UMERC+OREC
2025 Conference

12-14 August | Corvallis, OR USA

Experimental Validation and Analysis of Deep Reinforcement Learning Control for Wave Energy Converters

Abishek Subramanian[a], Shangyan Zou[a1], Bret Bosma[b], Junhui Lou[b], Courtney Beringer[b], Bryson Robertson[b], Pedro Lomonaco[b]

[a] *Michigan Technological University, 1400 Townsend Drive, Houghton, 49931, USA*
[b] *Oregon State University, 1600 SW Pioneer Place, Corvallis, 97331 , USA*

Abstract

This paper presents the real-time implementation and experimental validation of a Deep Reinforcement Learning (DRL) control strategy for a Wave Energy Converter (WEC). Wave energy technology is still underdeveloped. One important challenge is the inefficient Power Take-Off (PTO) control technologies. To address this challenge, the research team has developed a DRL control which shows promising performance in wave power production and outperforms existing model-based controls on paper. However, the practical performance of this control remains unknown. Therefore, the team at Michigan Technological University (MTU) collaborated with Oregon State University (OSU) in testing the control performance in the wave flume. The Laboratory Upgrade Point Absorber (LUPA) WEC is utilized for control demonstration. Varied wave tank tests have been conducted, including both regular and irregular wave tests. The testing results suggest: (1) the real-time implementation of the DRL control is relatively straightforward with a very small computational cost; (2) the robustness of the control is a challenging problem in practice, which can be addressed by incorporating significant randomness in training; (3) the control shows promising performance in practice. Specifically, the power produced in regular waves is close to the practical maximum, and the power produced in irregular waves outperforms an existing control.

Keywords: WEC Control; Deep Reinforcement Learning; Real-time experimental testing

---

[1] ∗ Corresponding author. E-mail address: shangyan@mtu.edu

1 Introduction

In recent years, wave energy has gained increasing attention as a promising renewable energy source due to its high energy density, predictability, and vast potential [1]. However, numerous technological and economic challenges need to be addressed before Wave Energy Converters (WECs) can become commercially viable [2]. A critical barrier among these challenges remains the development of an efficient PTO control strategy to maximize the generation of electricity from wave resources [3]. Conventional designs of the PTO control typically apply reduced-order models (e.g., linearized hydrodynamics) without considering other subsystems (e.g., the dynamics of the PTO unit) [4, 5]. Although these controls demonstrate promising performance in energy harvesting, this may be misleading [6, 7]. On the other hand, directly developing a model-based control that considers all the necessary subsystems is very cumbersome.

Recent advancements in machine learning shed new light on this challenging problem. In particular, the Deep Reinforcement Learning (DRL) approach is well-suited for WEC control problems [8]. This is due to the unknown optimality of the WEC system, especially when considering nonlinearities and uncertainties, which can be identified by DRL-based control through direct interaction with the environment. In addition, unlike model-based controls, DRL control does not require precise knowledge of the system dynamics (model-free), and therefore makes it possible to optimize WEC performance from a global perspective, taking necessary subsystems into account [9]. The research team has previously developed a DRL-based control for WEC and demonstrated promising performance in terms of wave power production, outperforming multiple state-of-the-art model-based controls [10]. Despite the performance shown on paper, the practical performance of the DRL control remains unknown.

To address this gap, this study focuses on presenting the practical performance of the DRL control. The main research questions we are trying to answer are: (1) What is the practical performance of the DRL control in terms of power production, adaptivity, robustness, computational speed, and losses? and (2) What are the challenges and limitations of practically implementing DRL control? The research team at MTU collaborated with the team at OSU on the wave tank testing of the DRL control with support from TEAMER. The experimental campaign took place at the O.H. Hinsdale Wave Research Laboratory's Large Wave Flume. The LUPA WEC is used as the platform for control implementation and demonstration (the single-body, heave-only configuration is adopted) [11]. The control is validated with both regular and irregular waves and compared with an Optimal Feedback Control (OFC) provided by the facility (with optimal feedback gains obtained by directly sweeping a range of values during tank testing) [12].

The rest of the paper is organized as follows. Section II introduces the detailed methodology of DRL control implementation. Section III covers the discussion of the results. Finally, the conclusion is drawn in Section IV.

2 Methodology

In this section, we will introduce the numerical modeling of LUPA and the DRL control.

2.1 Numerical Model

The LUPA WEC is a point absorber wave energy converter capable of operating in three configurations: a single-float heaving system, a two-body heaving system, and a two-body system with six degrees of freedom. The device is composed of two bodies, including a float and a heaving plate (Fig. 1 (a) and (b)). In this study, only the single-body heaving configuration is applied, and the heaving plate is simply locked in place to limit movement. In the presence of incoming waves, the float is designed to be excited and "follow" the water surface, while the heave plate remains more static. The PTO system captures the relative motion between the float and the heave plate to generate electrical power.

