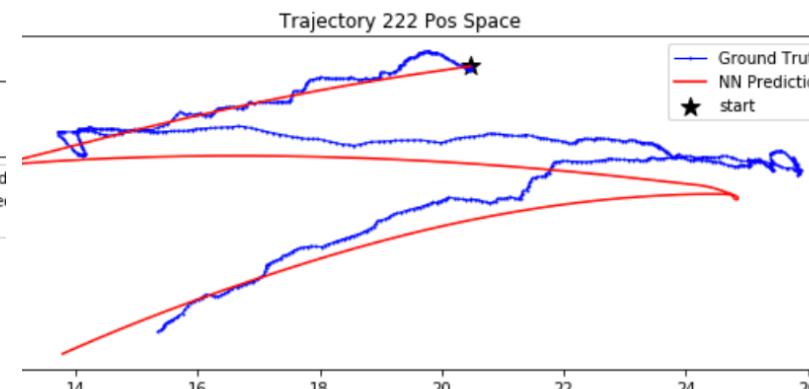
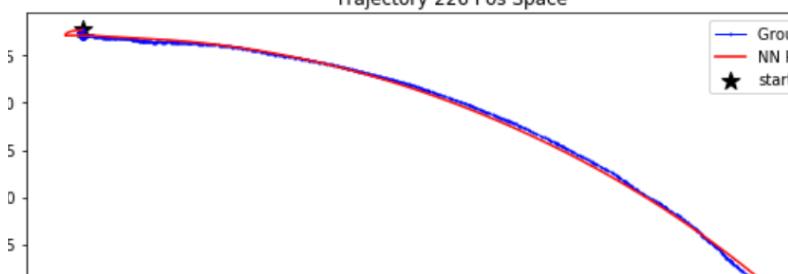
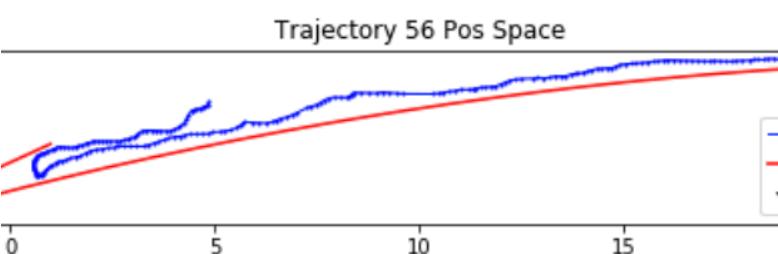
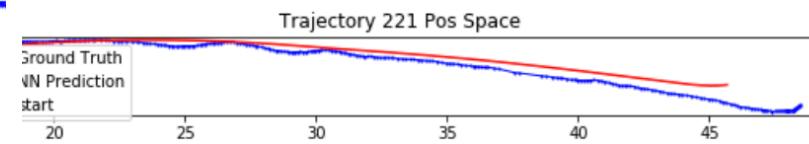
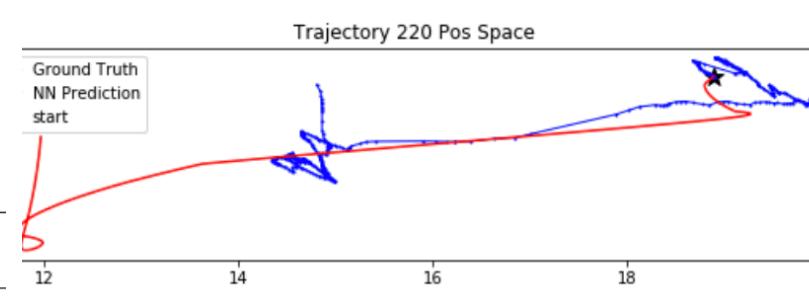
