MECE6397: Learning-Based Control
Author: Higor Felipe de Medeiros

Optimizing Offshore Logistics: Reinforcement Learning-Based Control for Offloading Operations from Spread-Mooring FPSOs Using Conventional Oil Tankers

---

1. Introduction

Offshore oil and gas operations face critical logistical challenges when transferring crude from FPSOs—large Floating, Production, Storage and Offloading platforms positioned over 300 km offshore. Traditionally, the offloading operation in FPSOs relies on specialized Dynamic Positioning systems ("DP3") shuttle tankers, which advanced, triple-redundant set of thruster systems ensures precise station-keeping, but come with steep acquisition and operating costs and limited availability. Given that a FPSO holds only six to eight days of production and must offload roughly one million barrels every four days, these constraints create frequent bottlenecks and reduce operational flexibility.

To address this issue, we evaluate the feasibility of replacing scarce DP3 shuttle tankers with readily available conventional oil tankers (equipped solely with standard onboard components—rudder and propeller) by employing a Reinforcement Learning (RL)–based control system. In our dynamic analysis and modeling, the FPSO is assumed fixed and the tanker is the only moving body, see Figure 1; the hawser is modeled as a rigid, inelastic link, and only steady ocean currents are considered (neglecting wind and waves).

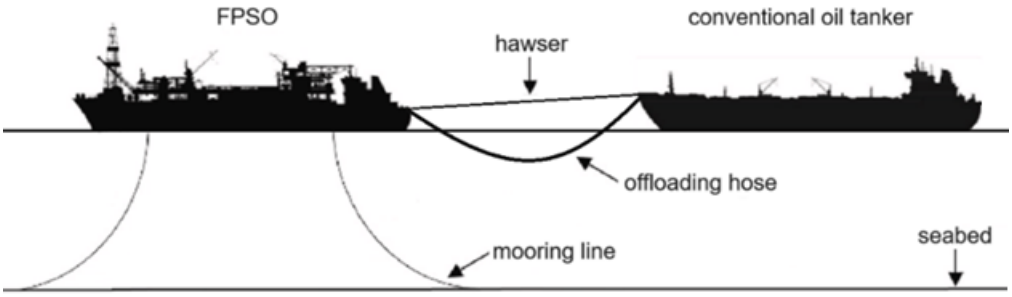

Figure 1 – Offloading with Conventional Oil Tanker

This project utilizes RL agent to correct tanker's position and heading by issuing continuous adjustments for the rudder and propeller-induced sway forces. The RL agent's reward function is to punctuate positional deviations—specifically lateral drift—so that, over training, the system learns to maintain station effectively.

2. Literature Review

To estimate current forces acting on the vessel, this work adopts the static short wing model by Leite et al. [1] , which simplifies the interaction by focusing on sway movements for vessels in steady conditions. For the fishtailing phenomenon, the methodology follows Simos et al. [2] , who applied a quasi-explicit hydrodynamic model validated with physical experiments using a VLCC model and a rigid hawser.

Propeller and rudder forces are estimated based on the modular mathematical model by Sukas et al. [3] , which predicts maneuvering performance for twin-propeller and twin-rudder ships. The development of a reinforcement learning-based control system will be guided by the foundational principles presented by Sutton and Barto [4] .

3. Physical Problem Formulation

The methodology proposed in this paper is based on the time-series results of ship heading (ψ), hawser angle (γ) and hawser force (F_h) reported in [2] , which were obtained at a model current speed of U = 0.5 m/s (equivalent to 4.74 m/s full-scale using a 90:1 scale factor).

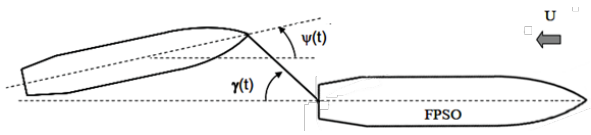

Figure 2 - Schematic representation of rigid hawser experimental setup

Using this data, the objective is to compute the actual hawser and current-induced forces, evaluate the maximum rudder and propeller responses, determine the resulting total sway forces and moments, and compare them with the vessel's inherent inertial resistance to assess whether they can alter the sway direction.

