POAC'25
St. John's, Newfoundland and Labrador, Canada

Proceedings of the 28th International Conference on Port and Ocean Engineering under Arctic Conditions
Jul 13-17, 2025
St. John's, Newfoundland and Labrador
Canada

Enhancing Nonlinear Model Predictive Controller for Autonomous Surface Vessels Using Reinforcement Learning: A Trajectory Tracking Study

Charuka Amarappulige[1], Eranga Fernando[1], Kevin Murrant[2], Syed Imtiaz[1], Salim Ahmed[1], Shameem Islam[2], Hasanat Zaman[2]
[1] Department of Process Engineering, Memorial University of Newfoundland, St. John's, NL, Canada
[2] Ocean, Coastal & River Engineering, National Research Council of Canada, St. John's, NL, Canada

ABSTRACT

The advancement of autonomous surface vessel (ASV) technology emphasizes the need for precise and reliable trajectory tracking, often achieved through Nonlinear Model Predictive Controllers (NMPC). However, NMPC performance depends on accurate vessel models, which are difficult to estimate due to marine environment uncertainties. This study explores integrating Reinforcement Learning (RL) with NMPC to reduce reliance on precise modeling. The approach involved developing an NMPC for a test vessel and designing a parametric approximation RL algorithm to correct state between actual and predicted outcomes. RL-NMPC performance was tested in both online and offline training scenarios. Experimental results demonstrated that the RL-enhanced NMPC effectively reduced state errors through online learning. The challenges faced by online learning due to the limitation of test facility was overcome by the introduction of simulation-based training. The final offline trained controller managed to reduce the state errors by 12% with respect to the conventional controller in the test track. However, the system's robustness against external disturbances was not evaluated in this study, leaving room for further investigation. The findings highlight RL's potential to improve NMPC usability and lay the foundation for future studies in complex marine conditions.

KEY WORDS: Nonlinear model predictive controller (NMPC); Reinforcement Learning (RL); Autonomous Surface Vessel (ASV)

INTRODUCTION

The research on autonomous ship navigation has taken significant interest due to its potential for marine surveillance, reduce accidents, enhance performance in harsh sea conditions, and improve overall operational efficiency. However, complexity in marine environment involving factors such as waves, currents, winds, and ice brings major challenge to achieving this objective. Researchers are continuously working to address these barriers and achieve successful autonomy to marine transportation.

Ship maneuvering can be categorized into two high level segments: controllers designed for low-speed maneuvers and position-keeping such as dynamic positioning and controllers designed for high-speed operations and trajectory tracking. This paper focuses predominantly on the trajectory tracking portion of autonomous surface vessels. According to Fossen (1994) trajectory tracking involves three interconnected components: guidance, navigation, and control. Guidance is related to generating desired parameters, such as the position, heading, and velocities that the vessel need to achieve. Navigation involves determining the vessel's current parameters relative to these desired states. Control portion consist of taking necessary actions to bridge the gap between the current and desired states. The primary emphasis of this work is to identify and design a suitable controller to enable precise trajectory tracking for autonomous surface vessel.

Control systems for autonomous surface vessels (ASVs) can be designed using various approaches. Classical control theory, fuzzy logic, neural networks and model predictive controllers are named few. Classical control methods, such as Proportional–integral–derivative (PID) controllers, has been one approach due to its simplicity and effectiveness in control industry. Moradi and Katebi (2001) proposed a predictive PID controller for ship autopilot design, followed by robustness and adaptability check in marine environments. Wang et al. (2019) developed a control system for ships using a modified PID algorithm, showing improved stability and responsiveness. However, PID controllers had its own limitations such as handling nonlinearity of the models and external disturbances. Majid and Arshad (2015) presented a fuzzy self-adaptive PID tracking controller for ASVs, achieving enhanced trajectory tracking performance. Rae et al. (1993) utilized fuzzy rule-based techniques for docking autonomous underwater vehicles. In recent literature, application of neural network-based controllers caught attention of many. Cheng et al. (2020) explored neural network-based control for underactuated surface vessels, ensuring transient performance through experimental validation. The dependency on the training data and adapting to situation beyond normal operating conditions has always been a challenging task for Artificial Neural Network (ANN) which is quite common in complex ocean environment.

