Graphical Abstract

Trust Region Policy Optimization-Based Pitch Control for Floating Offshore Wind Turbines in Above-Rated Wind Conditions

Flavie Didier, Salah Laghrouche, Daniel Depernet

Highlights

Trust Region Policy Optimization-Based Pitch Control for Floating Offshore Wind Turbines in Above-Rated Wind Conditions

Flavie Didier, Salah Laghrouche, Daniel Depernet

- Research highlight 1
- Research highlight 2

Trust Region Policy Optimization-Based Pitch Control for Floating Offshore Wind Turbines in Above-Rated Wind Conditions

Flavie Didier[1,], Salah Laghrouche[1], Daniel Depernet[1]

[a]*Energy Department, Université Marie et Louis Pasteur, UTBM, CNRS, FEMTO-ST Institute (UMR 6174), Belfort, 90010, France*

Abstract

The control of Floating Offshore Wind Turbines (FOWTs) in Region III is challenging due to complex aerodynamic, hydrodynamic, and structural interactions. This paper presents a fully data-driven, model-free Deep Reinforcement Learning (DRL) controller based on the Trust Region Policy Optimization (TRPO) algorithm to regulate the collective blade pitch of a 5 MW semi-submersible FOWT. The controller was trained in high-fidelity simulations and experimentally validated in a wave basin using a Software-In-the-Loop (SIL) approach. Results show improved generator speed regulation and platform stability compared to a baseline Gain-Scheduling Proportional-Integral (GSPI) controller. However, performance degradation with generator speed overshoots was observed under extreme wind conditions. This study highlights the potential of DRL for FOWT control and identifies future directions to enhance robustness in harsh environments.

Keywords: Floating offshore wind turbine, collective blade pitch control, model-free control, deep reinforcement learning, Trust Region Policy Optimization, wave basin validation

1. Introduction

Global temperatures are projected to rise by up to 1.5°C by 2030 [?], threatening ecosystems, biodiversity, infrastructure, and human health. This trend is driven by greenhouse gas emissions, primarily from human activities. Therefore, transitioning to cleaner energy sources, such as renewables,

*Corresponding author.
1

is essential to mitigate this crisis. Wind energy stands out as a promising solution, with offshore deployment addressing the spatial limitations of onshore wind farms. Among these, Floating Offshore Wind Turbines (FOWTs) offer significant advantages by operating in deeper waters with stronger, steadier winds, thus enhancing power generation efficiency.

Despite technological advancements and several notable floating wind deployments, such as the Kincardine Offshore Wind Farm by Principle Power [?], Hywind Tampen by Equinor [?], and the EolMed project led by BW Ideol [?], the control of FOWTs remains a significant challenge due to their complex, nonlinear, and highly coupled dynamics. The latest outlook by IRENA [?] highlights the rapid growth of floating offshore wind, while the World Energy Council [?] documents the increasing number of large-scale global projects. The offshore environment introduces continuous stochastic disturbances from wind, waves, and currents, which interact with the floating platform's six Degrees of Freedom (DoFs). The offshore environment introduces continuous stochastic disturbances from wind, waves, and currents, which interact with the floating platform's six Degrees of Freedom (DoFs). Among the three standard operating regions, Region III—associated with above-rated wind speeds—presents the most critical control challenges. In this region, the control objective shifts from maximizing energy capture to maintaining generator speed at its rated value, commonly through pitch-to-feather control. However, while this strategy effectively reduces rotor thrust, it can also introduce negative aerodynamic damping [?]. This phenomenon occurs when blade pitch adjustments inadvertently excite, rather than suppress, platform pitch oscillations—particularly near the structure's natural frequency. Such destabilizing feedback may lead to resonant behavior, posing risks to both structural integrity and power quality. These challenges call for advanced control strategies capable of simultaneously regulating generator speed and suppressing platform motion, while accounting for the nonlinear and coupled aero-hydro-servo-elastic dynamics of the system.

Traditional control methods primarily use Collective Blade Pitch (CBP) to adjust blade angles and regulate power output based on operating conditions. Jonkman [?] provided a detailed dynamic modeling framework for offshore floating wind turbines, laying the foundation for control system development. Based on this framework, Larsen and Hanson [?] proposed a method to mitigate low-frequency tower vibrations induced by pitch control, addressing a key limitation in early floating wind platforms. The widely referenced 5-MW baseline controller developed by Jonkman et al. [?], based on a Gain-Scheduled Proportional-Integral (GSPI) approach, is often adapted from bottom-fixed wind turbines and used as a benchmark. Furthermore, [?] analyzed how control strategies influence pitch damping characteristics, highlighting the sensitivity of floating systems to feedback design. These traditional approaches, while foundational, remain highly sensitive to environmental disturbances and struggle with the nonlinear, coupled dynamics specific to FOWTs.

Recent advancements emphasize Multi-Input Multi-Output (MIMO) control models that account for the coupled dynamics of FOWTs. Linear control strategies such as the Linear Quadratic Regulator (LQR) have been explored extensively. Namik and Stol contributed significantly to this field through a series of studies: in [?], they introduced an individual blade pitch control scheme for FOWTs, in [?], they analyzed the performance of such control strategies on different floating platforms, and in [?], they extended the approach to spar-buoy configurations, highlighting control performance under platform motion. Christiansen et al. later proposed an optimal control design tailored to ballast-stabilized floating turbines [?] and further investigated wave disturbance reduction [?]. Lemmer et al. [?] complemented these works by offering a systematic comparison of linear control methods for disturbance rejection. These approaches typically rely on linearized models derived around operating points to simplify the floating wind turbine dynamics.

Linear Parameter-Varying (LPV) control strategies have also emerged, allowing gain scheduling across operating regions. Bagherieh et al. [?] proposed LPV control above rated wind speed, while Zhao et .al [?] extended it with switching mechanisms to adapt to platform motion.

To enhance robustness, H_∞ control has been investigated using various formulations. Bakka et al. [?] developed an LMI-based synthesis approach for output feedback control, while Li and Gao [?] applied generalized H_∞ structural control to mitigate loads. Cortes Sanchez [?] explored disturbance rejection strategies under wind and wave conditions. Further studies by Bakka et al. [?] and Hara et al. [?] highlighted the effectiveness of gain-scheduled and experimentally validated H_∞ control in offshore settings.

Model Predictive Control (MPC) has also emerged as a powerful alternative due to its ability to manage constraints and optimize multivariable responses. Mahmoud and Oyedeji [?] provided a comprehensive survey on

MPC applications in wind turbine systems. Lemmer et al. [?] and Cunha et al. [?] implemented MPC strategies to reduce blade fatigue and suppress platform motion. Okada et al. [?] extended MPC to parameter-varying models for greater flexibility, while Wakui et al. [?] introduced preview-based MPC techniques to further enhance platform stabilization and power output regulation.

Nonlinear Model Predictive Control (NMPC) strategies aim to overcome the limitations of linear assumptions. Schlipf et al. [?] compared feed-forward and MPC approaches using LIDAR measurements to enhance wind preview capabilities. In a subsequent study, Schlipf et al. [?] applied NMPC to floating wind turbines, demonstrating its effectiveness in managing platform dynamics. The work in [?] further extended NMPC formulations with real-time LIDAR inputs, improving the responsiveness of turbine control systems. Raach et al. [?] focused on the integration of individual pitch control within NMPC frameworks to mitigate loads and enhance stability. Shah et al. [?] later proposed an NMPC-based solution aimed specifically at minimizing platform motions, validating its capability under various offshore operating conditions.

Sliding Mode Control (SMC) has also gained traction for its robustness to model uncertainties and external disturbances. Bagherieh et al. [?] pioneered the application of SMC with nonlinear input-output feedback linearization to enhance control precision in floating offshore systems. Zhang et al. [?] introduced an adaptive robust control approach using SMC, while their later work [?] proposed a super-twisting version tailored for FOWTs with CBP strategy. A foundational contribution to the field came from Shtessel et al. [?], who developed an adaptive-gain formulation of the super-twisting SMC, improving chattering mitigation and convergence. Building upon these foundations, Zhang and Plestan [?] applied an adaptive SMC to floating wind turbines equipped with permanent magnet synchronous generators, demonstrating improved dynamic response. Taleb et al. [?] introduced a novel adaptation law designed specifically for FOWT control. Finally, Taleb and Plestan [?] proposed a reduced-parameter version of the adaptive super-twisting controller, which facilitates practical implementation while preserving robustness. Despite their strengths, many of these model-based controllers depend on linearized or reduced-order representations — often derived from OpenFAST [?] — which limits their performance when the system operates far from nominal conditions.

While nonlinear controllers based on reduced-order models of FOWTs have improved disturbance rejection and handling of platform dynamics, their performance often hinges on the accuracy of the underlying model. Basbas et al. [?] provided a comprehensive review of modeling strategies for nonlinear control design, highlighting the trade-offs between fidelity and complexity. Among them, Betti et al. [?] developed a control-oriented model that simplifies platform and turbine interactions while maintaining essential dynamics. Similarly, Lemmer [?] proposed low-order modeling techniques tailored for control and optimization purposes, emphasizing practical implementability. Homer et al. [?] introduced a 3D physics-based modeling framework designed specifically for control synthesis, accounting for spatial effects in floating platforms. Despite these advances, the dependency on accurate models leaves such approaches sensitive to structural uncertainties and environmental variability, which limits robustness and adaptability in real-world applications. Recent studies by Basbas et al. [?] and Liu et al. [?] have proposed super-twisting-based control laws tailored to reduced-order, control-oriented models of floating offshore wind turbines. While these approaches improve disturbance rejection and dynamic response, their performance remains sensitive to the accuracy of the underlying modeling assumptions and simplifications. To overcome this limitation, Didier et al. [?] introduced neural network-based observers with adaptive laws derived from Lyapunov analysis, enabling real-time compensation of model uncertainties and unmodeled dynamics. Furthermore, in [?], a higher-order sliding mode controller was integrated with a neural network observer to address the full-order FOWT dynamics, including the second-order behavior of the blade pitch actuator.

Building on these hybrid approaches, data-driven, model-free control methods offer a promising alternative by leveraging data to bypass explicit system dynamics, making them well suited for highly nonlinear systems [?]. Recent advances in computing and data processing enable techniques like fuzzy logic, Machine Learning (ML), Deep Learning (DL), and genetic programming [?]. In the context of FOWTs, Kane [?] demonstrated the feasibility of using ML-based control strategies for individual blade pitch regulation, showing improved adaptability compared to conventional methods. Roh [?] proposed a DL-based controller that compensates for actuator delay, enhancing tracking performance in floating wind systems. These approaches design controllers using input-output data, enhancing adaptability to real-time changes in FOWT dynamics. Although their application in FOWTs remains limited, data-driven methods hold significant potential to address traditional control challenges, making them a key area for future research.

Among the realm of ML, Reinforcement Learning (RL) stands out as a particularly promising approach. Rooted in the principles of the Markov Decision Process (MDP), RL agents are designed to learn optimal control policies through direct interaction with their environment, optimizing long-term rewards based on observed data. A policy, in the context of RL, defines the strategy or mapping from an agent's observed states to its actions within the environment. Deep RL (DRL), which integrates neural network structures into the RL framework, exhibits a significant ability to address the intricate control challenges of conventional model-based methods, such as dependence on analytical models and lack of robustness to modeling errors and uncertainties. Notably, the adaptability of neural networks to dynamic environments reduces the need for extensive tuning and enables the controller to generalize behaviors across varying operating conditions. The actor-critic architecture, a popular approach within DRL, facilitates this by employing two neural networks—the actor and the critic—to respectively improve the policy and evaluate its performance. The learning process unfolds in two main steps: first, the critic evaluates the current policy, and then the actor improves the policy to better meet control objectives.

Although applications of DRL in the FOWT control domain are emerging, most existing works have focused on model-based approaches, where either the FOWT dynamics are learned via neural networks or approximators are embedded within the control structure. For instance, Xie et al. [?] proposes a model-based Incremental Dual Heuristic Programming (IDHP) framework that requires an internal model of the system. Chen et al. [?] implement a Software-in-the-Loop (SIL) architecture combining DRL with dynamic response analysis for FOWTs. In a subsequent study, Chen and Hu [?] investigate the influence of key parameters on dynamic response prediction using artificial intelligence-based methods. Additionally, Chen et al. [?] propose a simulation annealing-based optimization algorithm to improve the forecast accuracy of dynamic responses in FOWTs. While these studies demonstrate the potential of DRL, many rely on deterministic policy gradient methods such as DDPG, which are often sensitive to hyperparameter tuning and less robust to exploration noise.

This paper addresses this gap by developing a DRL-based CBP control system specifically for Region III operation. The controller is based on the actor-critic Trust Region Policy Optimization (TRPO) algorithm, which optimizes control performance without relying on an internal dynamics model. TRPO introduces a trust region constraint using the Kullback-Leibler (KL) divergence between successive policies, enabling more stable learning than first-order methods like DDPG or Twin Delayed DDPG (TD3). This stability is particularly important for floating wind systems, where abrupt policy updates can exacerbate nonlinear effects such as negative damping in the platform pitch dynamics. Furthermore, TRPO employs a stochastic policy formulation, promoting more effective exploration of the state-action space—a key advantage in offshore environments characterized by stochastic disturbances and coupled dynamics. While TD3 partially addresses DDPG's limitations through twin critics and target smoothing, its deterministic policy structure may still hinder robustness under variable conditions.

Thus, given the safety-critical nature of pitch control in Region III and the experimental validation required for deployment, TRPO was selected for its favorable balance of stability, safety, and exploration capacity. To the best of our knowledge, this study is the first to propose a fully data-driven, model-free DRL controller validated in a realistic wave basin environment for FOWTs operating in Region III. This work demonstrates the practical feasibility of DRL-based control strategies through both high-fidelity simulations and experimental implementation, offering new insights into floating wind turbine control under complex offshore conditions. The proposed controller is benchmarked against the conventional GSPI controller [?] using a validated wave basin platform that replicates Region III hydrodynamic scenarios.

The main contributions of this paper are as follows:

- The development of a novel, fully data-driven, model-free DRL controller based on the TRPO algorithm within an actor-critic framework, specifically designed to optimize the control of FOWTs in Region III.