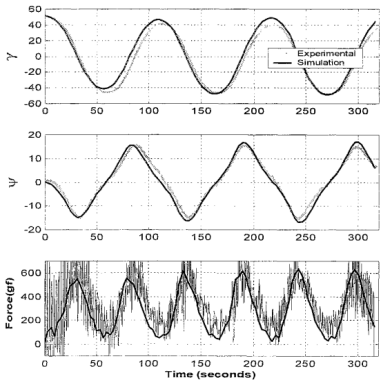

Figure 3 - Time series results [2]

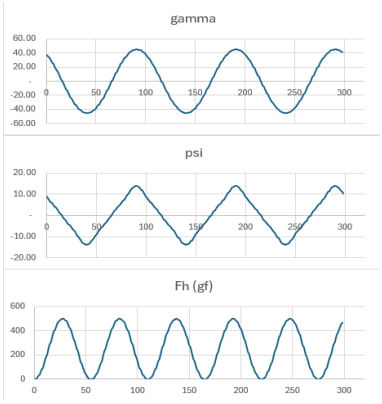

Figure 4 - Fitted functions

3.1 Motion with Rigid Hawser

The equations which represent the dynamics of the Spread Mooring System are:

$$(M+M_{11})\dot{u}-(M+M_{22})vr-(Mx_{CG}+M_{26})r^2=F_{X,R}(u;v;r)+F_X(u;v;r)$$

$$(M+M_{22})\dot{v}+(M+M_{11})ur+(Mx_{CG}+M_{26})\dot{r}=F_{Y,R}(u;v;r)+F_Y(u;v;r)$$

$$(I_Z^\circ+M_{66})\dot{r}+(Mx_{CG}+M_{26})(\dot{v}+ur)=N_{Z,R}^\circ(u;v;r)+N_Z^\circ(u;v;r)$$

The presence of the rigid hawser introduces the following kinematic relationships:

$$u(t)=U\cos\psi(t)+L_H\dot{\gamma}\,\mathrm{sen}(\gamma(t)+\psi(t))$$

$$v(t)=-U\,\mathrm{sen}\,\psi(t)+L_H\dot{\gamma}\cos(\gamma(t)+\psi(t))-x_A\dot{\psi}(t)$$

$$r(t)=\dot{\psi}(t)$$

And the forces and moments imparted by the action of the hawser can be written in terms of the tension force acting on it (F_h), in the form given by expressions:

$$F_X=F_H\cos(\gamma+\psi)$$

$$F_Y=-F_H\,\mathrm{sen}(\gamma+\psi)$$

$$N_Z^\circ=-x_AF_H\,\mathrm{sen}(\gamma+\psi)$$

The dynamic system admits a trivial equilibrium in which the vessel travels at its nominal forward speed with zero lateral and yaw velocity, aligned with the incoming current. To assess the stability of this equilibrium, the model is linearized around that operating point under the assumptions of small attack angles near zero or 180 degrees, negligible second-order lateral and yaw terms compared to products of forward and lateral motions, and very slow

variation of the forward speed. The resulting linear dynamics are presented in matrix form and can then be reformulated in terms of the two angular deviation variables.

$$\bar{F}_X = -\bar{F}_{X,R}$$

$$\begin{bmatrix} A_{22}.L_H/L & (-A_{22}\,x_A/L + A_{26}) \\ A_{26}.L_H/L & (-A_{26}\,x_A/L + A_{66}) \end{bmatrix} \frac{d^2}{d\bar{t}^2} \begin{Bmatrix} \gamma(\bar{t}) \\ \psi(\bar{t}) \end{Bmatrix} + \begin{bmatrix} 0 & (A_{11} - A_{22}) \\ 0 & 0 \end{bmatrix} \frac{d}{d\bar{t}} \begin{Bmatrix} \gamma(\bar{t}) \\ \psi(\bar{t}) \end{Bmatrix} = \begin{Bmatrix} \bar{F}_{Y,R} \\ \bar{N}^{\circ}_{Z,R} \end{Bmatrix} + \begin{Bmatrix} \bar{F}_Y \\ \bar{N}^{\circ}_Z \end{Bmatrix}$$

3.2 Ocean Current

Current-induced forces were evaluated using the static "short-wing" model of Leite et al. (1998) [1] , in which the tanker hull is idealized as a low-aspect-ratio wing in a steady, uniform current, with yaw, free-surface, and finite-depth effects all neglected.