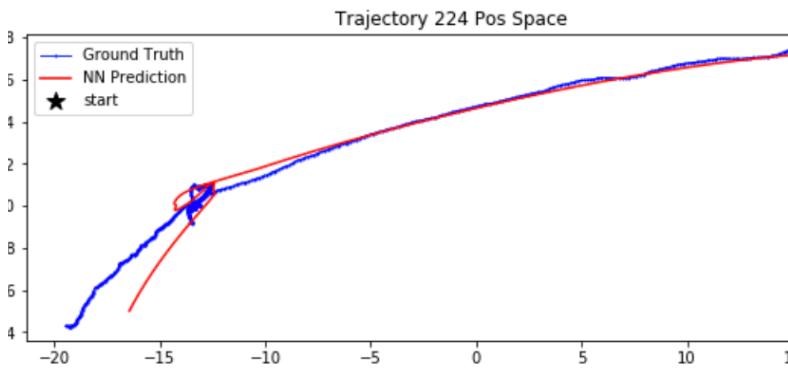
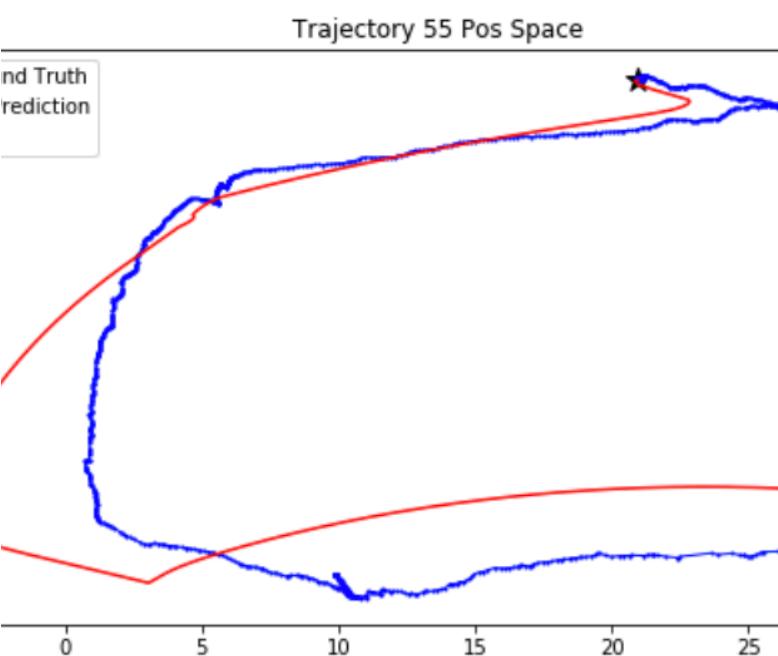
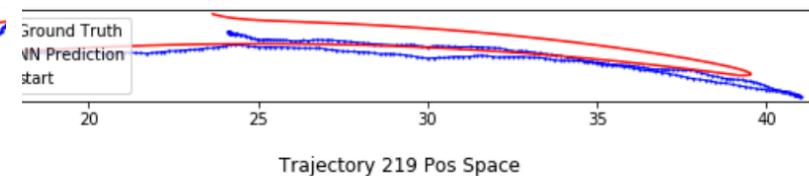
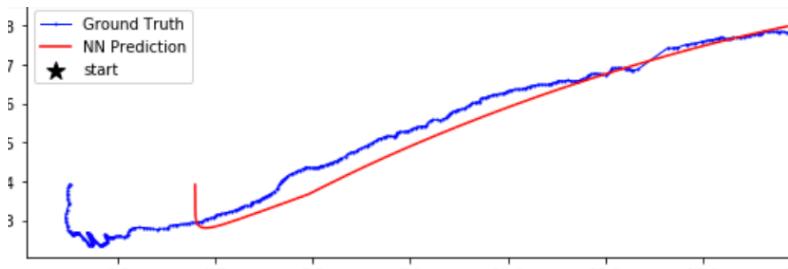
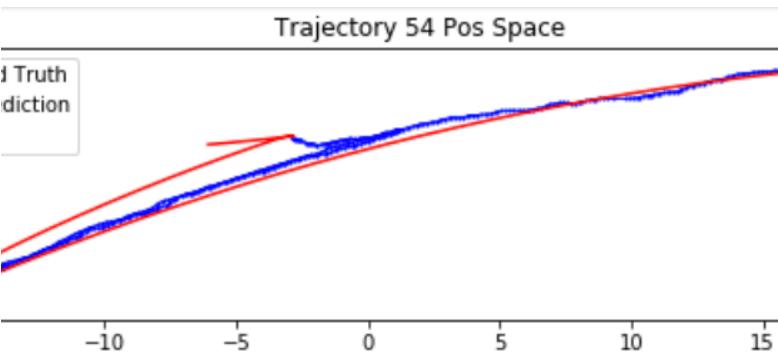