- The implementation of a control strategy that eliminates the need for detailed dynamic modeling, enabling the controller to adapt directly to complex environmental conditions through learning from simulation data.

- The experimental validation of the proposed DRL controller in a realistic wave basin setup using a Software-In-the-Loop (SIL) approach, demonstrating its superior performance in generator speed regulation and platform stability compared to the conventional GSPI controller.

- A comprehensive performance analysis under various wind and wave conditions, highlighting both the robustness of the DRL controller and areas for improvement, particularly under extreme offshore scenarios.

The structure of this paper is as follows: Section II outlined the primary control objectives for the considered semi-submersible FOWT system. Section III details the design of the proposed DRL controller, including the implementation specifics and training process. Section IV discusses the experimental setup within the wave basin facility and presents a comparative analysis of the controller's performance under various environmental conditions. Finally, Section V concludes the study and suggests avenues for future research.

2. Problem Formulation

This section presents the FOWT system considered for the pitch control problem and outlines the associated control objectives.

2.1. Floating Wind Turbine System

This study considers the NREL OC4-DeepCWind 5 MW semi-submersible FOWT, illustrated in Figure **??**. The turbine specifications, summarized in Table **??**, are derived from the NREL 5 MW baseline turbine [**?**], with additional platform characteristics from [**?**]. The FOWT system consists of a 5 MW wind turbine mounted on a semi-submersible platform, stabilized by ballast and a large waterplane area. The platform's mooring system maintains its position while resisting drift caused by wind and waves.

Table 1: Specifications of the NREL OC4-DeepCwind 5 MW Semi-submersible FOWT.

Parameter	**Value**
Rated power	5 MW
Rotor orientation, configuration	Upwind, 3 blades
Rotor diameter	126 m
Hub diameter	3 m
Hub height	90 m
Cut-in wind speed	3 m/s
Rated wind speed	11.4 m/s
Cut-out wind speed	25 m/s
Rated rotor speed	12.1 rpm
Rated generator speed	1173.7 rpm
Rated generator torque	43,093.55 Nm
Gearbox ratio	1:97
Generator efficiency	0.944
Minimum blade pitch setting	0°
Maximum blade pitch setting	90°
Maximum absolute blade pitch rate	8°/s

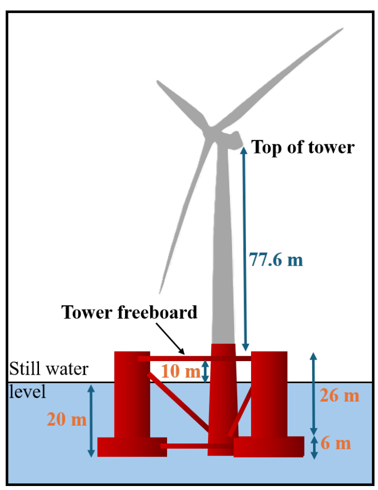

Figure 1: Semi-submersible FOWT structure.

The fundamental principle of energy transmission within the wind turbine, shown in Figure ??, involves converting wind kinetic energy into electrical power.

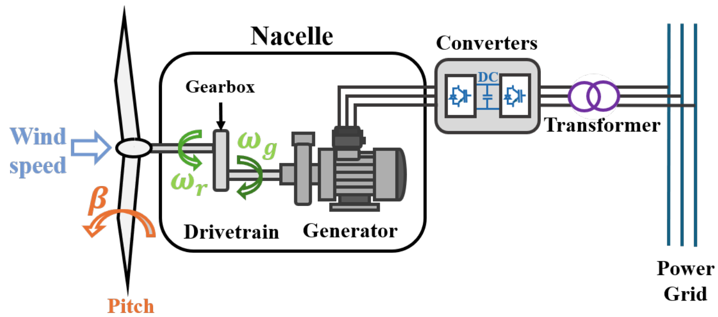

Figure 2: Wind turbine energy transmission system.

The wind exerts aerodynamic forces on the turbine blades, generating rotational motion:

$$P_{\text{wind}} = \frac{1}{2}\rho_a \pi R_r^2 v^3, \tag{1}$$

where ρ_a is the air density, R_r is the rotor radius, and v is the wind speed. Not all wind energy is captured due to aerodynamic inefficiencies, quantified by the power coefficient C_p, which depends on the blade pitch angle β and the tip-speed ratio λ:

$$C_p = f(\beta, \lambda), \quad \text{with } \lambda = \frac{\omega_r R_r}{v}, \tag{2}$$

where ω_r is the rotor speed. The aerodynamic power P_A extracted from the wind becomes:

$$P_A = \frac{1}{2}\rho_a C_p \pi R_r^2 v^3. \tag{3}$$

The drivetrain transmits mechanical energy from the rotor to the generator. Neglecting friction, its dynamics can be modeled as a rigid one-mass shaft, as described in [?]:

$$\dot{\omega}_r = \frac{1}{J_l}\left(\frac{P_A}{\omega_r} - \eta_g T_g\right), \tag{4}$$

where η_g is the gearbox ratio, T_g is the generator torque, and $J_l = J_r + \eta_g^2 J_g$ represents the equivalent shaft inertia. Finally, the generator converts mechanical energy into electrical power P_g:

$$P_g = T_g\omega_g, \quad \text{with } \omega_g = \eta_g\omega_r. \tag{5}$$

While the energy transmission dynamics provide valuable insight into FOWT behavior, the DRL-based controller proposed in this study does not require explicit dynamic modeling. Instead, it learns optimal control policies directly from data, enabling a fully model-free approach to tackle the challenges in FOWT control.

2.2. Control Objectives

Given the slower role of the nacelle's yaw motion in immediate control responses, it is neglected in this study. The primary control inputs are the blade pitch angle and generator torque, with the pitch-to-feather method [?] being employed in Region III. This method regulates generator speed by adjusting blade pitch while maintaining constant torque, offering adaptability to wind fluctuations and reducing structural loads.

Unlike bottom-fixed turbines, FOWTs introduce additional complexity due to their floating platforms, which exhibit six DoFs. The platform's motion makes FOWTs more vulnerable to disturbances from wind, waves, and currents. At above-rated wind speeds, pitch-to-feather control reduces rotor thrust, potentially leading to negative damping, where control actions amplify platform pitching motion and cause resonant oscillations [?]. Balancing power regulation and platform stability is particularly challenging, as these objectives often conflict. Aggressive pitch adjustments can improve power regulation but exacerbate platform instability and fatigue loads.

This study addresses these challenges by designing a DRL-based pitch controller that considers two primary control objectives:

- **Maximize power output:** The goal is to maintain the generator power at its rated value of 5 MW. Since the generator torque is constant, this translates to maintaining the generator rotational speed (ω_g) at its rated value ($\omega_{gd} = 122.9096\,\text{rad/s}$). The error e_1 quantifies the deviation of the current rotational speed from the rated speed:

$$e_1 = \omega_g - \omega_{gd}. \tag{6}$$

- **Ensure platform stability:** Minimizing pitch oscillations is critical, as platform motion directly impacts generator speed and power output. While eliminating these oscillations entirely in Region III is not feasible, reducing their magnitude enhances system stability. The error e_2 measures the deviation of the platform pitch rate (ω_y) from its reference value ($\omega_{y,\text{ref}} = 0\,\text{rad/s}$):

$$e_2 = \omega_y - \omega_{y,\text{ref}}. \tag{7}$$

The controller aims to minimize both e_1 and e_2, achieving an optimal trade-off between power regulation and platform stability.

To meet these objectives, a DRL-based controller is developed using the TRPO algorithm, an on-policy, actor-critic, model-free approach known for its stability and effectiveness in complex control environments. The DRL controller dynamically adjusts the CBP angle while maintaining the generator torque at its rated value, ensuring robust performance in the challenging conditions of Region III.

3. DRL-based Collective Blade Pitch Control Design in Region III

This section presents the design of the DRL agent, serving as a CBP controller for the introduced 5 MW semi-submersible FOWT. The OpenFAST simulation software, incorporating the NREL OC4-DeepCWind model, serves as the agent's training environment. This black-box approach enables the development of control strategies without requiring prior knowledge of the system's underlying dynamics.

The proposed controller, illustrated in Figure **??**, employs an artificial neural network architecture to optimize CBP control. It integrates actor and critic networks that interact with the simulated FOWT environment, leveraging the TRPO algorithm for policy updates. Additionally, the architecture includes an action selector to generate control inputs and a reward calculator to evaluate performance.

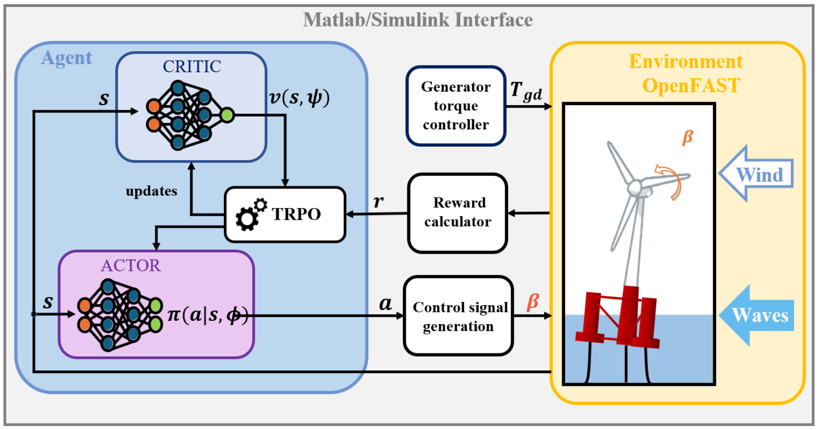

Figure 3: Interaction between the DRL agent and the OpenFAST simulation software.

3.1. MDP Formulation for FOWT Control in Region III

The interaction between the DRL agent and the simulated FOWT environment is represented using the MDP framework. MDP models the interaction between a decision-making agent and its environment [?], where the environment includes the system to be controlled and any external disturbances. The agent interacts with the environment through three primary signals, as depicted in Figure ??:

- **Observations s:** Information received from the environment describing its current state at a given time.
- **Actions a:** Decisions made by the agent that directly affect the environment's state.
- **Rewards r:** Feedback from the environment evaluating the effectiveness of the agent's actions.

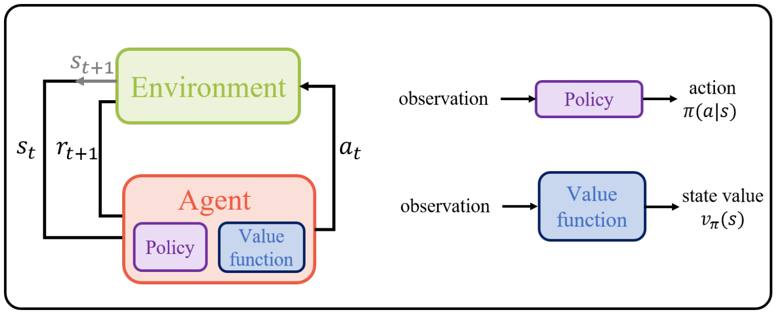

Figure 4: Interaction between the agent and environment in MDP.

By effectively defining the MDP elements—state space (S), action space (A), reward function (R), discount factor (γ), and transition probability (p)—the FOWT control problem is formulated as a robust foundation for developing a DRL-based control strategy.

3.1.1. States (S)

According to the energy transmission model (**??**)-(**??**), the power generated by the FOWT system depends on the generator rotational speed ω_g. This speed, in turn, is influenced by wind speed v, blade pitch angle β, and indirectly by the platform pitch θ_y and its angular velocity ω_y. Hence, the generated power P can be expressed as:

$$P = f(\omega_g, v, \theta_y, \omega_y, \beta). \tag{8}$$

However, when employing a model-free TRPO algorithm, explicit knowledge of the complete system dynamics is not required for controller design. This means that the function f can remain unknown. Instead, the state space S is designed to include all essential aspects needed for the DRL agent to make informed decisions:

- **Generator speed** (ω_g) [rad/s]: Indicates energy capture and power generation.
- **Generator speed error** ($e_1 = \omega_g - \omega_{gd}$) [rad/s]: Tracks deviation from the rated speed.
- **Platform pitch angle** (θ_y) [rad]: Reflects platform motion due to wind and waves.

- **Platform pitch angular velocity** ($e_2 = \omega_y$) [rad/s]: Monitors the rate of platform pitch changes.
- **Previous blade pitch control** (β_{t-1}) [rad]: Captures the controller's last action.

The observation vector s is then:

$$s = [\omega_g, e_1, \theta_y, \omega_y, \beta_{t-1}]^\top. \tag{9}$$

3.1.2. Actions (A)

The action space is defined as the continuous normalized range $A = [-1, 1]$. Accordingly, the policy network within the DRL controller outputs the mean and standard deviation of a Gaussian distribution at each time step. A normalized continuous action is then stochastically sampled as:

$$a_{\text{norm}} \sim \mathcal{N}(\mu(s), \sigma(s)^2),$$

where $\mu(s)$ and $\sigma(s)$ are the state-dependent mean and standard deviation produced by the actor network. During training, this sampling encourages exploration, while during evaluation, only the mean action is used to ensure deterministic behavior.

The sampled action $a_{\text{norm}} \in [-1, 1]$ is linearly mapped to the physical blade pitch range $[0, \frac{\pi}{2}]$ rad (i.e., 0° to 90°). To ensure safe and stable interaction with the FOWT environment—particularly during early exploratory phases—the resulting action is saturated within the valid pitch range, and a rate limiter is applied to constrain variation between successive time steps. This limiter is consistent with the Baseline controller [?] and is expressed as:

$$\dot{u} \leq \beta_{\text{rate}}, \tag{10}$$

where $\beta_{\text{rate}} = 8^\circ/\text{s}$ denotes the maximum allowable blade pitch rate.

3.1.3. Reward function (R)

The reward function guides the DRL controller by quantifying the immediate outcomes of actions to achieve the presented key control objectives: power regulation and platform stability.

