<u>Reviewer #2:</u>

1. The language of the manuscript needs to be significantly optimized.

Answer : *Need your suggestions on this part sir*

2. This manuscript needs to clearly articulate the advantages of deep reinforcement learning algorithms in comparison to existing methods.

Answer :

The ship MMG Model is a control non affine system which makes difficult to design a controller such SMC, LQR and any other online adaptive learning controls. The DRL methods shows the possibility of design a control system which gives better results for control non affine system. Even though the controller such as PID, MPC could be capable to achieve the controlling tasks, these algorithms depend on system parameters highly. The PID controller for KVLCC2 will behave differently for the ship's full load condition and no-load condition. The computation cost for the MPC and the extension to multi action control makes it challenging in cost function design. Meanwhile DRL algorithms are model free control techniques. Due to the advantage of neural network's nonlinear learning character, the DRL algorithms will give optimal control irrespective of system properties when we train the DRL agent with the different system properties (such as different mass, different wave heading angle, propeller speed, etc.,)

3. It is important to provide a clear motivation for adopting four deep reinforcement learning methods to achieve path following.

Answer :

The motivation behind choosing the different RL algorithms is to investigate the robustness of ON policy (PPO), OFF policy algorithms (DDPG, TD3) and Entropy minimizing algorithm (SAC). PPO is an ON policy algorithm which will learn from the possible sample trajectory taken during the episode. DDPG, TD3 and SAC are the OFF-policy algorithms. These algorithms store the transition details in memory buffer. Form the stored memory, the agent will learn about the optimality. DPPG is using the combination of ON and OFF policy methods to learn. To the extend, TD3 will give an additional exploration to the agent in terms of estimating the value functions. Finally, SAC is an algorithm which learn about the environment from minimizing the entropy. Form the results, the inference is made that the OFF policy algorithms are working comparatively better than ON policy algorithm.

4. Regarding the innovation and contribution of the manuscript, it is important to go beyond a mere implementation of different algorithms and provide a clear and innovative approach to the problem.

Answer :

A major part of the innovation of the paper lies in the environment construction, reward function, designing of the agent's input and output vectors. selection of hyper parameters such as episode length, termination criteria of an episode. In agent construction as well, it requires meticulous attention in the hyper parameter selection.

5. Try to cite the published official journal as much as possible in the reference section.
I will try to do it sir. may I cite the Non-DRL papers?

Answer: --- ?

6. The result figures should be arranged before the references

Answer :

I have arranged it with different colors and annotation.

Reviewer #3:

1. The letters in Figure 1 should be explained in the caption

Answer :

I have done it.

2. Subplots in Figure 7, 8 and 17, 18 should be labelled with different methods. Right now, it is difficult to tell which plot corresponds to the method describing in the text.

Answer :

I have done it with different colors and annotation.

3. In introduction, the last paragraph "This paper is divided into the following sections. Section 2 briefly describes the nonlinear ship dynamics and wave disturbances. The reward function and the guidance algorithm is mentioned in section 3. The neural network structure and the RL environment are discussed in section 4. Section 5 briefly discuss the four different DRL algorithms used in this paper. Path following results in calm water and waves are discussed in section 6. Finally, the conclusions are drawn in section 7." , -> "...The neural network structure and the four different DRL algorithms used in this paper are discussed in Section 4. Path following results in calm water and waves are discussed in Section 5. Section 6 presents the hardware in-loop simulation results... Finally, the conclusions are drawn in Section 7."

Answer :

I don't understand this part.

4. "Eqn" -> "Eq." throughout the paper

Answer :

I have changed it.

5. When describing the equations, it's better to punctuate it and explain the letters in the equation later, e.g., "where xxx stands for...". This applies for most equations in the paper.

Answer :

This we already had done it. I've included few more as well.

6. The "Acknowledge" section is empty.

Answer :

I have no idea about this part sir.