The relative flow in the ship's axes is resolved into longitudinal (u) and transverse (v) components, from which the resultant speed Vr and angle of attack αr are computed.

$$\mathbf{U}_r(t) = u(t)\mathbf{i} + v(t)\mathbf{j}$$
$$u(t) = \dot{x}_1(t) - V_C \cos\alpha(t)$$
$$v(t) = \dot{x}_2(t) - V_C \,\mathrm{sen}\,\alpha(t)$$

Omitting slow temporal variation gives the relative current speed and incidence angle:

$$V_{cr} = \sqrt{u^2 + v^2}$$
$$\alpha_r = \pi + \arctan(v/u)$$

The static surge, sway and yaw contributions are:

$$F_{1C}(u;v;0) = \frac{1}{2}\rho V_{cr}{}^2 LT.C_{1C}(\alpha_r)$$
$$F_{2C}(u;v;0) = \frac{1}{2}\rho V_{cr}{}^2 LT.C_{2C}(\alpha_r)$$
$$F_{6C}(u;v;0) = \frac{1}{2}\rho V_{cr}{}^2 L^2 T.C_{6C}(\alpha_r)$$

The longitudinal-force coefficient combines hull resistance and a small-wing lift correction:

$$C_{1C}(\alpha_r) = C_0 \cos\alpha_r + \frac{1}{8}\frac{\pi T}{L}(\cos 3\alpha_r - \cos\alpha_r)$$

where the baseline resistance term is estimated from model-scale data by

$$C_0 \cong \frac{0{,}094}{(\log_{10}\mathrm{Re} - 2)^2}\frac{S}{TL}$$

Heuristic expressions are given for the lateral-force and yaw-moment coefficients:

$$C_{2C}(\alpha_r) = \left(C_Y - \frac{\pi T}{2L}\right)\mathrm{sen}\,\alpha_r|\mathrm{sen}\,\alpha_r| + \frac{\pi T}{2L}\mathrm{sen}^3\alpha_r + \frac{\pi T}{L}\left(1 + 0{,}4\frac{C_B B}{T}\right)\mathrm{sen}\,\alpha_r|\cos\alpha_r|$$
$$C_{6C}(\alpha_r) = -\frac{l_P}{L}C_Y\,\mathrm{sen}\,\alpha_r|\mathrm{sen}\,\alpha_r| - \frac{\pi T}{L}\mathrm{sen}\,\alpha_r\cos\alpha_r - \left(\frac{1+|\cos\alpha_r|}{2}\right)^2\frac{\pi T}{L}\left(\frac{1}{2} - 2{,}4\frac{T}{L}\right)\mathrm{sen}\,\alpha_r|\cos\alpha_r|$$

3.3 Rudder and Propeller

The lateral ("sway") force generated by the rudder–propeller system of a large tanker is predicted using the modular MMG (Maneuvering Modeling Group) framework, which decomposes total hydrodynamic forces into hull, propeller, and rudder contributions.

$$(m + m_y)\,\dot{v} + (m + m_x)\,u\,r + m_x\,x_G\,r^2 \;=\; Y_H + Y_R$$

where YH is the bare-hull contribution and YR is the additional sway generated by the rudder. The rudder term YR relies on the inflow modified by the propeller's wake fraction, considering β ≈ v/U (for small sway v ≪ U)

$$w_P = w_{P0}\,\exp(-4\,\beta^2)$$

which reduces the nominal flow speed U to an effective propeller inflow UP. This inflow sets the advance ratio

$$J_P = \frac{u_P}{n_P\,D_P}$$

and hence the thrust coefficient

$$K_T = k_0 + k_1\,J_P + k_2\,J_P^2$$

At the rudder, velocities (uR, vR) are further corrected for wake and yaw effects, producing an effective rudder angle

$$\delta_R = \gamma_R\,\beta_R - \arctan\!\left(\tfrac{y_R}{x_P}\right), \quad \beta_R = \beta - \ell_R\,r$$

at which each section develops a normal force

$$F_N = \tfrac{1}{2}\,\rho\,A_R\,(u_R^2 + v_R^2)^{\frac{6.13\,\Lambda}{\Lambda+2.25}}\;\sin(\alpha), \quad \alpha = \delta - \delta_R$$

Summing the port and starboard section forces and applying the hull-amplification factor aH (with the MMG sign convention) yields

$$Y_R = -\big(1 + a_H\big)\,\big(F_{N,P}\,\cos\delta_P + F_{N,S}\,\cos\delta_S\big)$$

For this study, the total rudder-induced sway force YR was computed according to the MMG convention (hence the leading negative sign) by setting the port and starboard deflections to δP=-30◦ and δS=+30◦, respectively, yielding the two vector components of YR.

$$F_{\text{thru\_sway}} = \big[Y_{\text{portside}},\, Y_{\text{starboard}}\big]$$

4. Markov Decision Process Formulation

In this work, station-keeping under fishtailing is posed as a Markov Decision Process (MDP) defined by the tuple (S,A,T,R). The state space S comprises 100 discrete configurations, each encoding a unique combination of ship heading ψ, hawser angle γ, and instantaneous sway direction (port-side or starboard drift). Two of these—states 15 and 65—are designated as equilibrium points, representing perfectly balanced upstream and downstream alignments, respectively.