The general reward signal r is expressed as:

$$r = W_1 \cdot f(e_1) + W_2 \cdot g(e_2) + W_3 \cdot h(\dot{u}), \tag{11}$$

where W_1, W_2, W_3 are weights representing the relative importance of each objective:

- **Generator speed tracking** ($f(e_1)$): Penalizes deviations of generator speed ω_g from its rated value ω_{gd}, ensuring efficient power generation.
- **Platform stability** ($g(e_2)$): Discourages large pitch rates ω_y to reduce platform motion.
- **Control smoothness** ($h(\dot{u})$): Limits abrupt control changes to minimize actuator wear.

The specific base reward signal used in this study, obtained through manual tuning, is:

$$r_{\text{base}} = \max(0, 15 - |e_1|) - W_1 \cdot |e_1| - W_2 \cdot |e_2| - W_3 \cdot |\dot{u}| + G_1, \tag{12}$$

where $G_1 = 10$ if $|e_1| \leq 2$, otherwise $G_1 = 0$. This bonus incentivizes operation near the rated generator speed. The weights are tuned as $W_1 = 0.5$, $W_2 = 0.08$, and $W_3 = 0.01$, achieving a balanced trade-off among objectives.

To further ensure safe operation, a constraint-based penalty is introduced in the reward function. Specifically, a penalty term C is added to the base reward r_{base}, resulting in the final reward signal used by the DRL agent:

$$r = r_{\text{base}} + C, \tag{13}$$

where the constraint penalty C is defined as:

$$C = \begin{cases} -1, & \text{if } \omega_g \notin [80, 163], \\ 0, & \text{otherwise.} \end{cases} \tag{14}$$

This constraint discourages the agent from operating outside the generator's allowable speed range and helps guide policy learning toward safe and effective control actions.

3.1.4. Discount factor γ

The discount factor, set to $\gamma = 0.99$, balances short- and long-term rewards, ensuring the agent values both immediate power generation and the long-term stability of the system.

3.1.5. Transition probability p

The transition probability describes how system states change in response to actions, influenced by wind, waves, and hydrodynamic effects. Using the TRPO algorithm, the DRL agent empirically learns these dynamics through interactions with the environment. By iteratively observing state-action-reward transitions (s, a, r, s'), the agent refines its policy to adapt to the nonlinear and stochastic nature of FOWT dynamics.

3.2. Neural Networks Architectures

The DRL controller includes two neural networks: the actor $\pi(a|s, \phi)$ and the critic $v(s, \psi)$, where ϕ and ψ represent the parameters of the actor and critic networks, respectively. The critic network evaluates the quality of states by estimating the expected return, while the actor network determines the action probabilities.

3.2.1. Critic network

The architecture of the actor network is represented in Figure **??** and is organized as:

- **Input layer:** Receives the state vector s.
- **Hidden layers:** Two fully connected layers with 256 units each, using ReLU (Rectified Linear Unit) activation functions.
- **Output layer:** A fully connected layer that produces a scalar output, representing the value $v(s, \phi)$, which estimates the expected return from state s.

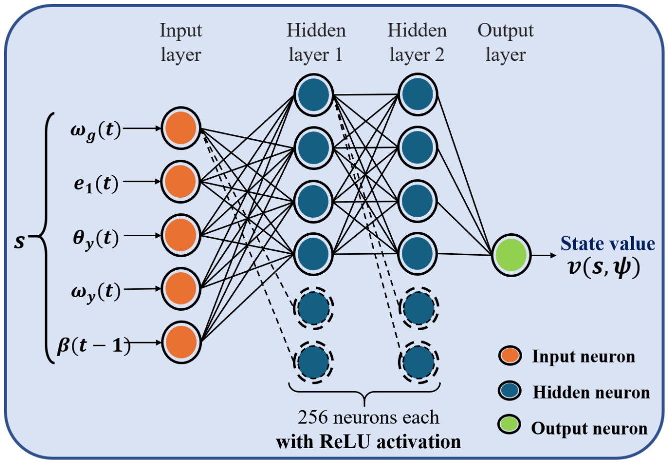

Figure 5: Architecture of the critic neural network employed within the DRL controller.

The total number of learnable parameters ψ in the critic network is calculated based on the layer sizes:

$$\text{First fully connected layer: } (5 \times 256) + 256 = 1,536 \text{ parameters}$$

$$\text{Second fully connected layer: } (256 \times 256) + 256 = 65,792 \text{ parameters}$$

$$\text{Output layer: } 256 \times 1 = 256 \text{ parameters}$$

$$\text{Total: } 1,536 + 65,792 + 256 = 67,584 \text{ parameters.}$$

3.2.2. Actor network

The architecture of the actor network is represented in Figure **??** and is organized as:

- **Input layer:** Receives the state vector s.
- **Hidden and output layers:** The architecture consists of fully connected layers, branching into two separate paths:
  - **Mean path:** Outputs the mean values for actions using the $Tanh$ activation function, ensuring outputs are bounded within [-1, 1], scaled appropriately for the action space.
  - **Standard deviation path:** Outputs the standard deviation using the Softplus activation function, ensuring that the values remain positive.

Together, the mean and standard deviation outputs define the parameters of the Gaussian action distribution from which the action is sampled.

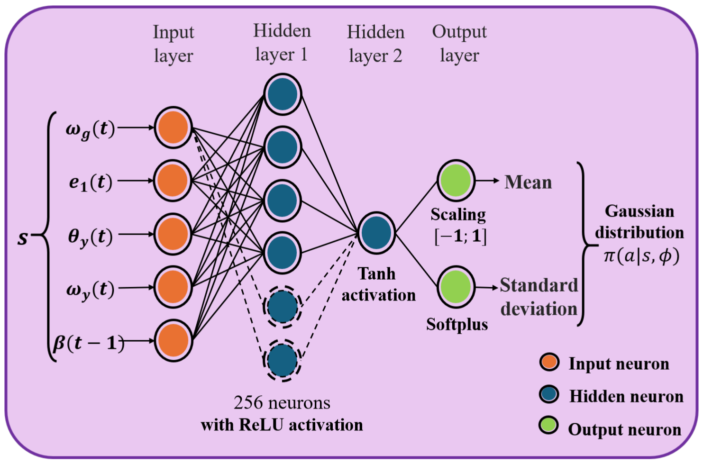

Figure 6: Architecture of the actor neural network employed within the DRL controller.

The total number of learnable parameters ϕ in the actor network is computed based on the layer sizes:

Input fully connected layer: $(5+1) \times 256 = 1,536$ parameters

Second fully connected layer: $(256+1) \times 1 = 257$ parameters

Mean path fully connected layer: $(1+1) \times 1 = 2$ parameters

Standard deviation path fully connected layer: $(1+1) \times 1 = 2$ parameters

Total: $1,536 + 257 + 2 + 2 = 1,797$ parameters.

3.2.3. State normalization

State normalization is a key preprocessing step in DRL, especially when neural networks are used as function approximators. In this implementation, observed states are scaled to the range $[-1, 1]$ using Min-Max scaling:

$$s_i^{norm} = \frac{2 \cdot (s_i - s_{i,\min})}{s_{i,\max} - s_{i,\min}} - 1. \tag{15}$$

This scaling ensures that each normalized state s_i^{norm} falls within $[-1, 1]$, stabilizing the learning and mitigating gradient instability from widely varying input magnitudes. Normalization is crucial for the semi-submersible FOWT, where state variables have diverse units and ranges. The ranges used in this study are:

- **Generator speed (ω_g)**: [80, 163] rad/s
- **Platform pitch (θ_y)**: [-0.035, 0.105] radians
- **Platform pitch rate (ω_y)**: [-0.035, 0.035] rad/s
- **Blade pitch angle (β)**: [0, π] radians

3.3. Training Process of the DRL Controller

The TRPO algorithm is used to optimize the control policy of the 5 MW semi-submersible FOWT. The actor network $\pi(a|s, \phi)$ and critic network $v(s, \psi)$ are trained iteratively to maximize the expected reward. Training is guided by gradient-based optimization using experiences collected from the simulated environment.

3.3.1. Overview of TRPO algorithm

The TRPO algorithm, introduced by Schulman et al. in 2015 [**?**], is a robust actor-critic method designed to ensure stable policy updates. TRPO achieves stability by constraining the Kullback-Leibler (KL) divergence between successive policies, preventing abrupt changes that could destabilize the control system.

The policy update process under TRPO is summarized as follows:

1. **Initialization:** The critic $v(s, \psi)$ and actor $\pi(a|s, \phi)$ networks are initialized with random weights ψ and ϕ, respectively.
2. **Experience generation:** The agent interacts with the simulated FOWT environment using the current policy, generating episodes of state-action-reward tuples:

$$(s_{ts}, a_{ts}, r_{ts+1}, s_{ts+1}, \ldots, s_{ts+N-1}, a_{ts+N-1}, r_{ts+N}, s_{ts+N}). \qquad (16)$$

   Here, each tuple $(s_t, a_t, r_{t+1}, s_{t+1})$ represents the state, action, reward, and next state. The starting time step ts is incremented after each set of N experiences, as $ts \leftarrow ts + N$. The agent selects actions based on a probability distribution derived from the current policy: $\pi(a|s_t, \phi)$.

3. **Advantage function and Return calculation:** For each step t, where $t = 1, 2, \ldots, N$, calculate the Advantage function A_t and Return G_t:

Generalized Advantage Estimation (GAE):

$$A_t = \sum_{k=t}^{ts+N-1} (\gamma\lambda)^{k-t} \left(r_{k+1} + \gamma v(s_{k+1}, \psi) - v(s_k, \psi)\right), \tag{17}$$

where λ is the smoothing factor, and γ is the discount factor.

Return G_t:

$$G_t = A_t + v(s_t, \psi). \tag{18}$$

4. **Mini-batch selection:** Randomly select a subset of experience data to create mini-batches of size M, which will be used to update the networks.
5. **Critic network update:** Update the critic network by minimizing the mean squared error between the predicted values and the returns:

$$\psi \leftarrow \psi - \alpha_{\text{critic}} \nabla_{\psi} L_{\text{critic}}(\psi), \tag{19}$$

where α_{critic} is the learning rate for the critic, controlling the step size of each update, and the loss function $L_{\text{critic}}(\psi)$ is defined as:

$$L_{\text{critic}}(\psi) = \frac{1}{M} \sum_{i=1}^{M} \left(G_i - v(s_i, \psi)\right)^2. \tag{20}$$

6. **Actor network update:** Update the actor network to maximize the expected advantage, subject to the KL-divergence constraint. The objective function is defined as:

$$L_{\text{actor}}(\phi) = -\frac{1}{M} \sum_{i=1}^{M} \left(\frac{\pi(a_i|s_i, \phi)}{\pi(a_i|s_i, \phi_{\text{old}})} A_i + w\mathcal{E}_i(\phi, s_i) \right), \tag{21}$$

Here $\pi(a_i|s_i, \phi)$ is the probability of taking action a_i following the current policy, $\pi(a_i|s_i, \phi_{\text{old}})$ is the probability of taking action a_i following the old policy. The entropy term $\mathcal{E}_i(\phi, s_i)$ is defined as follows, where w is the entropy loss weight:

$$\mathcal{E}_i(\phi, s_i) = \frac{1}{2} ln \left(2\pi \cdot \mathrm{e} \cdot \sigma_i^2\right), \tag{22}$$

where σ_i is the standard deviation for the output action when in state s_i following the current policy. This entropy loss term encourages exploration by preventing the policy from becoming too deterministic.

7. **KL-divergence constraint satisfaction:** Ensure that the updated policy remains close to the old one by enforcing the following constraint:

$$\frac{1}{M}\sum_{i=1}^{M} D_{\mathrm{KL}}(\phi_{\mathrm{old}}, \phi, s_i) \leq \delta, \tag{23}$$

where δ controls the size of the policy update. The KL-divergence $D_{\mathrm{KL}}(\phi_{\mathrm{old}}, \phi, s_i)$ between the old policy and current policy is computed as:

$$D_{\mathrm{KL}}(\phi_{\mathrm{old}}, \phi, s_i) = \ln\left(\frac{\sigma_\phi}{\sigma_{\phi_{\mathrm{old}}}}\right) + \frac{\sigma^2_{\phi_{\mathrm{old}}} + (\mu_{\phi_{\mathrm{old}}} - \mu_\phi)^2}{2\sigma^2_\phi} - \frac{1}{2}. \tag{24}$$

Here, μ_ϕ and σ_ϕ represent the mean and standard deviation of the action distribution output by the current actor policy, while $\mu_{\phi_{\mathrm{old}}}$ and $\sigma_{\phi_{\mathrm{old}}}$ correspond to the mean and standard deviation of the action distribution under the old policy.

8. **Actor parameter update:** Perform the actor network parameter update by solving the optimization problem using the conjugate gradient descent:

$$\phi = \phi_{\mathrm{old}} + \alpha\sqrt{\frac{2\delta}{(H^{-1}g)^\top H^{-1}(H^{-1}g)}}H^{-1}g, \tag{25}$$

where H is the Hessian matrix, g is the gradient of the objective function $L_{\mathrm{actor}}(\phi)$, and α is the step size determined via line search:

$$\alpha \in \left\{1, \frac{1}{2}, \frac{1}{2^2}, \ldots, \frac{1}{2^{n-1}}\right\}. \tag{26}$$

9. **Iteration:** Repeat the steps iteratively, allowing the TRPO algorithm to continuously improve the policy.

The DRL controller is implemented using the **Reinforcement Learning Toolbox** in MATLAB/Simulink R2023a, where the actor and critic networks, as previously defined, are trained using the built-in TRPO agent framework. A custom environment interface was developed to integrate the TRPO agent with the OpenFAST simulator, allowing the agent to reset episodes, advance simulation steps, and compute rewards based on real-time system states.