Model predictive controller (MPC) in another well-established controller mechanism for autonomous surface vessels with good capability to handle nonlinearities and disturbances. Alagili et al. (2024) showed ability of nonlinear model predictive controller (NMPC) performance of Dynamic Positioning (DP) system under different sea conditions. Islam et al. (2023) presented offset free tracking control for both disturbed and disturbance free environment to maintain trajectory tracking accuracy. However, these approaches are heavily dependent on the accuracy of the model used. The ability of obtaining accurate vessel model has always been a difficult task due to the complexities of ocean dynamics. Martinsen et al. (2022) demonstrated combination of both Reinforcement Learning (RL) and NMPC for overcome the model inaccuracies and time-varying disturbances in DP operations.

This paper focuses on developing NMPC for a model ship and using the RL to address the mathematical model inaccuracies via parameter tuning for trajectory tracking operation. The major objective of above activity is to reduce model vessel's actual and predicted states. Then, the developed method validated with model ship in Ocean Engineering Basin of National Research Council of Canada. The rest of the article is formulated as follows. The mathematical modeling of vessel, development of NMPC and RL algorithm are discussed in the section 2. Experimental setup results and comparison is summarized in section 3. Future works and conclusion are described in finally in section 4.

MATERIAL AND METHODS

This work aims to design and implement a controller for a model ship with the help of reinforcement learning (RL) to enhance its performance. Furthermore, this paper is being used to address the real-time implementation challenges faced as well. The methodology used to achieve this goal consists of several steps. Initially the approximated mathematical model of the ship is estimated to serve as the model. Next, a NMPC is developed for trajectory tracking requirements. Then, the controller performance will be validated in simulation-based testing. Parallelly, the RL algorithm is introduced to the controller to minimize predicted and actual state errors. As per the next step the simulation-based testing will be conducted to evaluate its performance. Finally, the algorithm is tested on the actual model ship for performance. Simultaneously, this paper is used to address practical implementation issues such as computational constraints and hardware limitations. This approach ensures safe and reliable implementation while addressing both theoretical and practical challenges. Materials used in this study were discussed in subsequent sections.

Vessel Modeling

The vessel particulars have been extracted from the People Supply Vessel (PSV) model ship (1:45), which is a testing model in the offshore engineering basin research facility of the National Research Council Canada. Vessel particulars are attached in Table 1. The vessel consists of two main propellers and two tunnel thrusters. It is considered symmetrical along the keel and the same torque is applied to both symmetrical pairs when there are no disturbances in the system.

Table 1. PSV model ship particulars

Particulars	Full Scale	Unit	Model Scale	Unit
Length, OA	87.39	m	1.942	m
Length, WL	84.285	m	1.873	m
Beam	19.98	m	0.444	m
Draft, Mean	6.5	m	0.1444	m
Displacement	7737.328	Tonne	82.838	Kg

According to Fossen (2011), marine vessels consist of 6 Degrees of Freedom (DOF), as illustrated with the PSV model ship in Figure 1. Since roll, pitch and heave motions are not considered as controllable maneuvers. In order to simplify the approach, we have taken the main 3 DOF, namely, surge, sway and yaw into consideration.

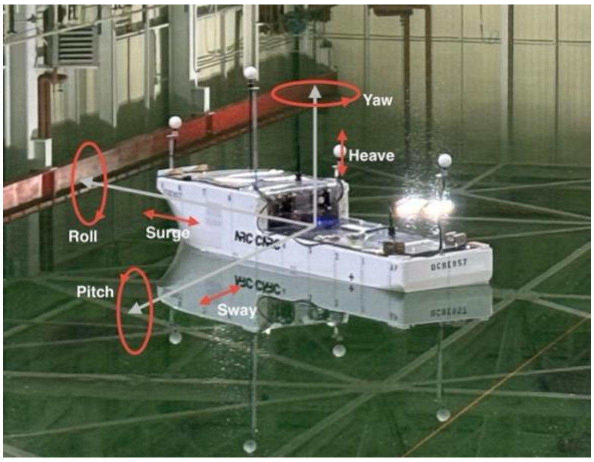

Figure 1. Mariane vessel motions

Kinematics and dynamics developed according to Fossen (2011) keeping the earth based coordinates with heading angle and ship body based velocities as states. Notations are further described in Figure 2. The earth-based position vector denoted as $\eta = [x \quad y \quad \psi]^T$ and ship body based velocity vector denoted as $\upsilon = [u \quad v \quad r]^T$. The motion equations are further described in (1) and (2)