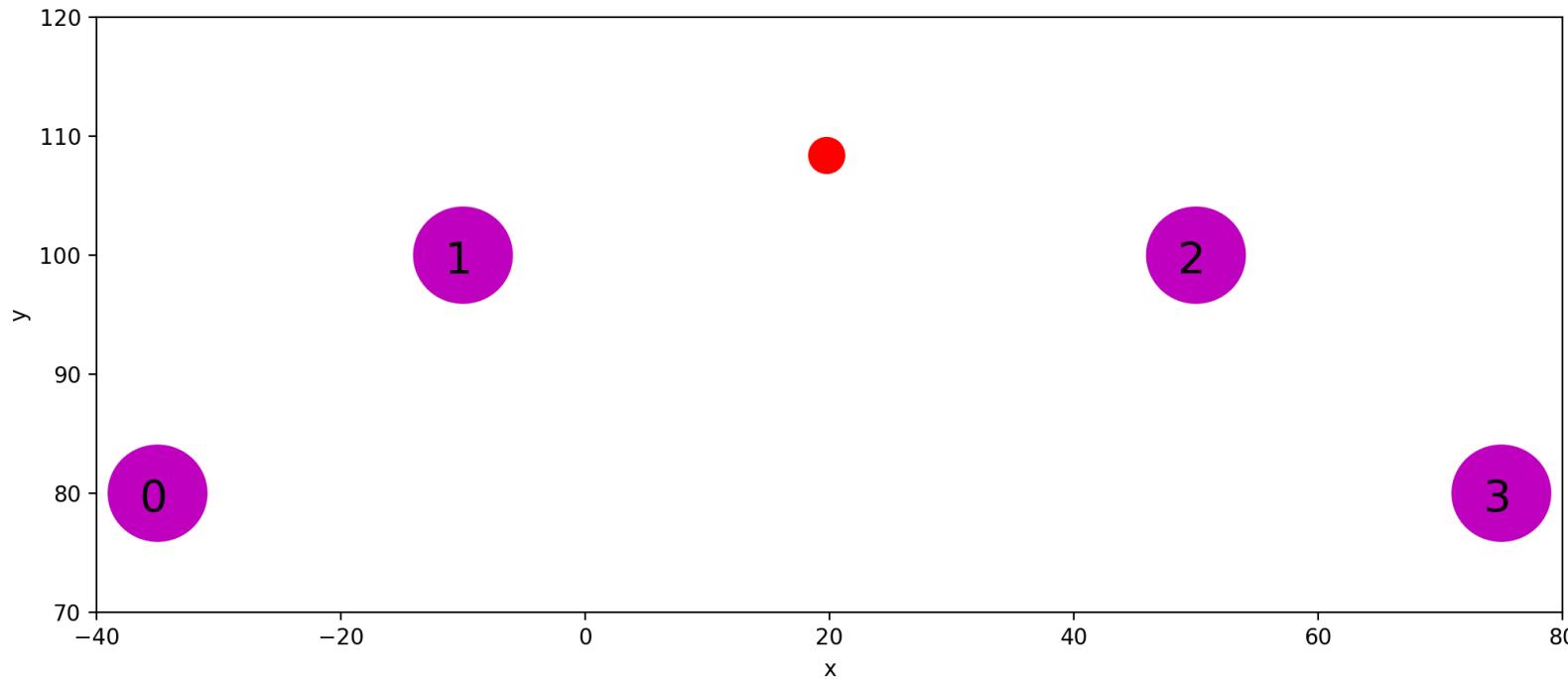




# Real Hand PPO Training Issue

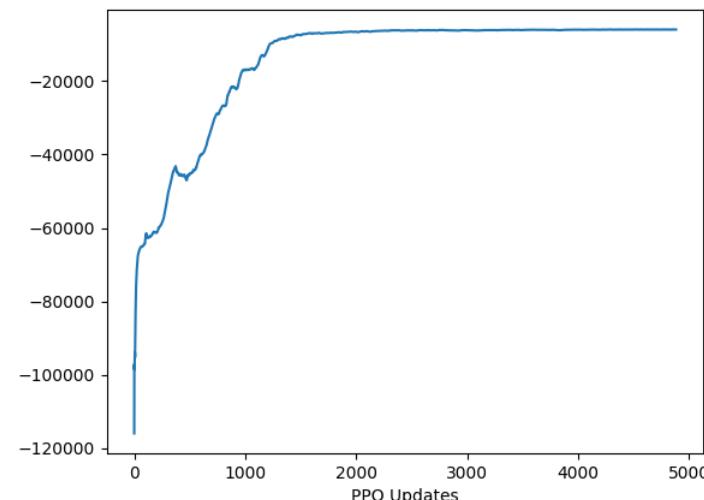