3.3.2. Simulated environment

For a model-free approach, a robust and accurate simulation platform is essential to emulate the NREL OC4-DeepCwind 5 MW semi-submersible FOWT. The high-fidelity OpenFAST simulation software [**?**] is chosen for its detailed and reliable modeling of aerodynamic, hydrodynamic, and structural dynamics, as well as mooring system behavior. It incorporates a servo-elastic structural model to capture the complex forces acting on floating wind turbines. Figure **??** illustrates the key OpenFAST modules:

- **InflowWind:** Provides wind field inputs at the rotor, including speed and turbulence intensity.
- **ElastoDyn:** Simulates elastic structural dynamics of the drivetrain, tower, and nacelle.
- **BeamDyn:** Models blade flexibility using beam-type finite element methods.
- **AeroDyn:** Calculates aerodynamic loads using Blade Element Momentum Theory (BEMT).
- **HydroDyn:** Simulates hydrodynamic forces acting on floating structures.
- **Mooring:** Models mooring system dynamics via MAP++, FEAMooring, or MoorDyn.
- **ServoDyn:** Simulates pitch and torque actuator dynamics for control systems.

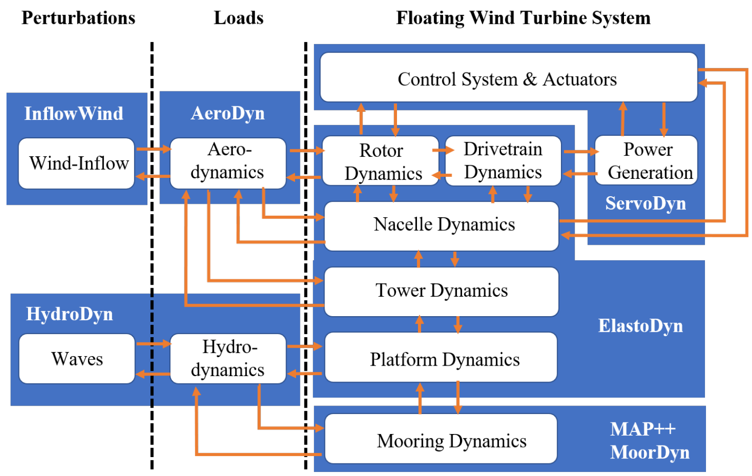

Figure 7: Architecture of the OpenFAST simulation software, depicting key modules used in FOWT system emulation.

3.3.3. *Training process*

The training process of the DRL controller follows these steps:

1. **Initialization:** Initialize the actor and critic networks with random weights to begin the exploration process.
2. **Episode execution:** For each training episode:
   (a) **Environment reset:** Reset the OpenFAST simulation environment using the `ElastoDyn` module and set initial conditions.
   (b) **Initial observation:** Retrieve the initial state s_0 and compute the initial action a_0 using the current policy.
   (c) **Interaction loop:** While the episode has not terminated, repeat the following steps:
      - Select an action $a \sim \pi(a|s)$ based on the current state s and the actor network.
      - Apply a to the environment and observe the next state s' and the reward r.
      - Store the transition (s, a, r, s') for policy and value updates.
      - Update the current state $s \leftarrow s'$.
3. **Return calculation:** For each trajectory of transitions compute the cumulative return G_t.

4. **Critic update:** Update the critic network by minimizing the loss function, which measures the discrepancy between the predicted value and the cumulative return (**??**).
5. **Actor update:** Optimize the actor network parameters using a policy gradient method with a trust region constraint to ensure stable updates (**??**).
6. **Training loop:** Repeat the process until convergence to an optimal policy that maximizes the long-term reward while respecting FOWT operational constraints.

The training is conducted episodically, with standardized initial conditions to enhance learning robustness. Each episode begins with a collective blade pitch angle set to 16.8773°, perturbed with random noise, and a platform pitch angle randomly initialized between 1.5° and 2.5°. Episodes last $T_f = 600\,\mathrm{s}$, with control actions executed at a fixed time step $T_s = 0.0125\,\mathrm{s}$. This setup introduces variability, improving the policy's ability to generalize across different initial conditions.

Table **??** outlines the specific settings employed during the training of the TRPO agent.

Table 2: Training Parameters for the DRL Controller.

Parameters	Value	Description
Total simulation time (T_f)	600 s	Duration of each training episode.
Time step (T_s)	0.0125 s	Frequency for updates to capture environment dynamics.
Discount factor (γ)	0.99	Ensures focus on long-term rewards.
Experience horizon (N)	512	Trajectory length for balancing immediately vs. future outcomes.
Entropy Loss weight (w)	0.01	Promotes exploration by preventing early convergence.
KL Divergence limit (δ)	0.01	Stabilizes policy by controlling policy update deviation.
Line Search Iterations (n)	10	Optimizes policy update step size.
Mini-batch size (M)	128	Sample size for each update.
Smoothing factor (GAE) (λ)	0.95	Manages bias variance in advantage estimation.
Critic learning rate (α_{critic})	0.001	Adjusts critic network weights.

3.3.4. Environmental conditions

The efficacy of the DRL controller depends on the Environmental Conditions (ECs) used during training, simulated with OpenFAST.

The key environmental factors considered are:

- **Wind speed range:** Training scenarios span wind speeds from the rated speed ($V_{rated} = 11.4\,\text{m/s}$) to the cut-out speed ($V_{cut-out} = 25\,\text{m/s}$) of the NREL 5 MW reference turbine, covering typical operational conditions in Region III.

- **Turbulence intensity:** To capture the stochastic nature of offshore wind, turbulent wind profiles are generated using NREL's TurbSim software [?] with the Kaimal turbulence model [?]. The Kaimal model, widely adopted for offshore environments, accurately simulates gusty and variable wind conditions [?].

- **Wave conditions:** Complex wave dynamics are modeled using the HydroDyn module in OpenFAST. The wave profiles are generated based on the Pierson-Moskowitz spectrum, a standard for fully developed sea states [**?**], and characterized by significant wave height and peak period.

A specific training environment, EC0, is defined to standardize the training setup. Table **??** summarizes its characteristics, and the corresponding wind and wave profiles are shown in Figure **??**. Both wind and wave directions are aligned along the downwind axis of the FOWT.

Table 3: Characteristics of Environmental Condition EC0

Parameter	**Value**
Simulated Wind	
Mean Wind Speed	18 m/s
Turbulence Intensity	15%
Wind Speed Range	10.72 to 25.32 m/s
Simulated Wave	
Significant Wave Height	1.2646 m
Peak Period	10 s
Wave Type	Irregular

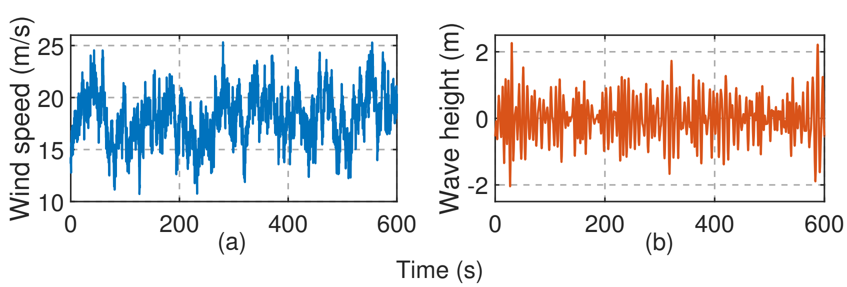

Figure 8: Wind speed (a) and wave height (b) profiles used in the simulation environment for training the agent.

3.4. Training Results

This subsection presents the outcomes of the training simulations using the DRL controller in the Matlab/Simulink environment coupled with OpenFAST. The turbine specifications are detailed in Table **??**, and the training parameters are summarized in Table **??**. For these simulations, all DoFs in OpenFAST were enabled, except for yaw motion, which was neglected. A constant generator torque strategy was adopted, fixing the generator torque at its rated value $T_{\text{gd}} = 43,093.55\,\text{N.m}$. The DRL controller applied a CBP strategy, uniformly adjusting all three blade pitch angles, as illustrated in Figure **??**.

Training was conducted on a workstation equipped with an 11^{th} Gen Intel® Core™ i7-11850H processor (8 cores, 2.50 GHz), with 6 cores allocated for parallel environment simulation. A total of 500 training episodes were executed. The agent's performance was evaluated every 10 episodes, focusing on generator speed tracking and platform pitch motion mitigation. Root Mean Square Error (RMSE) metrics were computed for both generator speed and platform pitch rate to quantitatively assess control accuracy and motion suppression. The best control policy emerged around episode 138, corresponding to approximately 6.6 million simulation steps. The total training time to reach this performance level was approximately 28 hours.

The performance of the trained DRL controller was compared to the Baseline GSPI controller, implemented as an external dynamic link library (.dll) within OpenFAST [**?**]. Table **??** summarizes the GSPI controller parameters.

Table 4: Baseline GSPI: Pitch Control Parameters.

Parameter	**Value**	**Description**
K_p	0.0063	Proportional gain for pitch controller at rated pitch (s)
K_i	0.0009	Integral gain for pitch controller at rated pitch (-)
K_k	0.11	Pitch angle where aerodynamic power derivative w.r.t. pitch doubles (rad)
β_{max}	1.57	Max pitch setting (rad)
β_{min}	0.0	Min pitch setting (rad)
β_{rate}	0.14	Max absolute blade pitch rate (rad/s)

3.4.1. Simulation results under training conditions

The results, shown in Figure **??**, include response curves for generator speed, platform pitch angle, platform pitch rate, and blade pitch angle under EC0 conditions. These variables serve as key indicators for evaluating control performance.

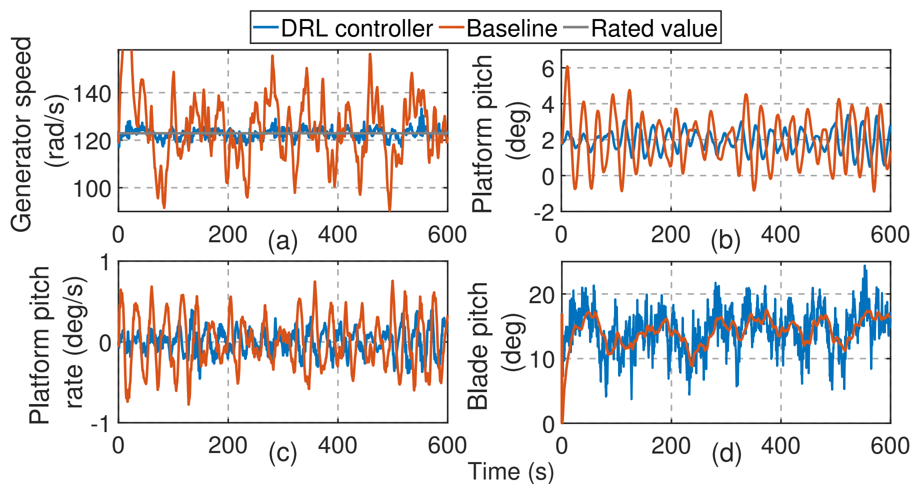

Figure 9: Training results of the DRL controller compared with the GSPI baseline: (a) Generator speed ω_g, (b) Platform pitch angle θ_y, (c) Platform pitch rate ω_y, and (d) Blade pitch angle β.

To quantitatively assess the controller's performance, Figure **??** presents statistical diagrams summarizing the mean, STandard Deviation (STD), and min-max values for rotor speed, platform pitch angle and rate, and blade pitch angle. Table **??** provides the RMSE values for generator speed and platform pitch rate. As RMSE is highly sensitive to errors in the data, it serves as a valuable metric for evaluating the accuracy of the trained controller. The statistical analysis covers the period from 100 seconds to 600 seconds to exclude the effects of initial conditions.

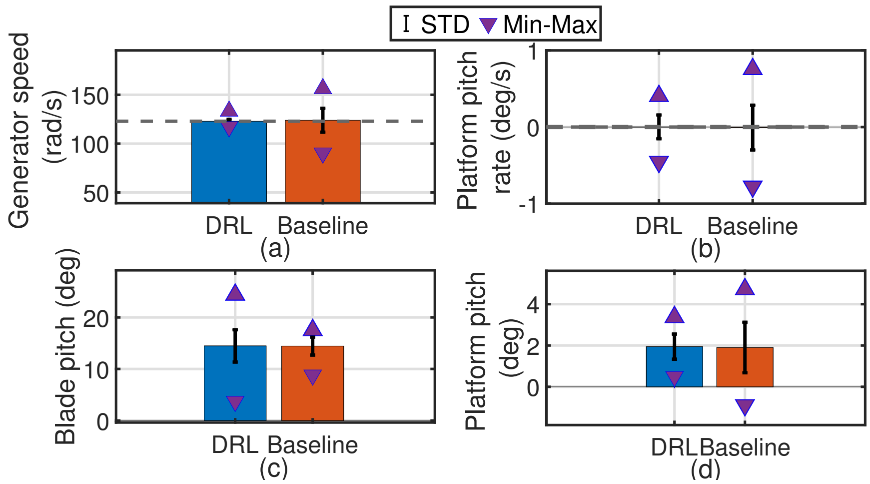

Figure 10: Statistical diagrams of mean, STD, and min-max values under EC0 for the DRL and Baseline controllers: (a) Generator speed ω_g, (b) Platform pitch rate ω_y, (c) Blade pitch angle β, (d) Platform pitch angle θ_y.

Table 5: RMSE for Genrator Speed (ω_g) and Platform Pitch Rate (ω_y).

Controller	**RMSE ω_g (rad/s)**	**RMSE ω_y (deg/s)**
DRL Controller	1.91043	0.16138
Baseline	12.3083	0.30103

3.4.2. Discussion

The performance evaluation of the DRL controller under EC0 conditions, as compared to the Baseline GSPI controller, provides several key insights into its effectiveness in managing both generator speed and platform stability.

Generator speed tracking: The results in Figure **??**(a) and the statistical summary in Figure **??**(a), as well as the RMSE values in Table **??** demonstrate that the DRL controller significantly outperforms the Baseline in terms of generator speed tracking. The DRL controller maintains a much lower RMSE value of 1.91 rad/s, compared to 12.31 rad/s for the Baseline. This suggests that the DRL approach is more effective at optimizing power generation, particularly under the high turbulence intensity conditions of EC0, by dynamically adjusting the blade pitch to maintain rated generator speed. Moreover, the reduced STD in the DRL controller's generator speed further underscores its ability to maintain more stable speed tracking with less variability than the Baseline.