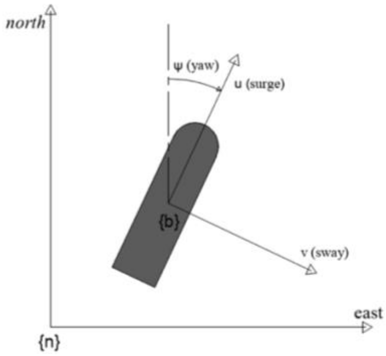

Figure 2. Earth based {n} and ship body-based {b} notations

$$\dot{\eta} = J(\psi)\nu \tag{1}$$

$$M\dot{\nu}_r + C(\nu_r)\nu_r + D(\nu_r)\nu_r = \tau_c + \tau_{env} \tag{2}$$

$$J(\psi) = \begin{bmatrix} \cos(\psi) & -\sin(\psi) & 0 \\ \sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

$M \in R^{3\times3}, C(\nu) \in R^{3\times3}, D(\nu) \in R^{3\times3}$ are estimated by system identification procedure described as in (Alagili et al., 2024). The $J(\psi)$ acts as the transformation matrix from ship body based coordinates to earth based coordination. The $\tau_c \in R^3$ and $\tau_{env} \in R^3$ are forces exerted towards the ship in the main DOF via thrusters and environment.

Nonlinear Model Predictive Controller (NMPC)

Initially, the controller is developed as a constrained open-loop optimal control problem in a MATLAB environment. The state of the vessel is constructed as a combination of earth-based position and ship body-based velocities $X = [\eta^T \quad \upsilon^T]^T$. The fmincon interior-point algorithm in

MATLAB is being used with fixed size moving horizon (Jayasiri et al., 2017) to solve the objective function in equation (4). The Q, R and P are tuning parameters in this equation and IC and X_T are constant and trajectory respectively. U_T is the control input generated at time T.

$$J = \underset{X,U}{argmin}\ IC + \sum_{k=0}^{N-1} Q(X_k - X_{kT})^2 + R(X_N - X_{NT})^2 + \sum_{k=0}^{N-1} P(U_k - U_{k+1})^2 \tag{4}$$

$$s.t. X_k = f(\mathrm{X}_{\mathrm{k-1}}, U_{\mathrm{k-1}})$$

Trajectory Generation

Trajectory points are generated by discretizing the continuous path for a given speed and time interval for motion planning. Initially, the trajectory is parameterized to generate x and y coordinates and heading to outline the trajectory. This activity is followed by computing the velocity profiles by taking numerical derivatives of the position data, ensuring smooth transitions between states. Then, the velocities are transformed to ship body based velocities to provide complete state representation. In this study, we have selected the figure-8 shaped trajectory due to its availability of the entire set of motions of the vessel and its ability to iterate multiple cycles in training episodes.

Reinforcement Learning Algorithm

Reinforcement learning (RL) can be categorized as a learning technique which learns what to do according to the situation in order to maximize the numerical reward. Here, we work on a semi-gradient Sarsa algorithm with parametric approximation of the action-value function $\hat{q}(X, U, W) \approx q_*(X, U)$, where $W \in R^2$ is a finite dimensional weight vector (Sutton and Barto, 2018). Given the optimization problem, we define the parametric action value function as equation 5 and the update for one-step Sarsa as in equation 6. The pseudo code for the learning algorithm is given in the below box.

$$\hat{q}(X, U, W) = \underset{X,U}{min}\ w_1 + \sum_{k=0}^{N-1} Q(X_k - X_{kT})^2 + w_2(X_N - X_{NT})^2 + \sum_{k=0}^{N-1} P(U_k - U_{k+1})^2 \tag{5}$$

$$W_{t+1} = W_t + \alpha[R_{t+1} + \gamma\, \hat{q}(X_{t+1}, U_{t+1}, W_t) - \hat{q}(X_t, U_t, W_t)]\, \nabla\hat{q}(X_{t+1}, U_{t+1}, W_t) \tag{6}$$