The action set A consists of ten control inputs formed by the Cartesian product of steering direction (port or starboard) and thrust intensity levels (0 %, 25 %, 50 %, 75 %, 100 %). Thus, at each decision epoch the agent selects one of these ten discrete rudder–propeller commands.

The transition dynamics T capture how applied forces interact with both hydrodynamic resistance and the vessel's inertia. If the chosen action commands zero thrust, the system advances cyclically among sway states, reproducing the natural fishtailing oscillation. For nonzero thrust, the next state is determined by comparing the commanded force against the sum of current-induced and hawser sway resistances and inertial resistance: insufficient forces are accumulated in a buffer, while forces that exceed total resistance reset the buffer and deactivate peripheral sway states of the next state.

The reward function R is designed to encourage prompt suppression of fishtailing. Every time step incurs a temporal penalty of −1. Applying thrust that opposes the current sway direction yields +1, and applying thrust that reinforces the drift is penalized −1. A reward of +2 is granted whenever the applied force overcomes the combined resistance, and an additional +5 bonus is awarded if that action drives the system into one of the equilibrium states (S15 or S65) immediately thereafter. This structure guides the agent toward control policies that both arrest lateral oscillations and achieve stable alignment.

5. Main Results

Figure 5 shows a heatmap of the learned Q-values for all 100 states (rows) and ten discrete actions (columns). The two equilibrium states—state 15 and state 65—stand out with the highest Q-values across multiple actions (highlighted by the brightest bands), confirming that the agent has correctly learned to favor transitions into these equilibria. This strong concentration of value at states 15 and 65 indicates excellent agreement between the learned policy and the expected goal of driving the system toward its stable, aligned configurations.

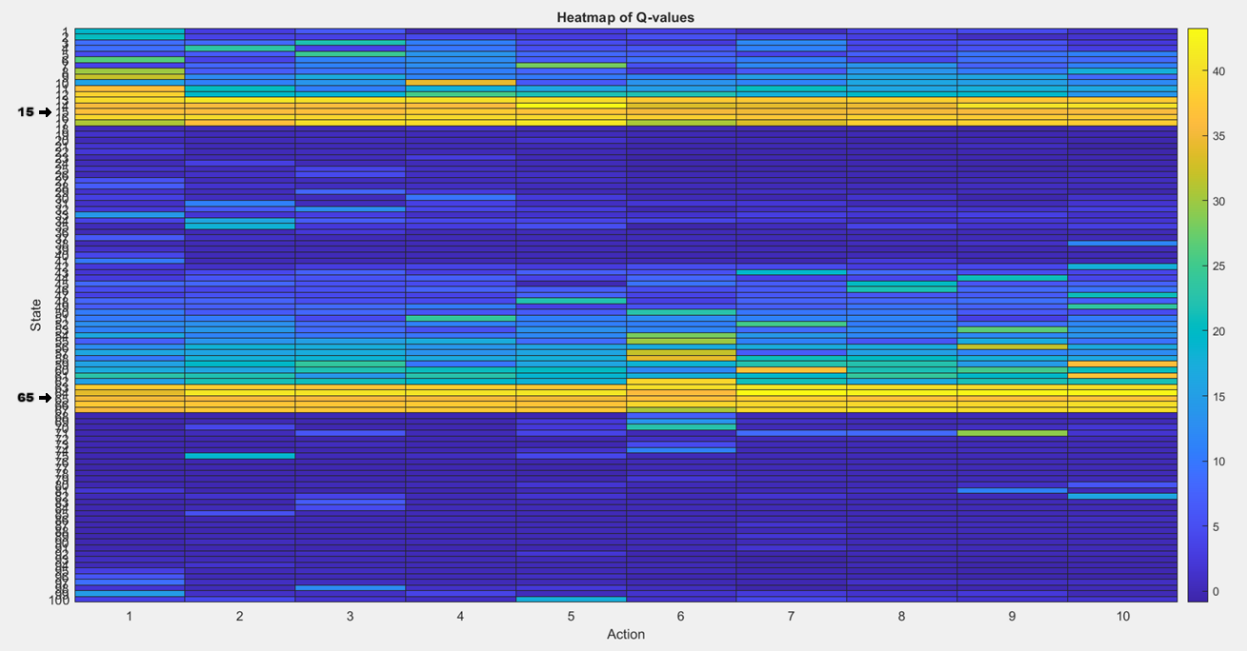

Figure 5 - Heatmap of Q-values

The following still frames, Figure 6, are taken from the video output of the vessel-motion simulation. Once again, it is evident that the ship quickly escapes the fishtailing oscillation and settles into one of the static equilibrium positions, state 15 or state 65.