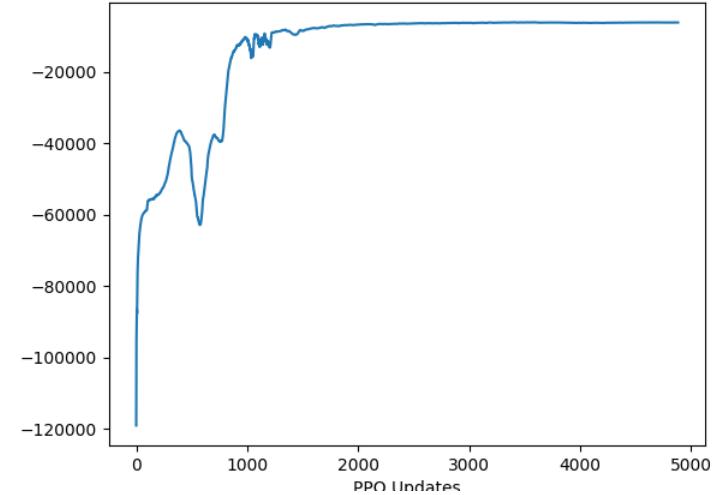
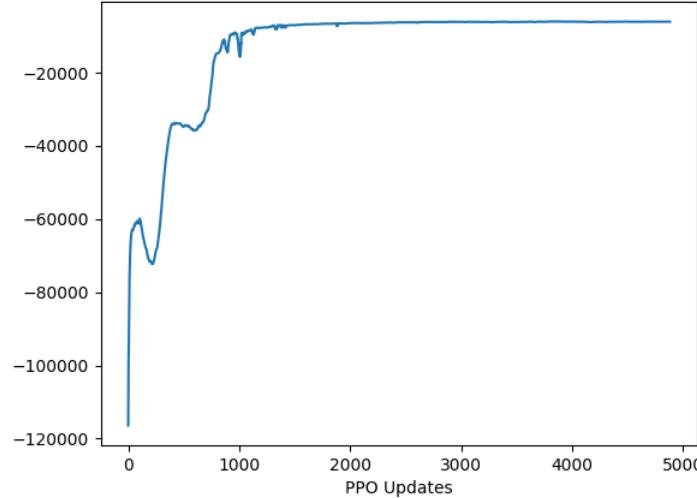
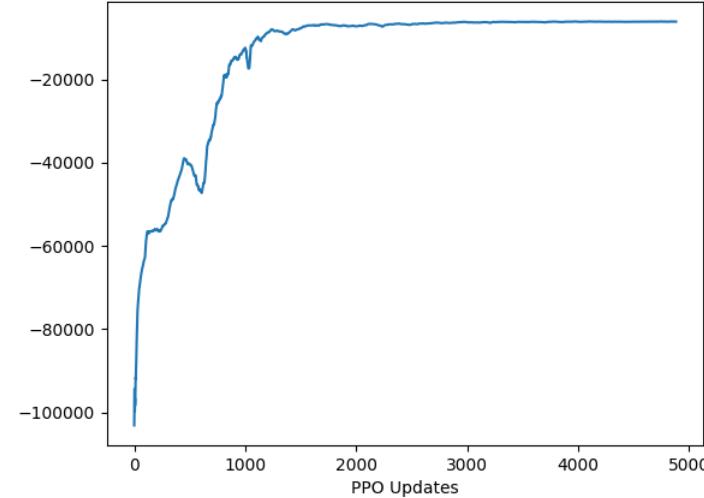
Returns are not stable.