Algorithm
Initialize state X, and action U
Loop for each step:
Choose next action, U' as a function of $\hat{q}(\mathrm{X, U, W})$ (greedy)
Reward $= \mathrm{K} \times \lVert \mathrm{X_{actual}} - \mathrm{X_{predicted}} \rVert$
$\delta \leftarrow$ Reward $+ \gamma\hat{q}(\mathrm{X', U', W}) - \hat{q}(\mathrm{X, U, W})$
$\mathrm{W} \leftarrow \mathrm{W} + \alpha\, \delta \boldsymbol{\nabla} \hat{q}(\mathrm{X, U, W})$
$\mathrm{X, U} \leftarrow \mathrm{X', U'}$

RESULTS AND DISCUSSION

The experiment was carried out in the Ocean Engineering Basin of the National Research Council in St John's with the PSV hull. The basin is 55m in length and 25m in width. It consists of a Qualisys motion capture system to measure earth based position and vessel velocities. The active markers are placed on the vessel, allowing the motion capture system to identify its position and orientation. The basin and ship before test are shown below in Figure 3.

Figure 3. Ocean Engineering Basin and PSV model before testing

The objective of the experiment is to evaluate the performance of improvement of the controller after incorporating the reinforcement learning algorithm into the cost function. The figure-8 motion was selected as a test track with an 8m radius in each loop for the evaluation. The vessel was set to maneuver around the shape at a speed of 0.2m/s. Trajectory points were generated for the test track as described in the trajectory generation chapter. The initial test was carried out on conventional NMPC and NMPC+RL with online learning. The results are shown in Figure 4 and Figure 5 respectively. The predicted and actual state error for each time step is illustrated in Figure 6 and Figure 7. In the figures state errors for positions calculated in meters while angle calculated in radians, similarly speeds states calculated m/s and rad/s respectively. However, the test could not be completed full path due to the signal drop-out zones in the setup. This hinders the overall objective of performing the online RL training opportunity by conducting iterative test runs on the track. The list of challenges faced, the impact on the test, and remedies taken are summarized in the Table 2.