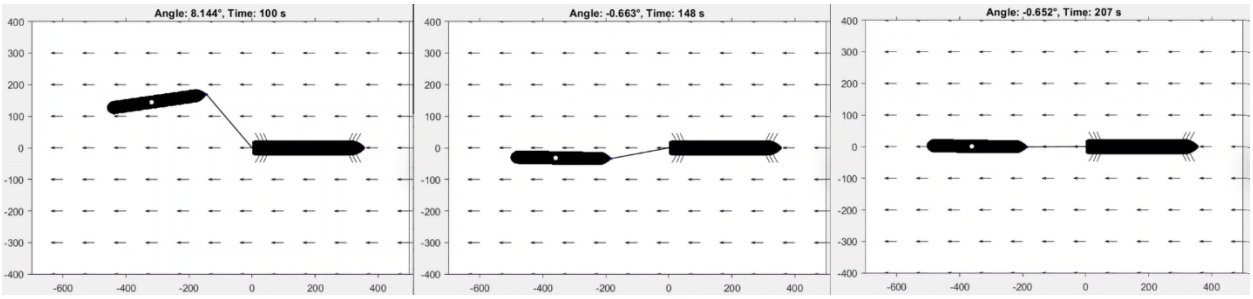

Figure 6 - Vessel-motion simulation

In the learned policy across 100 discrete states, the agent uses the neutral, zero-thrust command twenty times—most notably at the two equilibrium states (15 and 65)—demonstrating that it "knows" to hold position once stability is reached. At the extremes of the state space, where sway and resistance are greatest, it consistently selects high-intensity actions (Actions 4 and 6), each used fifteen times, to apply strong corrective forces. In the mid-range, the policy favors moderate thrust levels—Action 3 appears in twelve states, while Actions 5 and 7 are chosen in eight and seven states respectively—providing gradual adjustments as the system moves toward equilibrium. Rarely used options, such as Action 8 (only two states), reveal control maneuvers the agent seldom needed, and occasional bursts of full opposite-side thrust (Action 10 in ten states) indicate strategic efforts to break persistent drift. Overall, this pattern mirrors a human approach: aggressive corrections at the edges, smooth moderation in the center, and steady neutral holding at balance points.

6. Conclusion

This study has demonstrated that a Markov Decision Process framework can successfully train a control policy in which rudder and propeller forces rapidly counteract the combined effects of current and hawser induced loads that produce fishtailing. By relying on time-series inputs of ship heading, hawser angle, and hawser tension, the MDP finds discrete thrust commands that stabilize lateral motion in minimal time. Future work should extend this model by incorporating additional environmental disturbances— as wind and wave forces—and by computing heading, hawser angle, and tension online at each interaction through low-frequency horizontal-motion equations rather than relying on pre-computed results. Moreover, transitioning from a discrete to a continuous action space—adjusting rudder angle and propeller RPM directly—would further close the gap to real-world operations. It is also crucial to address practical safety considerations, such as system redundancy and failure-mitigation strategies (for example, using oceanic tugs), to ensure reliable deployment. Ultimately, this work has met its goal of validating reinforcement-learning–based station-keeping control as a viable alternative to DP3 for offloading operations, thereby laying the groundwork for optimizing offshore logistics through RL-driven control of conventional oil tankers at spread-moored FPSOs.

References

[1] Leite, A. J. P., Simos, A. N., Tannuri, E. A., & Aranha, J. A. P. (1998). Current forces in tankers and bifurcation of equilibrium of turret systems: Hydrodynamic model and experiments. *Applied Ocean Research, 20*, 145–156.

[2] Simos, A. N., Tannuri, E. A., Aranha, J. A. P., & Leite, A. J. P. (2001). *Theoretical analysis and experimental evaluation of the fishtailing phenomenon in a single-point moored tanker*. In *Proceedings of the International Society of Offshore and Polar Engineers Conference* (ISOPE). Cupertino, CA: International Society of Offshore and Polar Engineers.

[3] Sukas, O. F., Yasukawa, H., Kose, K., & Hino, T. (2019). System-based prediction of maneuvering performance of twin-propeller and twin-rudder ship using a modular mathematical model. *Applied Ocean Research, 86*, 282–297.

[4] Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press. https://www.andrew.cmu.edu/course/10-703/textbook/BartoSutton.pdf