Platform pitch angle and rate: The DRL controller also demonstrates superior performance in mitigating platform pitch motion, as shown in Figures **??**(b) and (c). Compared to the Baseline, the DRL controller achieves a noticeable reduction in both the pitch angle and rate, contributing to more stable platform behavior. The RMSE for the pitch rate is reduced from 0.30 deg/s (Baseline) to 0.16 deg/s (DRL), indicating better control of platform dynamics. Additionally, the lower STD of the platform pitch rate in Figure **??**(b) reinforces the DRL controller's ability to stabilize the platform more effectively.

Blade pitch angle behavior: A notable difference between the two controllers is the blade pitch angle behavior, as shown in Figures **??**(d) and **??**(c). The DRL controller exhibits a higher variation in the blade pitch angle, with a larger STD compared to the Baseline. This indicates that the DRL controller employs a more dynamic strategy, frequently adjusting the blade pitch to counteract environmental disturbances.

While this suggests a more aggressive control approach, the increased responsiveness in blade pitch adjustment contributes to better overall generator speed tracking and platform stability. However, it is important to note that excessive pitch variability may lead to increased wear on the blade actuators, potentially increasing maintenance costs over time.

4. Experimental Validation of the Trained DRL Controller

After completing the training process, the performance of the DRL controller was evaluated under novel wind and wave conditions, not encountered during training, trhough an experimental setup (Figure **??**). These tests assess the controller's ability to generalize across unseen scenarios.

To provide a comprehensive evaluation, the DRL controller's performance is also benchmarked against the Baseline GSPI controller. This comparative analysis highlights the advantages of the DRL-based approach while identifying any limitations under challenging operating conditions.

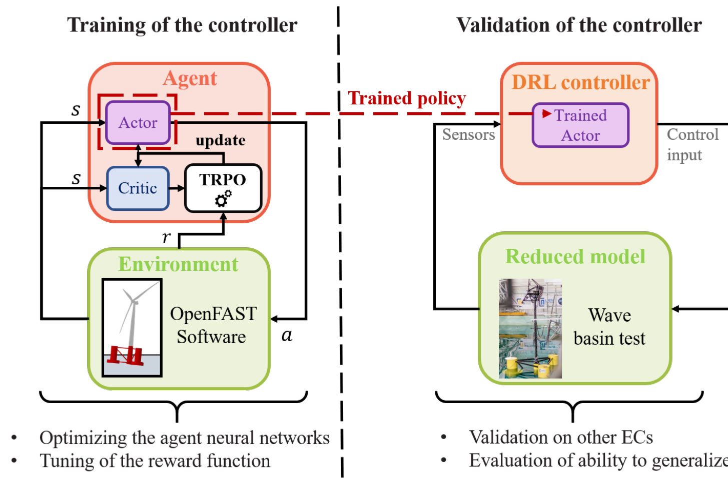

Figure 11: Implementation of the trained DRL controller for experimental validation.

4.1. Experimental Setup Description

The experimental validation was conducted at the LHEEA Laboratory (Laboratory of Hydrodynamics, Energetics, and Atmospheric Environment) at Centrale Nantes, France. This facility features advanced ocean engineering infrastructure, including a wave basin tailored for reduced-scale FOWT testing. Leroy et al. [?] performed an in-depth experimental analysis of the hydro-elastic response of a spar-type floating offshore wind turbine in such conditions. Building upon this, Bonnefoy et al. [?] proposed a hybrid SIL modeling approach to enhance the experimental assessment of FOWTs in wave tank environments. Additionally, Arnal [?] developed the experimental SIL framework used for real-time control validation, which was leveraged in the current study.

The experimental setup comprises a reduced-scale FOWT placed at the center of the wave basin, a Ni-compactRIO-9046 computer integrating both OpenFAST and the trained DRL controller, and a real-time control system. This control system simulate aerodynamic loads on the reduced model via force control, replicating the behavior of a full-scale FOWT under dynamic environmental conditions.

The wave basin, measuring 30 meters in width, 50 meters in length, and 5 meters in depth, is equipped with a wave generation system at one end and a wave absorption system at the other end. The wave generation system, composed of 48 hinged flaps, allows for the creation of multidirectional regular and irregular waves, characteristic of deep-water environments. On the opposite side, a wave absorption system minimizes reflections, ensuring clean wave profiles for accurate testing. Additionally, seven wave gauges installed within the basin provide precise measurements of wave elevation.

A 1/32 scale model of the NREL OC4 DeepCWind 5 MW semi-submersible FOWT was placed at the center of the basin (Figure **??**). This scaled model replicates the dynamics of the full-scale system and is equipped with the following sensors:

- Accelerometers located on the nacelle.
- Load cells with six measurement components are installed between the tower top and the nacelle, as well as between the platform and the tower, to capture forces and moments.
- A Qualisys Motion Tracking (QMT) system with four cameras to capture detailed position data of the scale model's floats and nacelle.

The model is anchored using four spring-loaded mooring lines that mimic the behavior of the three catenary mooring lines of the full-scale FOWT, ensuring realistic mooring forces. Wave gauges convert wave elevation data into voltage signals via an external acquisition system, while the QMT system continuously tracks the nacelle and platform positions, feeding real-time data to the NI-CompactRIO-9046 computer.

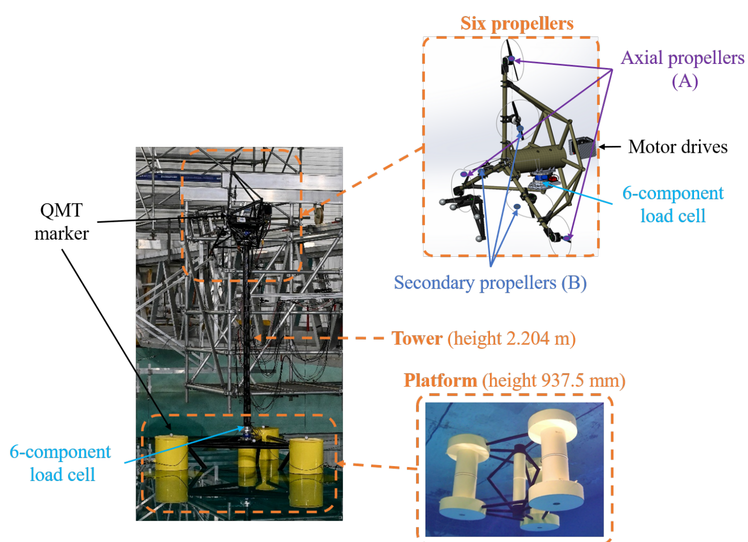

Figure 12: Scale model of the NREL OC4 DeepCWind 5 MW FOWT in the wave basin.

The experimental setup integrates the physical wave basin dynamics with OpenFAST for numerical simulations. The InflowWind module generates wind inflow, and AeroDyn calculates aerodynamic loads based on system states. The ServoDyn module manages turbine control (blade pitch and torque), while ElastoDyn computes platform motions and structural dynamics. Wave generation, platform motion, and mooring forces are physically modeled in the basin, while OpenFAST integrates real-time position data to simulate aerodynamic responses. At each simulation step, OpenFAST calculates rotor azimuth, rotational speed, and blade pitch angles, updating turbine states based on the imposed platform motions and tower deflection. For experimental deployment on the NI-compactRIO-9046, the trained DRL policy was exported from MATLAB/Simulink and re-implemented in `C++` as a shared library (`odiscon.so`), with configuration parameters defined in a `.yml` file. Thus, the external DRL controller, compiled as a shared library, computes the blade pitch angle and generator torque using both real-time measured data and simulated turbine states as inputs.

To address the challenge of scaling aerodynamic forces [?], caused by differences in the Reynolds number when scaling down wind turbines, the experiment employs a SIL approach. Instead of using a physical rotor, the aerodynamic forces computed in OpenFAST are applied using six propeller-based actuators (Figure ??). This configuration enables accurate reproduction of six-component aerodynamic forces with minimal actuator usage. Within the NI-compactRIO-9046 computer, the real-time control system of these propellers comprises two loops: an outer loop for OpenFAST calculations and an inner loop for actuator control based on a nonlinear SMC [?].

The outer loop receives the simulated wind profile and the real motion data of the FOWT's scale model from the QMT system via Ethernet cable, alongside blade pitch angle from the trained DRL controller. OpenFAST then computes the reference aerodynamic loads matrix M_{Ar}, which specifies the forces and moments to be applied by the six propellers.

The inner loop ensures that the actual loads M_{A} generated by the propellers closely follow this reference matrix M_{Ar}, using Pulse Width Modulation (PWM) signals sent to the electric motors driving the propellers. Two HBK MCS10 6D load cells are installed at the nacelle and tower base to measure multidimensional forces and moments. These load measurements are transferred to the NI-CompactRIO-9046 via a NI-9237 analog full-bridge input module, while control signals for the propeller motors are dispatched through a NI-9401 digital output module using PWM. This setup enables precise actuator control, allowing real-time aerodynamic forces to align seamlessly with the DRL control strategy.

Figure ?? presents a schematic diagram of this wave tank platform coupling framework.

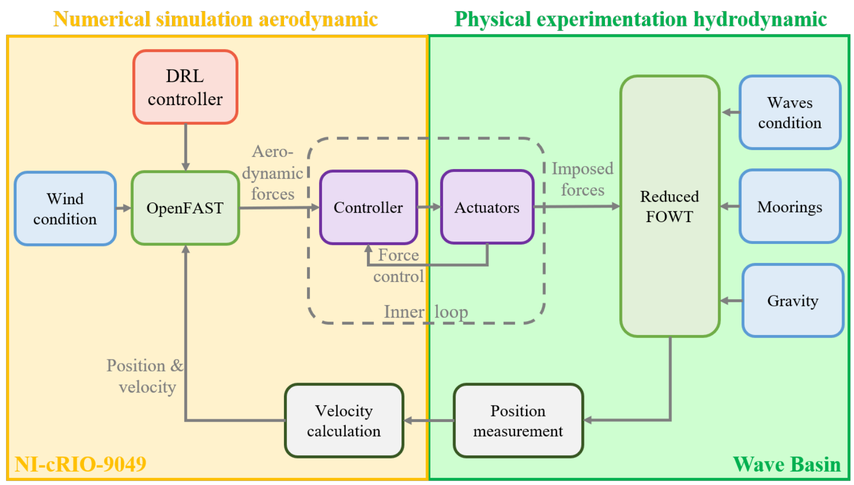

Figure 13: Schematic diagram of the wave tank platform SIL coupling framework.

4.2. Validation of the Experimental Setup

To ensure the experimental setup accurately mimics the hydrodynamic behavior of the OC4 DeepCWind semi-submersible 5 MW FOWT, a series of validation tests were conducted.

- **Dry tests:** Conducted to estimate the mass, the position of the center of gravity, and the inertia properties of the scale model.

- **Static pullout tests:** Performed to assess the stiffness of the mooring system, which is critical for accurately simulating the platform's stability and movements. The mooring equivalent linear stiffness around the equilibrium position on the surge axis for the entire mooring system was estimated after pullout tests as 77 N/m (at model scale).

- **Decay tests:** These tests measure the natural periods and damping coefficients, providing data on how the platform model's motion decays over time when perturbed. The hydrodynamic time-domain properties of the floating platform model were measured in all six DoFs. To achieve high accuracy, the connection stiffness and drag coefficients of the floating platform were meticulously calibrated. This calibration aligns the time-domain free-decay responses of the platform in the experimental setup with those predicted by OpenFAST simulations. This ensures that the mooring line behavior and other dynamic responses are correctly modeled. Figure **??** presents a comparison of the calibrated free-decay responses between the wave basin experiment and the OpenFAST simulation for surge, heave, and pitch motions of the platform. The specific initial conditions applied for the free-decay tests in still water (with no wind) are:

- **Surge:** Initial displacement of 25 cm.
- **Heave:** Initial displacement of 5 mm.
- **Pitch:** Initial amplitude of 6 degrees.

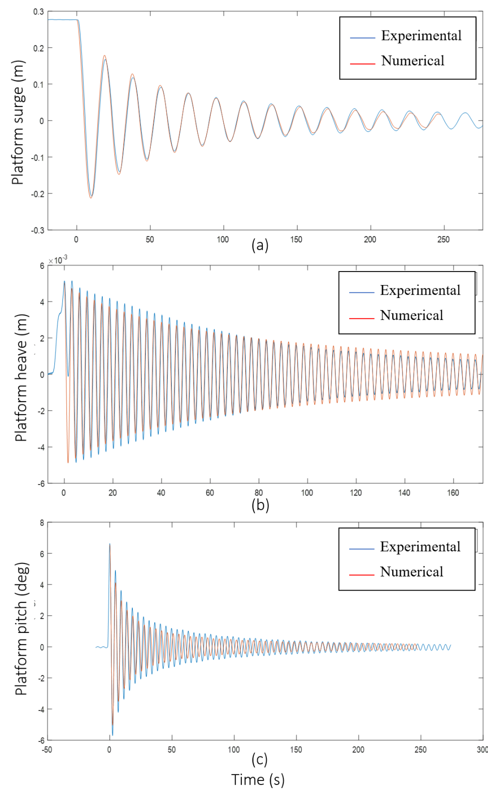

Figure 14: Comparison of free-decay responses between experimental setup and OpenFAST simulation in the surge, heave, and pitch DoFs: (a) Surge, (b) Heave, and (c) Pitch.

Table **??** gives the natural periods obtained with the reduced experimental OC4 platform in the wave basin compared to the OC4 specifications [?].

Table 6: Comparison of Natural Periods between OC4 Platform Specifications and Wave Tank Experiments.

DoF	OC4 FOWT specifications (s)	Wave basin experiments (s)
Heave	17	17.06
Surge	112	107.3
Pitch	25	26.20
Roll	25	26.16
Yaw	78	59.99
Sway	112	129.0

The observed periods for the heave DoF in the wave basin are closely aligned with the OC4 specifications, indicating an accurate replication of vertical dynamics. A slight deviation is noted for the surge period, with experimental results showing a shorter period than specified. This discrepancy may stem from differences in hydrodynamic modeling or environmental setup. Both pitch and roll exhibit minor variations but remain within an acceptable range, suggesting that rotational dynamics are adequately captured in the experimental setup. Larger deviations are observed in the yaw and sway periods, likely due to the specific mooring configuration employed in the experiments, as variations in mooring setups can significantly influence the platform's lateral and rotational movements.