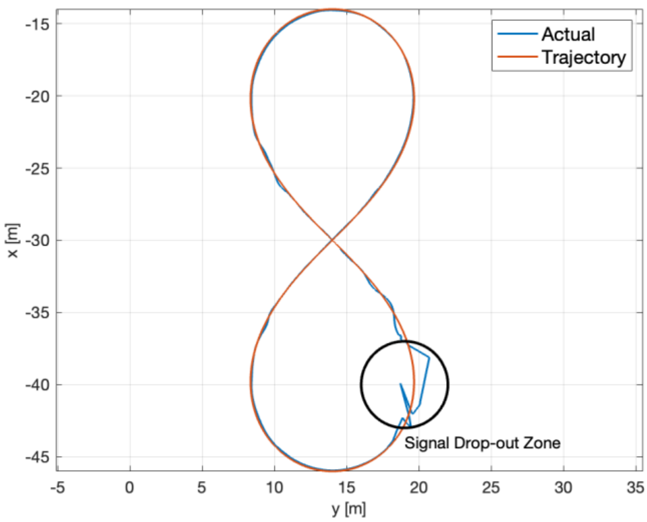

Figure 4. Conventional NMPC performing figure-8 motion

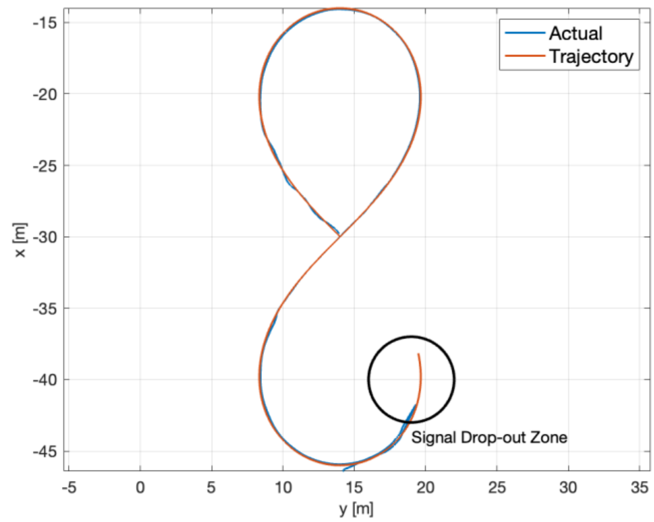

Figure 5. NMPC+RL performing figure-8 motion

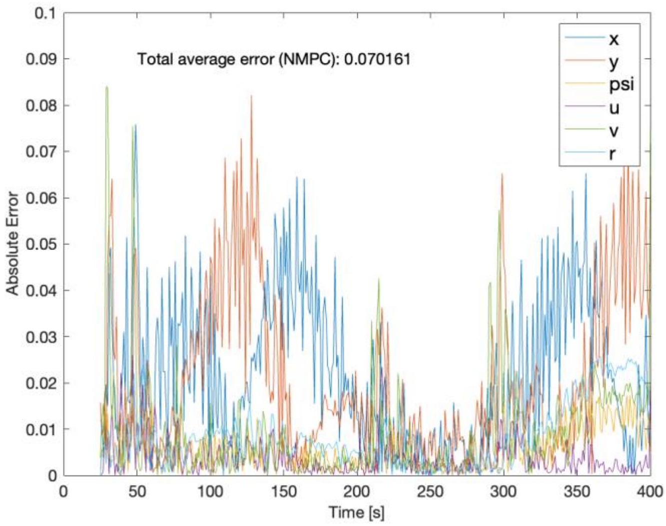

Figure 6. Predicted and actual state error for NMPC

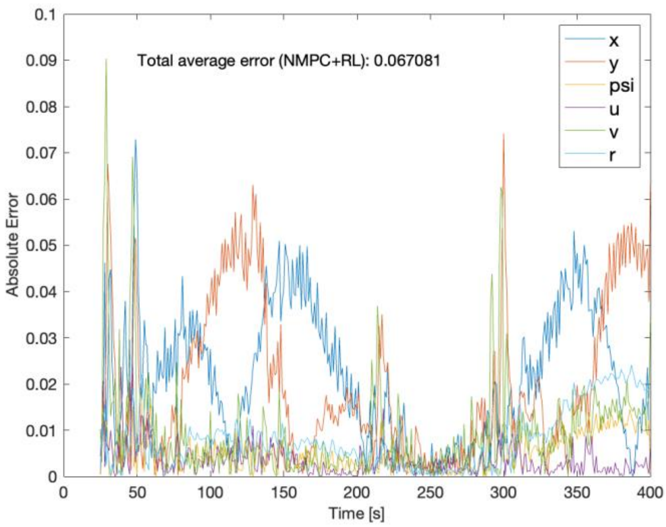

Figure 7. Predicted and actual state error for NMPC+RL

Table 2. Experiment related challenges and action taken

Challenge	Impact	Remedy
Signal dropout zones	Could not completed full figure-8 motion	Adaptation of offline iterative training
Inconsistency in starting point	Incomparable initial states	Exclude initial 25 timesteps from comparison
Asymmetry in thruster allocation	Difficulty in comparing asymmetric maneuvers	Avoid asymmetric maneuvers
Limited time availability of testing facility	Could not train the model with real vessel	Adaptation of offline training
Inconsistency in environmental condition	Change in outcomes in different days	Average out results of multiple days

In order to overcome signal dropout zones and limited time availability challenges, an offline training methodology was introduced. The purpose of the offline training was to mimic the ship's travel around the test track for multiple cycles and learn the optimum parameters in a simulated environment before launching the actual vessel on the basin for testing. The offline training is carried out in a simulation for 20,000 steps, which is equivalent to 40 cycles in a similar path. This session was conducted in two segments, one learning without noise and one with randomly generated noise included in the states and actions. Then, the trained parameters are fed into the experiment setup for actual performance evaluation. The vessel performance of the offline trained controller performing figure-8 is illustrated in Figure 8. The comparison of the total prediction and actual state error with respect to conventional NMPC controller average out for multiple runs is shown in Figure 9. Furthermore, the individual average state error distribution for each scenario is illustrated in Figure 10.