Maybe I need to use a smaller learning rate and train more steps.

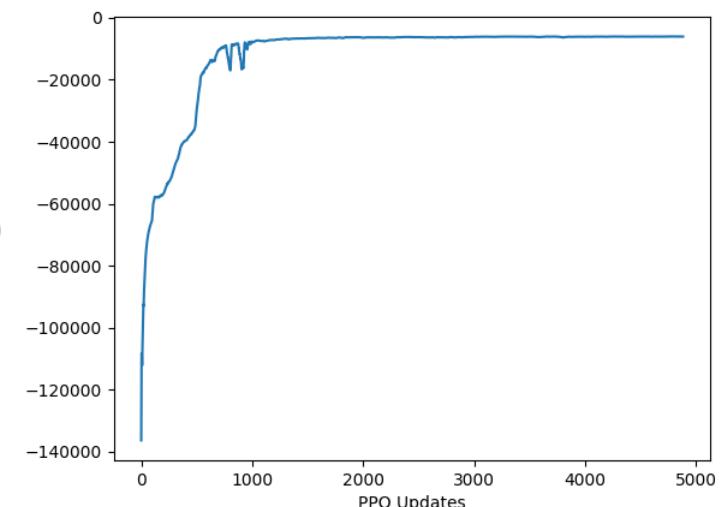


# Adaptive Hand PPO: Goal Loc 8

## learning rate: 1e-5; 10 million timesteps

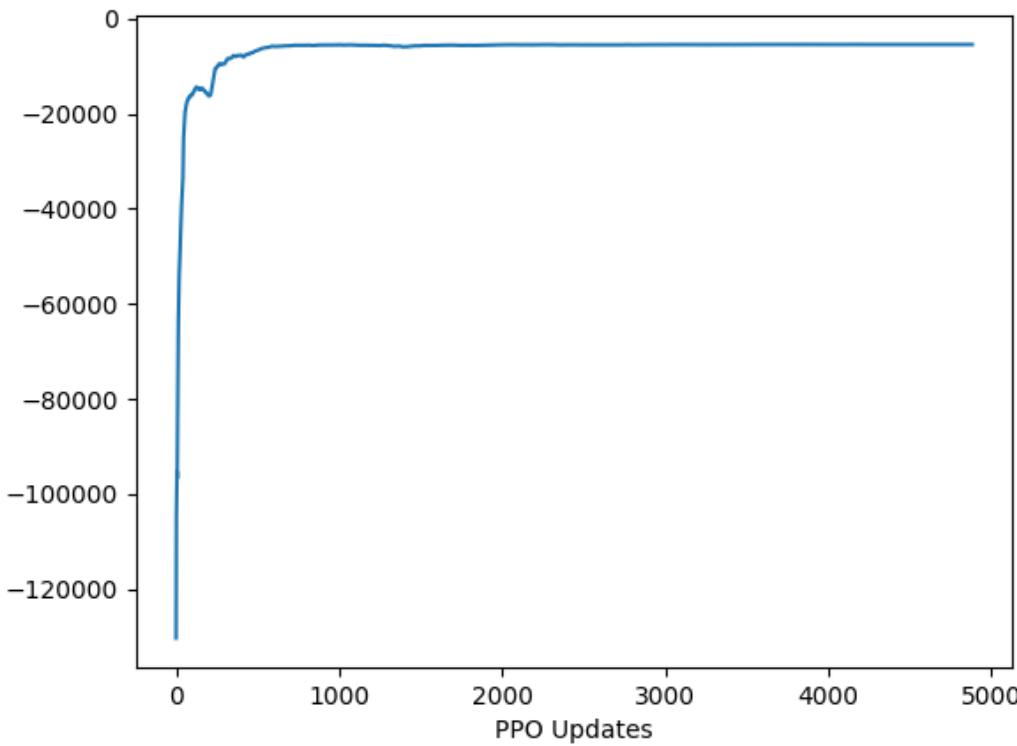


With control reward (from seed 0 to seed 4)

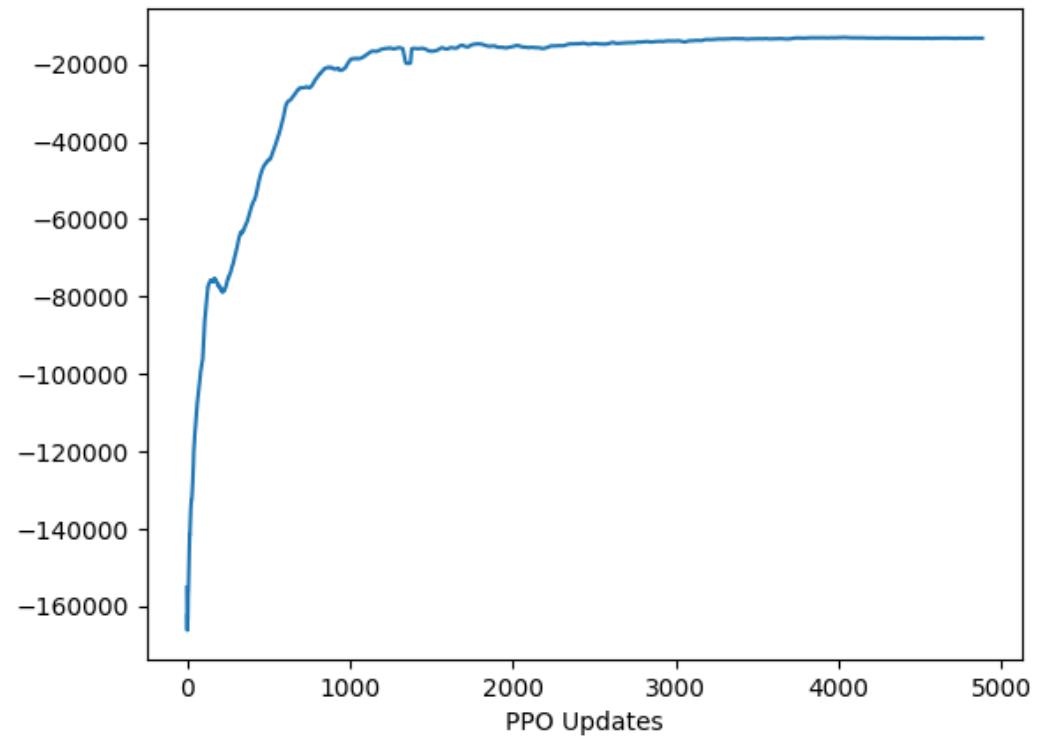


# Adaptive Hand PPO: Different Goal Locations

## learning rate: 1e-5; 10 million timesteps



With control reward: Goal Loc 7



With control reward: Goal Loc 0

# To Do List

- Fix the issues of PPO training on real hand model  
(Hyperparameter search: Learning rate, Number of train steps, etc.)
- Write a ros-package for rollouts of the real hand.  
(Avishai's previous rollout ros-package only works on the Motoman setup)
- Test whether the trained real hand dynamics model also works with Mridul's new data or not
- Do rollouts for A\* and PPO policy on real hand
- Work on our new online method

Mridul:

- Fix the issues of marker tracking
- Collect several episodes of data for testing my trained dynamics model