Overall, the experimental results from the wave basin setup are in good agreement with the OC4 platform's specifications. This validation confirms the accuracy of the experimental setup in replicating the dynamic behavior of the semi-submersible FOWT. As a result, the experimental platform is validated as reliable for further testing and for validating the DRL controller under realistic operating conditions.

4.3. Validation of the Trained DRL Controller

The performance of the trained DRL controller was rigorously evaluated through a series of tests conducted in the wave basin experimental setup. Each test lasted 36 minutes and was structured into three distinct phases:

- **Initialization (2 minutes)**: OpenFAST is initialized with the initial aerodynamic loads and wind conditions for the FOWT. The actuators of the propellers on the scaled model are activated to reproduce the initial forces. The testing begins once the forces acting on the model stabilize.

- **Test phase (14 minutes)**: The DRL controller operates under specified wind and wave conditions, with a sampling interval of 0.0565685 seconds. Given the model's $1/32^{\text{nd}}$ scale, aerodynamic parameters within OpenFAST are adjusted using Froude scaling laws to ensure realistic outcomes. This effectively extends the 14-minute test duration to a full-scale equivalent of approximately 79.2 minutes.

- **Post-test (20 minutes)**: This phase allows the scaled FOWT model to settle back to a resting state after the active testing period.

To evaluate the robustness and adaptability of the DRL controller, three distinct ECs were selected for testing, with their characteristics specified in Table **??**. EC3 was specifically chosen to simulate conditions near the operational limits of Region III, providing a thorough assessment of the DRL controller's performance under extreme wind profiles. These experimental setups and conditions ensure that the DRL controller is tested not only for efficacy but also for its ability to adapt to the dynamic and sometimes extreme conditions typical of offshore environments.

Table 7: ECs used during the Validation of the DRL Controller (Wave Heights are presented at Full Scale).

EC	**Mean Wind Speed (m/s)**	**Significant Wave Height (m)**	**Wave Spectral Peak Period (s)**
EC1	14	7	12.0
EC2	20	7	12.0
EC3	30	3.5	12.0

4.3.1. Experimental results

The experimental results for EC1, EC2, and EC3 are presented in Figures **??**, **??**, and **??**, respectively, showing response curves for generator speed, platform pitch angle, platform pitch rate, and blade pitch angle. The trained DRL controller is compared to the Baseline GSPI controller, both of which were compiled in a shared library and called at each time step within the NI-compactRIO-9049 system, tested under the same experimental setup.

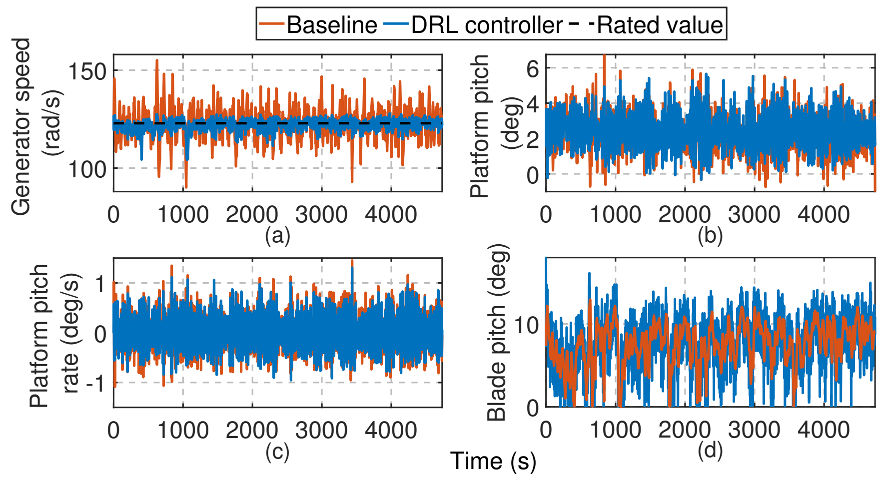

Figure 15: Experimental results comparing DRL controller (blue) and the Baseline (red) for EC1: (a) Generator speed ω_g, (b) Platform pitch angle θ_y, (c) Platform pitch rate ω_y, and (d) Blade pitch angle β.

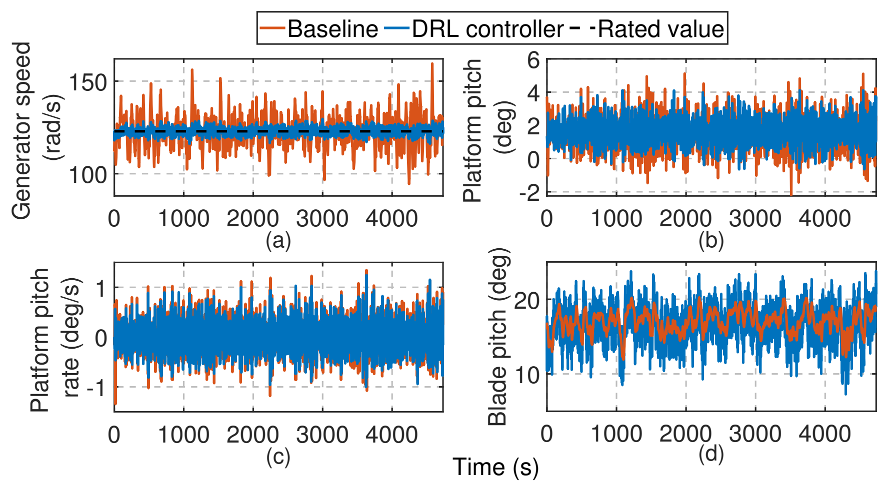

Figure 16: Experimental results comparing DRL controller (blue) and the Baseline (red) for EC2: (a) Generator speed ω_g, (b) Platform pitch angle θ_y, (c) Platform pitch rate ω_y, and (d) Blade pitch angle β.

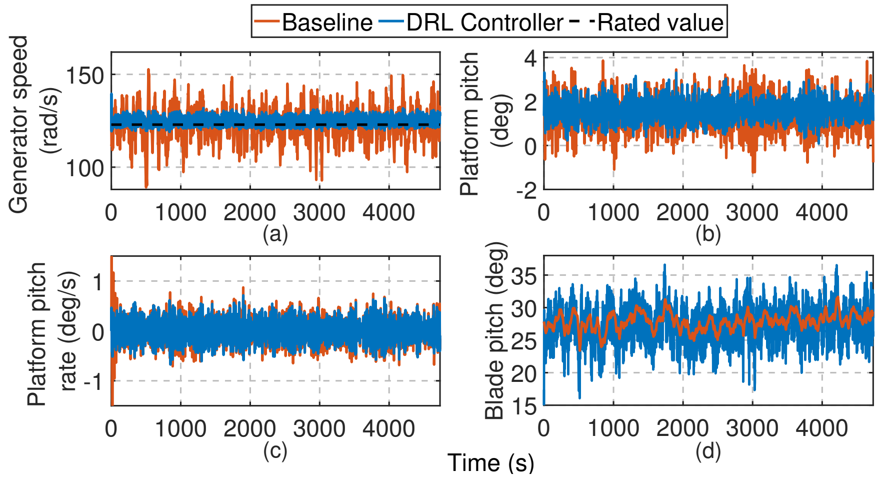

Figure 17: Experimental results comparing DRL controller (blue) and the Baseline (red) for EC3: (a) Generator speed ω_g, (b) Platform pitch angle θ_y, (c) Platform pitch rate ω_y, and (d) Blade pitch angle β.

Figure **??** presents statistical diagrams showing the mean, STD, and min-max values for generator speed, platform pitch angle and rate, and blade pitch angle across EC1, EC2, and EC3. Table **??** provides the RMSE values for generator speed and platform pitch rate. The first 100 seconds of experimental data are also excluded from the analysis to minimize the influence of initial conditions.

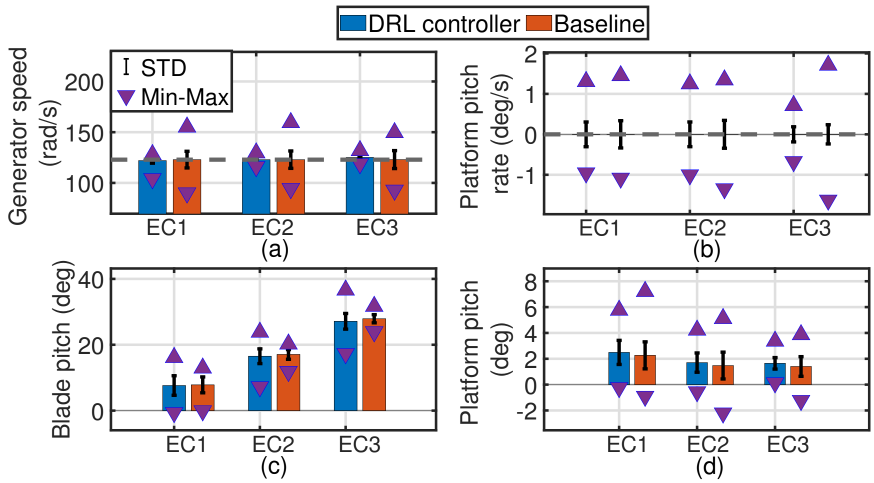

Figure 18: Statistical diagrams of mean, STD, and min-max values for EC1, EC2, and EC3: (a) Generator speed, (b) Platform pitch rate, (c) Blade pitch angle, and (d) Platform pitch angle.

Table 8: RMSE for Generator Speed ω_g and Platform Pitch Rate ω_y.

	Controller	**RMSE ω_g (rad/s)**	**RMSE ω_y (deg/s)**
EC1	DRL Controller	2.6984	0.3053
	GSPI	8.1879	0.3367
EC2	DRL Controller	1.7288	0.3050
	GSPI	8.5526	0.3481
EC3	DRL Controller	2.814	0.1869
	GSPI	8.8339	0.2245

The tower base side-to-side and fore-aft moments under EC1, EC2, and EC3 are shown in Figure **??**, Figure **??**, and Figure **??**, respectively. To highlight the simulation results, a quantitative analysis was conducted to assess the controller's impact on structural fatigue. Specifically, the fatigue Damage Equivalent Load (DEL) is used to characterize the bending moments. The DEL represents the equivalent load variation corresponding to the same damage level produced by a single load cycle, with the equivalent number of load cycles determined using the rainflow counting method. For the tower, a Wöhler exponent of 5 is applied [**?**].

To compare the tower base side-to-side and fore-aft moments under different ECs, the normalized DEL for each variable is calculated as:

$$\text{NormalizedDEL} = \frac{\text{DEL}}{\text{DEL}_b}, \tag{27}$$

where DEL_b is the DEL obtained using the Baseline controller for the respective variable.

The normalized DELs for the moments shown in Figures **??**, **??**, and **??** are presented in Figure **??**.

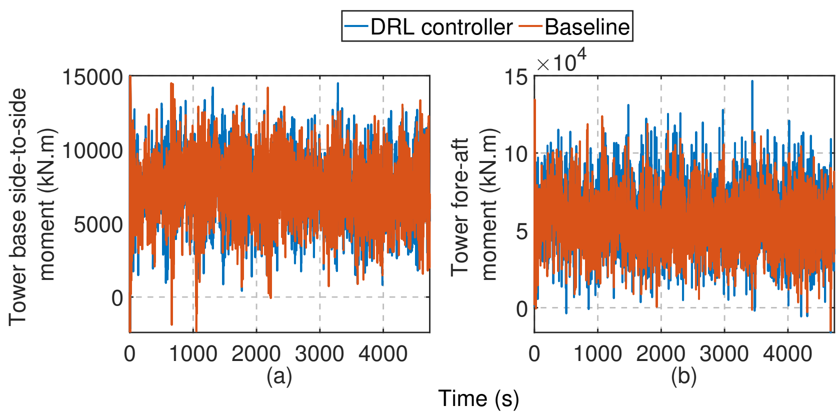

Figure 19: Bending moments under EC1 for the DRL and Baseline controllers: (a) Tower base side-to-side moment, (b) Tower base fore-aft moment.

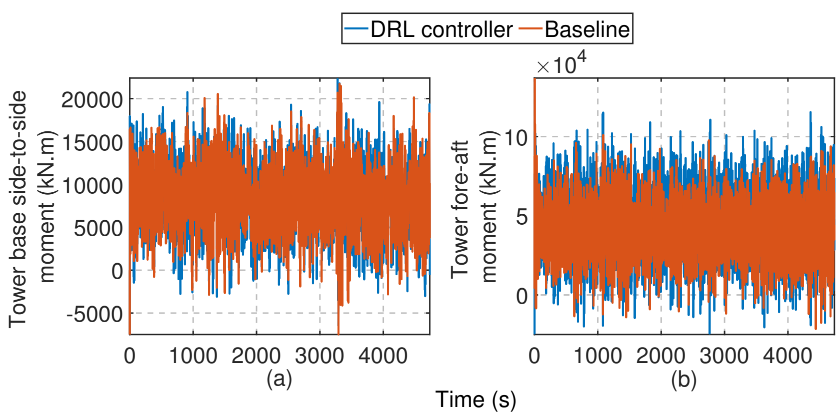

Figure 20: Bending moments under EC2 for the DRL and Baseline controllers: (a) Tower base side-to-side moment, (b) Tower base fore-aft moment.

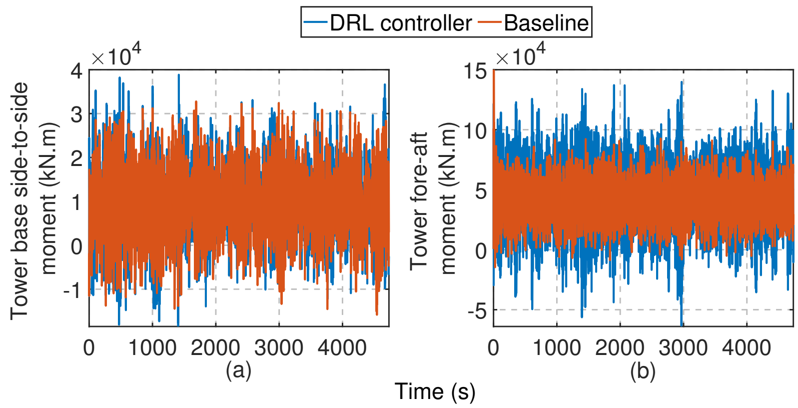

Figure 21: Bending moments under EC3 for the DRL and Baseline controllers: (a) Tower base side-to-side moment, (b) Tower base fore-aft moment.