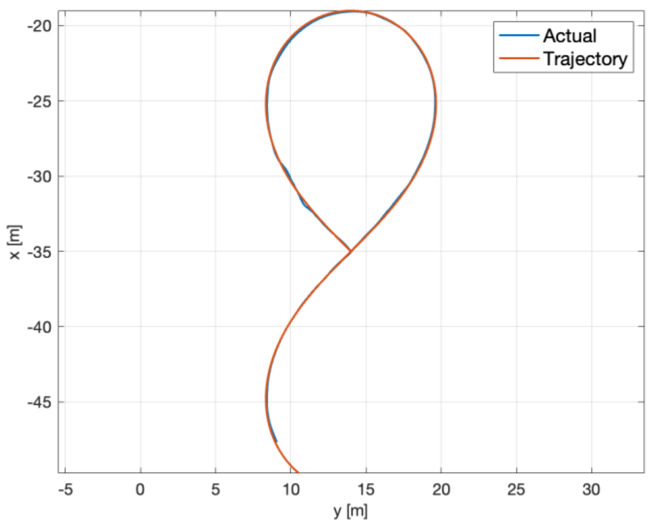

Figure 8. Offline trained (without noise) NMPC+RL performing figure-8 motion

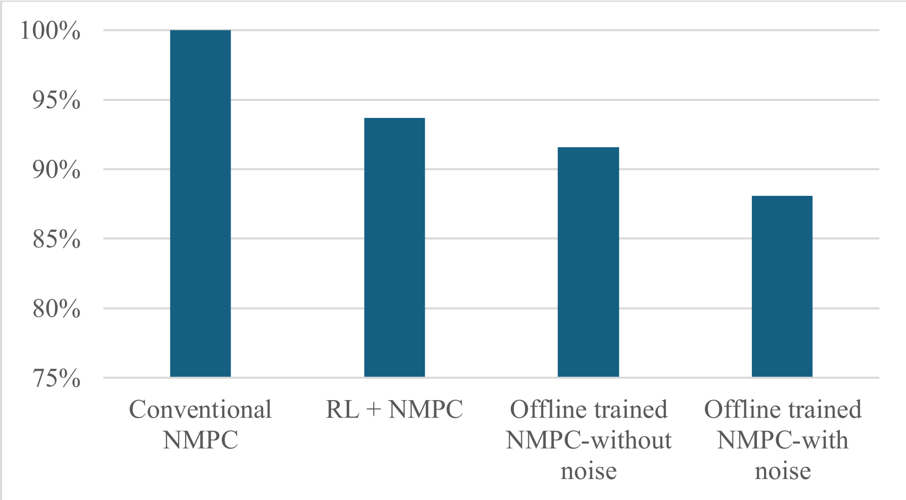

Figure 9. Total error comparison for each controller with multiple run average

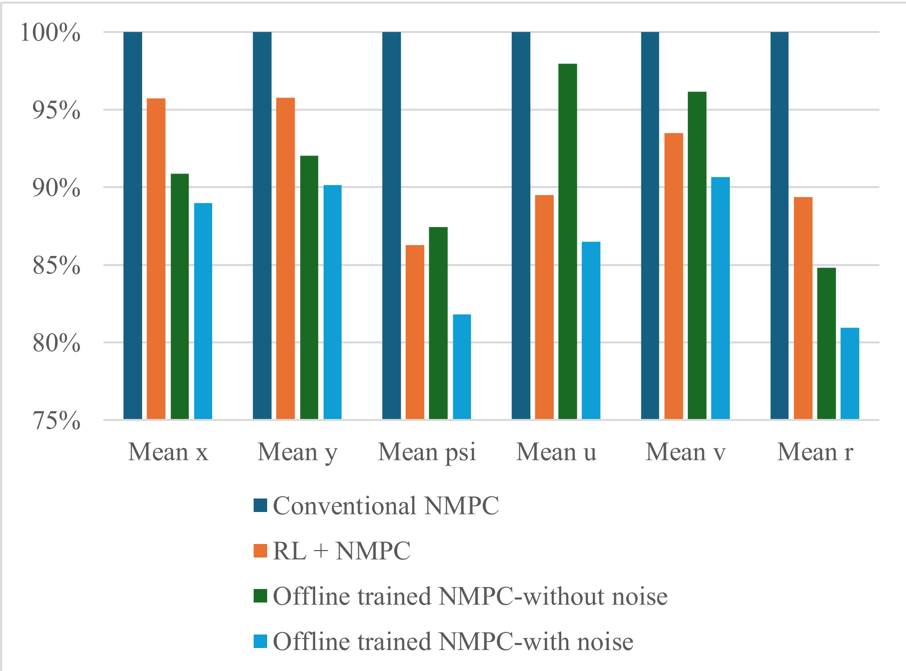

Figure 10. Total state error comparison for each controller with multiple run average

CONCLUSIONS AND FUTURE WORK

The study managed to successfully maneuver the ASV in a predefined path, overcoming real-time experimental challenges. The controller's performance was significantly improved with the usage of RL, achieving notable reduction in state errors. Additionally, offline-trained controller showed superior performance over online-learning controller which highlights the benefit of offline training.