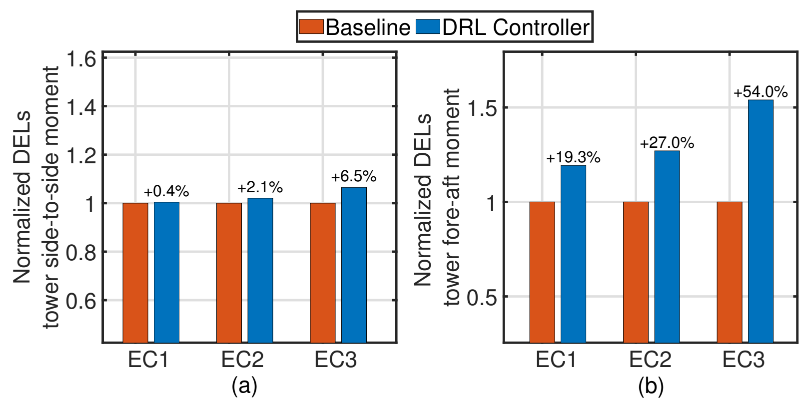

Figure 22: Normalized DELs for EC1, EC2 and EC3: (a) Tower base side-to-side moment, (b) Tower base fore-aft moment.

4.3.2. Discussion

Generator speed tracking: The generator speed response, illustrated in Figures **??**(a), **??**(a), and **??**(a), shows that the DRL controller maintains a narrower interval around the rated speed compared to Baseline GSPI controller across all ECs. In particular, under EC2 conditions, where the mean wind speed is 20 m/s, the DRL controller closely tracks the rated speed, achieving a much lower RMSE of 1.73 rad/s compared to 8.55 rad/s for the Baseline GSPI controller. This performance suggests that the DRL controller adapts well to wind speeds close to the wind conditions seen during training.

In EC3, which features more extreme wind conditions (mean wind speed of 30 m/s), the DRL controller maintains stability but shows a slight overshoot, with the generator speed consistently exceeding the rated value. This indicates that while the DRL controller can handle varying conditions, its ability to generalize under more extreme wind conditions is somewhat limited. This may be due to the relatively mild wind conditions encountered during training, as the maximum wind speed experienced by the DRL agent was 25.32 m/s.

Platform pitch angle and rate: The platform pitch angle and rate responses given in Figures **??** ((b), (c)), **??** ((b), (c)), and **??** ((b), (c)), indicate that the DRL controller reduces platform pitch angle and rate slightly more effectively than the Baseline GSPI controller. While the improvement is less pronounced than in generator speed tracking, the DRL controller demonstrates slightly better platform stability. In EC2, the DRL controller achieves an RMSE of 0.3050 deg/s for the platform pitch rate, compared to 0.3481 deg/s for the Baseline controller. This reduction in platform motion contributes to the structural integrity of the floating wind turbine, which is crucial for its long-term durability in offshore environments.

Although the reduction in platform pitch motion is not as significant as the improvements seen in the generator speed tracking, the DRL controller is still able to outperform the Baseline controller in mitigating platform dynamics.

Blade pitch angle behavior: As seen during the training phase, the DRL controller exhibits more dynamic adjustments in blade pitch angle, as illustrated Figures **??**(d), **??**(d), and **??**(d) for the blade pitch angle control. The DRL controller's strategy involves more frequent and larger blade pitch adjustments to counter environmental disturbances, which helps maintain generator speed and platform stability. However, this more aggressive control approach may lead to increased mechanical stress on the blade actuators and potentially higher maintenance costs over time.

Tower base moments: Across all environmental conditions (EC1, EC2, and EC3), the DRL controller's normalized DELs for the tower base side-to-side moments remain close to 1 (Figure **??**(a)), indicating its ability to maintain comparable fatigue loads to the baseline controller. For instance, under EC1, the normalized DEL is 1.004, reflecting only a +0.4% increase compared to the baseline. Similarly, under EC2 and EC3, the normalized DELs are 1.021 (+2.1%) and 1.065 (+6.5%), respectively. This slight increase in DEL, particularly in more severe environmental conditions (EC3), suggests that the DRL controller introduces additional minor structural fatigue, potentially due to its more aggressive control actions aimed at optimizing system performance.

In contrast, the fore-aft moments reveal a more pronounced increase in normalized DELs under all ECs, particularly in EC3 (Figure **??**(b)). For instance, under EC1, the normalized DEL reaches 1.193, representing a +19.3% increase compared to the baseline. This increase becomes more pronounced as platform pitching dynamics intensify.

Generalization across ECs: The DRL controller demonstrates strong performance in EC2, which has wind conditions similar to those used in training. In EC1, which has a lower wind speed mean (14 m/s), the DRL controller also generalizes well, though a minor deviation below the rated speed is observed. However, the extreme conditions of EC3 (30 m/s mean wind speed) pose a greater challenge to the DRL controller's generalization capabilities, as evidenced by the generator speed overshoot and increased blade pitch variation. Moreover, Figure **??** and Table **??** provide a quantitative comparison of the DRL and GSPI controllers. Across all ECs, the DRL controller exhibits lower RMSE values for both generator speed and platform pitch rate, demonstrating superior control accuracy. The statistical analysis further supports this, showing reduced standard deviations (STDs) for generator speed and platform pitch, which are more significant under EC2 (+27%) and EC3 (+54%).The substantial rise in the fore-aft DELs could stem from the controller's strategy to stabilize the platform while achieving generator speed tracking. These objectives may lead to higher loads in the fore-aft direction, where p rate with the DRL controller across EC1, EC2, and EC3. However, for blade pitch angle, the DRL controller shows higher STDs across all ECs, indicating more aggressive control inputs. The normalized DELs indicate that as conditions become more severe (EC3), the DRL controller tends to exert greater demands on the structure, especially in the fore-aft direction. This trend highlights a potential trade-off: while the DRL controller improves power regulation and system stability, it also introduces higher fatigue loads in some structural components.

Overall, the DRL controller shows good adaptability and robustness, outperforming the Baseline GSPI controller across a wide range of conditions typical of Region III. However, under more extreme conditions such as EC3, there is potential for a further refinement, particularly in terms of stabilizing generator speed and reducing structural fatigue. Strategies such as including a more explicit penalty for structural fatigue in the reward function or optimizing the control policy to balance performance and load reduction may mitigate these increases in DEL in future works.

5. Conclusion

This paper presented a fully data-driven, model-free Deep Reinforcement Learning (DRL) control strategy for regulating the collective blade pitch of a 5 MW semi-submersible Floating Offshore Wind Turbine (FOWT) in Region III. Using the Trust Region Policy Optimization (TRPO) algorithm within an actor-critic framework, the proposed controller effectively maintained rated generator speed and reduced platform pitch motion under varying offshore conditions. The controller was trained in a high-fidelity OpenFAST simulation environment and experimentally validated in a wave basin using a Software-In-the-Loop (SIL) approach, providing robust evidence of its performance benefits.

Experimental results demonstrated that the DRL controller outperforms the conventional Gain-Scheduling Proportional-Integral (GSPI) controller, particularly in generator speed regulation and platform stability. However, under extreme wind conditions, the controller exhibited performance degradation, characterized by generator speed overshoots and aggressive blade pitch variations, which could impact structural integrity over time.

This study highlights the feasibility and potential of DRL-based control strategies for FOWTs, offering improved adaptability and disturbance rejection without relying on explicit dynamic models. Future work will focus on enhancing the controller's generalization capabilities under extreme conditions by refining the reward function—especially to better manage the trade-off between control performance and structural fatigue—through automated tuning strategies such as Bayesian optimization, and by investigating advanced neural network architectures.

In addition to advancing model-free DRL-based pitch control, future research will explore hybrid control architectures that combine data-driven learning with model-based techniques to improve robustness and ensure stability. In particular, integrating DRL with nonlinear Sliding Mode Control (SMC) offers the potential to combine adaptability to unmodeled dynamics with formal stability guarantees. This line of work aims to develop hybrid control frameworks capable of safely and effectively managing the complex dynamics of floating wind turbines, especially under extreme offshore operating conditions.

Acknowledgments

This work was supported by the ANR Project (CREATIF, ANR-20-CE05-0039), the EIPHI Graduate School (contract ANR-17-EURE-0002) and the Region Bourgogne Franche-Comté.

References

[1] Intergovernmental Panel on Climate Change (IPCC), *Sixth Assessment Report*, Geneva, Switzerland, March 2023.

[2] Principle Power, *Kincardine Offshore Wind Farm*, https://www.principlepower.com/projects/kincardine-offshore-wind-farm.

[3] Equinor, *Hywind Tampen*, https://www.equinor.com/energy/hywind-tampen.

[4] BW Ideol, *EolMed Project*, https://www.bw-ideol.com/en/eolmed-project.

[5] International Renewable Energy Agency (IRENA), *Floating Offshore Wind Outlook*, Abu Dhabi, 2024.

[6] World Energy Council, *Global Offshore Wind Energy Projects*, https://www.worldenergy.org/impact-projects, 2023.

[7] Skaare, B., Hanson, T. D., and Nielsen, F. G., *Importance of Control Strategies on Fatigue Life of Floating Wind Turbines*, Volume 5: Ocean Space Utilization; Polar and Arctic Sciences and Technology; The Robert Dean Symposium on Coastal and Ocean Engineering; Special Symposium on Offshore Renewable Energy, International Conference on Offshore Mechanics and Arctic Engineering, pages 493-500, June 2007. doi: 10.1115/OMAE2007-29277.

[8] J.M. Jonkman, *Dynamics Modeling and Loads Analysis of an Offshore Floating Wind Turbine*, Technical Report, National Renewable Energy Laboratory (NREL), Report No. NREL/TP-500-41958, November 2007.

[9] T.J. Larsen and T.D. Hanson, *A method to avoid negative damped low frequent tower vibrations for a floating, pitch controlled wind turbine*, Journal of Physics: Conference Series, vol. 75, no. 1, pp. 012073, July 2007. doi: 10.1088/1742-6596/75/1/012073.

[10] J. Jonkman, S. Butterfield, W. Musial, and G. Scott, *Definition of a 5-MW Reference Wind Turbine for Offshore System Development*, Technical Report, National Renewable Energy Laboratory (NREL), Report No. NREL/TP-500-38060, Golden, Colorado, 2009. url: https://www.nrel.gov/docs/fy09osti/38060.pdf.

[11] J.M. Jonkman, *Influence of Control on the Pitch Damping of a Floating Wind Turbine*, Conference Paper, National Renewable Energy Laboratory (NREL), Report No. NREL/CP-500-42589, presented at the ASME Wind Energy Symposium, Reno, Nevada, January 7–10, 2008. url: http://www.osti.gov/bridge.

[12] H. Namik and K. Stol, *Individual blade pitch control of floating offshore wind turbines*, Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology, vol. 13, no. 1, pp. 74–85, 2010. Wiley Online Library.

[13] H. Namik and K. Stol, *Performance analysis of individual blade pitch control of offshore wind turbines on two floating platforms*, Mechatronics, vol. 21, no. 4, pp. 691–703, 2011. Elsevier.

[14] H. Namik and K. Stol, *Individual blade pitch control of a spar-buoy floating wind turbine*, IEEE Transactions on Control Systems Technology, vol. 22, no. 1, pp. 214–223, 2013. IEEE.

[15] S. Christiansen, T. Knudsen, and T. Bak, *Optimal control of a ballast-stabilized floating wind turbine*, in Proceedings of the 2011 IEEE International Symposium on Computer-Aided Control System Design (CACSD), pp. 1214–1219, 2011. IEEE.

[16] S. Christiansen, T. Knudsen, and T. Bak, *Extended onshore control of a floating wind turbine with wave disturbance reduction*, Journal of Physics: Conference Series, vol. 555, no. 1, pp. 012018, 2014. IOP Publishing.

[17] F. Lemmer, D. Schlipf, and P.W. Cheng, *Control design methods for floating wind turbines for optimal disturbance rejection*, Journal of Physics: Conference Series, vol. 753, no. 9, pp. 092006, 2016. IOP Publishing.

[18] O. Bagherieh and R. Nagamune, *Gain-scheduling control of a floating offshore wind turbine above rated wind speed*, Control Theory and Technology, vol. 13, no. 2, pp. 160–172, 2015. Springer.

[19] P. Zhao and R. Nagamune, *Switching LPV control of a floating offshore wind turbine on a semi-submersible platform*, in Proceedings of the 2019 IEEE 28th International Symposium on Industrial Electronics (ISIE), pp. 664–669, 2019. IEEE.

[20] T. Bakka and H.R. Karimi, *Robust output feedback H-infinity control synthesis with pole placement for offshore wind turbine system: An LMI approach*, in Proceedings of the 2012 IEEE International Conference on Control Applications, pp. 1467–1472, 2012. IEEE.

[21] X. Li and H. Gao, *Load mitigation for a floating wind turbine via generalized H_∞ structural control*, IEEE Transactions on Industrial Electronics, vol. 63, no. 1, pp. 332–342, 2015. IEEE.

[22] C.J. Cortes Sanchez, *Wind and wave disturbance rejection control of floating offshore wind turbines*, Master's Thesis, University of British Columbia, 2018.

[23] T. Bakka, H.R. Karimi, and N.A. Duffie, *Gain Scheduling for Output H8 Control of Offshore Wind Turbine*, in Proceedings of the ISOPE International Ocean and Polar Engineering Conference, pp. ISOPE–I, 2012. ISOPE.

[24] N. Hara, Y. Nihei, K. Iijima, and K. Konishi, *Blade pitch control for floating wind turbines: Design and experiments using a scale model*, in Proceedings of the 2017 IEEE Conference on Control Technology and Applications (CCTA), pp. 481–486, 2017. IEEE.

[25] M.S. Mahmoud and M.O. Oyedeji, *Adaptive and predictive control strategies for wind turbine systems: A survey*, IEEE/CAA Journal of Automatica Sinica, vol. 6, no. 2, pp. 364–378, 2019, doi: 10.1109/JAS.2019.1911375.

[26] F. Lemmer, S. Raach, D. Schlipf, and P.W. Cheng, *Prospects of linear model predictive control on a 10 MW floating wind turbine*, in International Conference on Offshore Mechanics and Arctic Engineering, vol. 56574, pp. V009T09A071, 2015, American Society of Mechanical Engineers.

[27] A. Cunha, E. Caetano, P. Ribeiro, and G. Müller, *Reducing blade fatigue and damping platform motions of floating wind turbines using model predictive control*, in International Conference on Structural Dynamics, 2014.

[28] Y. Okada, K. Haneda, T. Chujo, and T. Ohtsuka, *Parameter-varying Modeling and Nonlinear Model Predictive Control for Floating Offshore Wind Turbines*, IFAC-PapersOnLine, vol. 52, no. 16, pp. 382–387, 2019, doi: https://doi.org/10.1016/j.ifacol.2019.11.810.

[29] T. Wakui, A. Nagamura, and R. Yokoyama, *Stabilization of power output and platform motion of a floating offshore wind turbine-generator system using model predictive control based on previewed disturbances*, Renewable Energy, vol. 173, pp. 105–127, 2021, doi: https://doi.org/10.1016/j.renene.2021.03.112.

[30] D. Schlipf, L.Y. Pao, and P.W. Cheng, *Comparison of feedforward and model predictive control of wind turbines using LIDAR*, 2012 IEEE 51st Conference on Decision and Control (CDC), pp. 3050–3055, 2012, url: https://api.semanticscholar.org/CorpusID:15017519.

[31] D. Schlipf, F. Sandner, S. Raach, D. Matha, and P.W. Cheng, *Nonlinear model predictive control of floating wind turbines*, in ISOPE International Ocean and Polar Engineering Conference, ISOPE-I, 2013.

[32] D. Schlipf, D.J. Schlipf, and M. Kühn, *Nonlinear model predictive control of wind turbines using LIDAR*, Wind Energy, vol. 16, no. 7, pp. 1107–1129, 2013.

[33] S. Raach, D. Schlipf, F. Sandner, D. Matha, and P.W. Cheng, *Nonlinear model predictive control of floating wind turbines with individual pitch control*, 2014 American Control Conference, pp. 4434–4439, 2014, doi: 10.1109/ACC.2014.6858718.

[34] K.A. Shah, Y. Li, R. Nagamune, Y. Zhou, and W.U. Rehman, *Platform motion minimization using model predictive control of a floating offshore wind turbine*, Theoretical and Applied Mechanics Letters, vol. 11, no. 5, pp. 100295, 2021, doi: https://doi.org/10.1016/j.taml.2021.100295.

[35] O. Bagherieh, K. Hedrick, and R. Horowitz, *Nonlinear Control of Floating Offshore Wind Turbines Using Input/Output Feedback Linearization and Sliding Control*, Proceedings of the ASME 2014 Dynamic Systems and Control Conference, Volume 2, pp. V002T18A004, 2014, doi: 10.1115/DSCC2014-5982.

[36] C. Zhang, E. Tahoumi, S. Gutierrez, F. Plestan, and J. DeLeón-Morales, *Adaptive robust control of floating offshore wind turbine based on sliding mode*, 2019 IEEE 58th Conference on Decision and Control (CDC), pp. 6936–6941, 2019, doi: 10.1109/CDC40024.2019.9029231.

[37] C. Zhang, S.V. Gutierrez, F. Plestan, and J. de León-Morales, *Adaptive super-twisting control of floating wind turbines with collective blade pitch control*, IFAC-PapersOnLine, vol. 52, no. 4, pp. 117–122, 2019, doi: https://doi.org/10.1016/j.ifacol.2019.08.165.

[38] Y. Shtessel, M. Taleb, and F. Plestan, *A novel adaptive-gain supertwisting sliding mode controller: Methodology and application*, Automatica, vol. 48, no. 5, pp. 759–769, 2012, doi: https://doi.org/10.1016/j.automatica.2012.02.024.

[39] C. Zhang and F. Plestan, *Adaptive sliding mode control of floating offshore wind turbine equipped by permanent magnet synchronous generator*, Wind Energy, vol. 24, no. 7, pp. 754–769, 2021, doi: https://doi.org/10.1002/we.2601.

[40] M. Taleb, A. Marie, C. Zhang, M.A. Hamida, and P.E. Testelin, *Adaptive nonlinear control of floating wind turbines: new adaptation law and comparison*, IECON 2021 – 47th Annual Conference of the IEEE Industrial Electronics Society, pp. 1–6, 2021, doi: 10.1109/IECON48115.2021.9589421.

[41] M. Taleb and F. Plestan, *Adaptive supertwisting controller with reduced set of parameters*, 2021 European Control Conference (ECC), pp. 2627–2632, 2021, doi: 10.23919/ECC54610.2021.9655180.

[42] National Renewable Energy Laboratory, *OpenFAST Documentation, Release 3.4.0*, 2023, pp. 6–8, https://openfast.readthedocs.io/en/main/.

[43] Hedi Basbas, Yong-Chao Liu, Salah Laghrouche, Mickaël Hilairet, and Franck Plestan, *Review on Floating Offshore Wind Turbine Models for Nonlinear Control Design*, Energies, Volume 15, Article 5477, 2022. DOI: 10.3390/en15155477.

[44] Betti, Giulio, Farina, Marcello, Guagliardi, Giuseppe A., Marzorati, Andrea, and Scattolini, Riccardo, *Development of a Control-Oriented Model of Floating Wind Turbines*, IEEE Transactions on Control Systems Technology, vol. 22, no. 1, pp. 69–82, 2013.

[45] Frank Lemmer, *Low-Order Modeling, Controller Design and Optimization of Floating Offshore Wind Turbines*, Doctoral Thesis, University of Stuttgart, 2018.

[46] Homer, Jeffrey R. and Nagamune, Ryozo, *Physics-Based 3-D Control-Oriented Modeling of Floating Wind Turbines*, IEEE Transactions on Control Systems Technology, vol. 26, no. 1, pp. 14–26, 2017.

[47] Hedi Basbas, Hussein Obeid, Salah Laghrouche, Mickael Hilairet, and Franck Plestan, *Barrier Function Based-Adaptive Super-Twisting Algorithm for Floating Offshore Wind Turbine*, 2022 16th International Workshop on Variable Structure Systems (VSS), IEEE, pp. 166–171, 2022.

[48] Yong-Chao Liu, Hedi Basbas, and Salah Laghrouche, *Robust blade pitch control of semi-submersible floating offshore wind turbines based on the modified super-twisting sliding-mode algorithm*, Journal of the Franklin Institute, 2024, Article 107279, ISSN 0016-0032, DOI: 10.1016/j.jfranklin.2024.107279.

[49] F. Didier, Y.-C. Liu, S. Laghrouche, and D. Depernet, *Radial Basis Function Neural Network-Based Super-Twisting Blade Pitch Controller for the Floating Offshore Wind Turbine*, 10th International Conference on Control, Decision and Information Technologies (CoDIT), IEEE, Valletta, Malta, 2024.

[50] F. Didier, H. Obeid, Y. Chitour, L. Fridman, and S. Laghrouche, *Adaptive Neural Network-Based Higher-Order Sliding Mode Control for Floating Offshore Wind Turbines*, 17th International Workshop on Variable Structure Systems (VSS), Abu Dhabi, UAE, 2024.

[51] Hou, Z.-S., and Wang, Z., *From model-based control to data-driven control: Survey, classification and perspective*, Information Sciences, vol. 235, 2013, pp. 3–35.

[52] Zhao, Shuai, Blaabjerg, Frede, and Wang, Huai, *An Overview of Artificial Intelligence Applications for Power Electronics*, IEEE Transactions on Power Electronics, vol. 36, no. 4, pp. 4633-4658, 2021. doi: 10.1109/TPEL.2020.3024914.

[53] Kane, Michael B., *Machine Learning Control for Floating Offshore Wind Turbine Individual Blade Pitch Control*, 2020 American Control Conference (ACC), pp. 237-241, 2020. doi: 10.23919/ACC45564.2020.9147912.

[54] Roh, Chan, *Deep-Learning-Based Pitch Controller for Floating Offshore Wind Turbine Systems with Compensation for Delay of Hydraulic Actuators*, Energies, vol. 15, no. 9, article 3136, 2022. doi: 10.3390/en15093136.

[55] Xie, Jingjie, Dong, Hongyang, and Zhao, Xiaowei, *Power Regulation and Load Mitigation of Floating Wind Turbines via Reinforcement Learning*, IEEE Transactions on Automation Science and Engineering, vol. 21, no. 3, pp. 4328-4339, 2024. doi: 10.1109/TASE.2023.3295576.

[56] Chen, Peng, Chen, Jiahao, and Hu, Zhiqiang, *Software-in-the-Loop Combined Reinforcement Learning Method for Dynamic Response Analysis of FOWTs*, Frontiers in Marine Science, vol. 7, 2021. doi: 10.3389/fmars.2020.628225.

[57] Chen, Peng and Hu, Zhi Qiang, *A Study on Key Disciplinary Parameters of Artificial Intelligent-Based Analysis Method for Dynamic Response Prediction of Floating Offshore Wind Turbines*, Journal of Offshore Mechanics and Arctic Engineering, vol. 145, no. 1, article 010906, 2022. doi: 10.1115/1.4055993.

[58] Chen, P., Song, L., Chen, J.-h., and Hu, Z., *Simulation Annealing Diagnosis Algorithm Method for Optimized Forecast of the Dynamic Response*

of Floating Offshore Wind Turbines, Journal of Hydrodynamics, vol. 33, no. 2, pp. 216-225, 2021. doi: 10.1007/s42241-021-0033-9.

[59] Robertson, A., *Definition of the Semisubmersible Floating System for Phase II of OC4*, NREL Technical Report, NREL/TP-5000-60601, 2014.

[60] Vidal, Y., Acho, L., Luo, N., Zapateiro, M., and Pozo, F., *Power Control Design for Variable-Speed Wind Turbines*, Energies, Vol. 5, No. 8, pp. 3033–3050, 2012. DOI: 10.3390/en5083033.

[61] Samani, A.E., De Kooning, J.D.M., Kayedpour, N., Singh, N., and Vandevelde, L., *The Impact of Pitch-To-Stall and Pitch-To-Feather Control on the Structural Loads and the Pitch Mechanism of a Wind Turbine*, Energies, Vol. 13, No. 17, Article 4503, 2020. DOI: 10.3390/en13174503.

[62] Puterman, M.L., *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, Wiley Series in Probability and Mathematical Statistics, Wiley, New York, 1994. ISBN: 978-0-471-61977-2.

[63] Schulman, J., Levine, S., Abbeel, P., Jordan, M.I., and Moritz, P., *Trust Region Policy Optimization*, ArXiv, vol. abs/1502.0547, Feb. 2015.

[64] Kelley, N., and Jonkman, B., *TurbSim User's Guide: Version 1.5*, NREL Report, 2009.

[65] Kaimal, J.C., Wyngaard, J.C., Izumi, Y., and Coté, O.R., *Spectral characteristics of surface-layer turbulence*, Quarterly Journal of the Royal Meteorological Society, vol. 98, no. 417, pp. 563-589, 1972. DOI: https://doi.org/10.1002/qj.49709841707.

[66] Krieger, A., Ramachandran, G.K.V., Vita, L., Alonso, P.G., and Almería, G.G., *LIFES50+ Deliverable: D7.2 Design Basis*, DNV GL, Project 640741, 2015.

[67] Pierson Jr., W.J. and Moskowitz, L., *A proposed spectral form for fully developed wind seas based on the similarity theory of S.A. Kitaigorodskii*, Journal of Geophysical Research (1896-1977), vol. 69, no. 24, pp. 5181-5190, 1964. DOI: https://doi.org/10.1029/JZ069i024p05181.

[68] Leroy, V., Delacroix, S., Merrien, A., Bachynski-Polić, E.E., and Gilloteaux, J.-C., *Experimental investigation of the hydro-elastic response of a spar-type floating offshore wind turbine*, Ocean Engineering, vol. 255, 2022, pp. 111430. DOI: https://doi.org/10.1016/j.oceaneng.2022.111430.

[69] Bonnefoy, F., Leroy, V., Mojallizadeh, M.R., Delacroix, S., Arnal, V., and Gilloteaux, J.-C., *Multidimensional hybrid software-in-the-loop modeling approach for experimental analysis of a floating offshore wind turbine in wave tank experiments*, Ocean Engineering, vol. 309, 2024, pp. 118390. DOI: https://doi.org/10.1016/j.oceaneng.2024.118390.

[70] Arnal, V., *Experimental modelling of a floating wind turbine using a "software-in-the-loop" approach*, PhD thesis, École Centrale de Nantes, France, 2020. URL: https://theses.hal.science/tel-03237441.

[71] Meng, F., Lio, A.W.H., and Bredmose, H., *Challenges and Perspectives in Experimental Study of Floating Offshore Wind Turbine Control: Insights from Recent Research*, IEEE Control Systems, vol. 44, no. 5, 2024, pp. 58-62. DOI: 10.1109/MCS.2024.3432286.

[72] Mojallizadeh, M.R., Bonnefoy, F., Plestan, F., Hamida, M.A., and Ohana, J., *Euler implicit time-discretization of multivariable sliding-mode controllers*, ISA Transactions, vol. 147, 2024, pp. 140-152. DOI: https://doi.org/10.1016/j.isatra.2024.01.031.

[73] Z. Cheng, H. A. Madsen, W. Chai, Z. Gao, and T. Moan, *A comparison of extreme structural responses and fatigue damage of semi-submersible type floating horizontal and vertical axis wind turbines*, Renewable Energy, vol. 108, pp. 207–219, 2017, doi: 10.1016/j.renene.2017.02.067.