Future work will extend to evaluate the robustness of the controller under external disturbances, such as wave and wind forces, to ensure practical applicability in dynamic maritime environments. Further research on RL or Hybrid controller to improve its performance is proposed to achieve more reliable and efficient vessel control systems

ACKNOWLEDGEMENTS

The authors sincerely express their gratitude to the National Research Council of Canada for its generous financial support provided through the Collaborative Research and Development project. This support has been instrumental in enabling the successful completion of this research.

REFERENCES

Alagili, O., Fernando, E., Ahmed, S., Imtiaz, S., Murrant, K., Gash, B., Islam, M., Zaman, H., 2024. Experimental investigations of an energy-efficient dynamic positioning controller for different sea conditions. Ocean Engineering 299, 117297. https://doi.org/10.1016/j.oceaneng.2024.117297

Cheng, X., Li, G., Skulstad, R., Chen, S., Hildre, H.P., Zhang, H., 2020. A Neural-Network-Based Sensitivity Analysis Approach for Data-Driven Modeling of Ship Motion. IEEE J. Oceanic Eng. 45, 451–461. https://doi.org/10.1109/JOE.2018.2882276

Fossen, T.I., 2011. Handbook of marine craft hydrodynamics and motion control. Chichester, West Sussex, U.K. ; Hoboken N.J. : Wiley.

Fossen, T.I., 1994. Guidance and Control of Ocean Vehicles. Wiley.

Islam, T., Imtiaz, S., Ahmed, S., Islam, M., Zaman, H., Gash, R., 2023. Offset-Free Tracking Control for a Marine Autonomous Surface Ship, in: Volume 1: Offshore Technology. Presented at the ASME 2023 42nd International Conference on Ocean, Offshore and Arctic Engineering, American Society of Mechanical Engineers, Melbourne, Australia, p. V001T01A034. https://doi.org/10.1115/OMAE2023-104936

Jayasiri, A., Nandan, A., Imtiaz, S., Spencer, D., Islam, S., Ahmed, S., 2017. Dynamic Positioning of Vessels Using a UKF-Based Observer and an NMPC-Based Controller. IEEE Transactions on Automation Science and Engineering 14, 1778–1785. https://doi.org/10.1109/TASE.2017.2698923

Majid, M.H.A., Arshad, M.R., 2015. A fuzzy self-adaptive PID tracking control of autonomous surface vehicle, in: 2015 IEEE International Conference on Control System, Computing and Engineering (ICCSCE). Presented at the 2015 IEEE International Conference on Control System, Computing and Engineering (ICCSCE), pp. 458–463. https://doi.org/10.1109/ICCSCE.2015.7482229

Martinsen, A.B., Lekkas, A.M., Gros, S., 2022. Reinforcement learning-based NMPC for tracking control of ASVs: Theory and experiments. Control Engineering Practice 120, 105024. https://doi.org/10.1016/j.conengprac.2021.105024

Moradi, M.H., Katebi, M.R., 2001. Predictive PID Control for Ship Autopilot Design. IFAC Proceedings Volumes, IFAC Conference on Control Applications in Marine Systems 2001, Glasgow, Scotland, 18-20 July 2001 34, 375–380. https://doi.org/10.1016/S1474-6670(17)35111-X

Rae, G.J.S., Smith, S.M., Anderson, D.T., Shein, A.M., 1993. A Fuzzy Rule Based Docking Procedure for Two Moving Autonomous Underwater Vehicles, in: 1993 Americ an Control Conference. Presented at the 1993 American Control Conference, pp. 580–584. https://doi.org/10.23919/ACC.1993.4792923

Sutton, R.S., Barto, A.G., 2018. Reinforcement Learning, second edition: An Introduction. MIT Press.

Wang, L., Wu, Q., Liu, J., Li, S., Negenborn, R.R., 2019. Ship Motion Control Based on AMBPS-PID Algorithm. IEEE Access 7, 183656–183671. https://doi.org/10.1109/ACCESS.2019.2960098