Table 1: LUPA WEC key device parameters

Specification	Value
Mass	436 kg
Float Diameter	1 m
PTO stroke	0.5m
Moment of Inertia (X)	66.17 kg.m^2
Moment of Inertia (Y)	65.33 kg.m^2
Moment of Inertia (Z)	17.16 kg.m^2
Quadratic Drag Coefficient	0.15

The power electronics of the LUPA are housed in a compartment (Fig. 1 (c)) within the floating body, serving both data acquisition and control of the PTO unit. The PTO unit, depicted in Fig. 1 (d), translates the linear relative motion

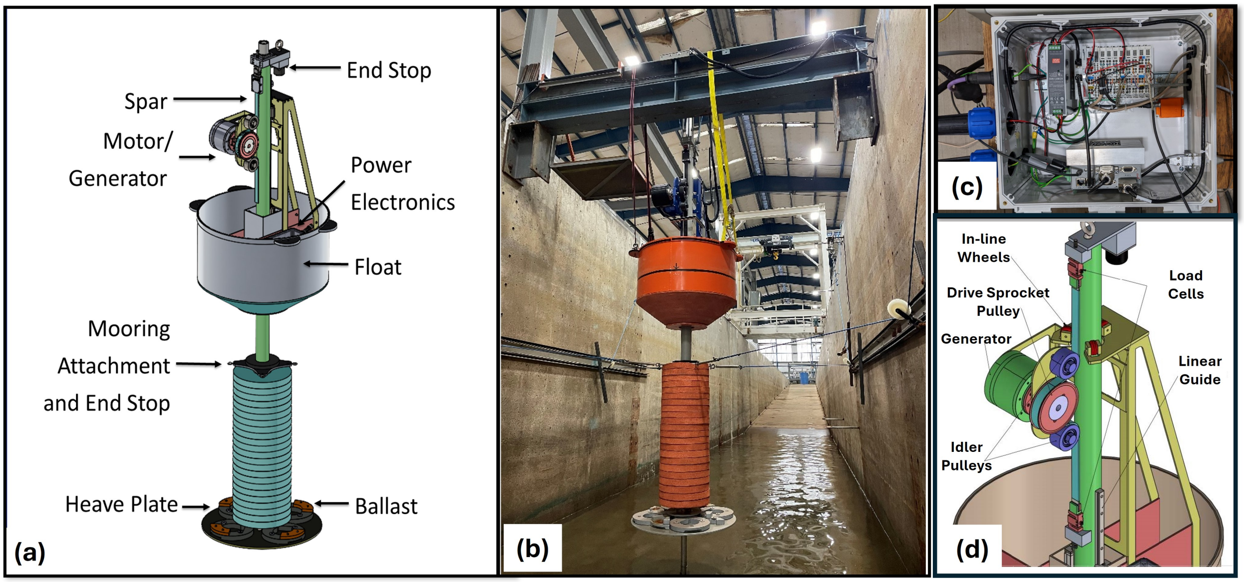

Figure 1: LUPA Hardware Setup

between the two bodies of LUPA into rotational motion through a pulley system, which then drives the generator/motor. This motor is controlled by a motor drive that applies the desired control force.

The numerical model of LUPA is developed in WECsim [12], where the hydrodynamics can be expressed as:

$$(m_r + m_\infty)\ddot{z} = F_{ex} + F_r + F_{hs} + F_{drag} + F_{PTO} + F_{loss} \tag{1}$$

where z represents the heave displacement, F_{ex} is the excitation force, F_r is the force due to radiation damping, F_{hs} is the hydrostatic force F_{drag} is the quadratic drag force and F_{loss} represents losses in the physical system that include the mechanical friction and electromagnetic damping/inertia in the motor. Moreover, m_r and m_∞ represent the rigid body mass and the added mass at infinity frequency, respectively. The PTO loss force is modelled as a linear damping:

$$F_{loss} = -C_{ld}\dot{z} \tag{2}$$

It is noted that the loss coefficient in the pretraining of the DRL control is initially assumed to be zero (in other words, this force is not considered). However, it is later found during testing that there is a significant model mismatch due to mechanical losses. Therefore, this force is necessary to be added to the numerical model for improved control training, and the damping coefficient is calibrated according to the testing data. More details of this force are discussed in Section 3.1. The key physical parameters of the LUPA WEC are summarized in Table 1.

2.2 DRL Control Algorithm

Next, the DRL control is developed for LUPA to maximize wave power production. The complete control framework is presented in Fig. 2. In general, the DRL agent collects real-time rewards and states to train the action-value function (Q function). Then, the next optimal action is determined by maximizing the Q function. It is noted that the action generated by the DRL agent is discrete, whereas the control of a WEC is typically continuous. Therefore, the action of the DRL agent is designed to be the variation of the gains of a time-varying Proportional and Integral (TVPI) PTO control, such that the WEC control remains continuous while maintaining a discrete and finite action space. The Deep Q-Network (DQN) is applied in this study as the DRL algorithm. More details of the presented DRL control can be found in [9, 10]; we will briefly introduce this control in this section.

In the DQN algorithm, the target and prediction Q-networks are formulated as:

$$\begin{aligned} \tilde{Q} &= r + \gamma \max_{a'} Q(s', a'; \theta_i^-) \\ Q &= Q(s, a; \theta_i) \end{aligned} \tag{3}$$

where s and a are the state and action at the current time step, s' and a' denote the state and action at the next time step. θ_i and θ_i^- represent the weight of the prediction and target Q network respectively. Moreover, r denotes the reward collected from the environment. In this paper, the state is the displacement and velocity of the WEC:

$$s = [z, \dot{z}] \tag{4}$$

and the reward is the mechanical power:

$$r = P_m = -F_{PTO} * \dot{z} \tag{5}$$

As we mentioned earlier, the PTO force adopts a TVPI control law:

$$F_{PTO} = -K_p(t)\dot{z} - K_i(t)z \tag{6}$$

where $K_p(t)$ and $K_i(t)$ are the time-varying control gains which are updated based on the optimal action ($a = [\delta K_i, \delta K_p]$) selected by the DRL agent at each RL decision step (t_{RL}). Finally, the neural network is trained by minimizing the error between the target and prediction network:

$$\nabla_{\theta_i} L(\theta_i) = \mathbb{E}_{s,a,r,s'} \left[\left(\tilde{Q} - Q \right) \nabla_{\theta_i} Q(s, a; \theta_i) \right] \tag{7}$$

where $L(\theta_i)$ represents the loss function.

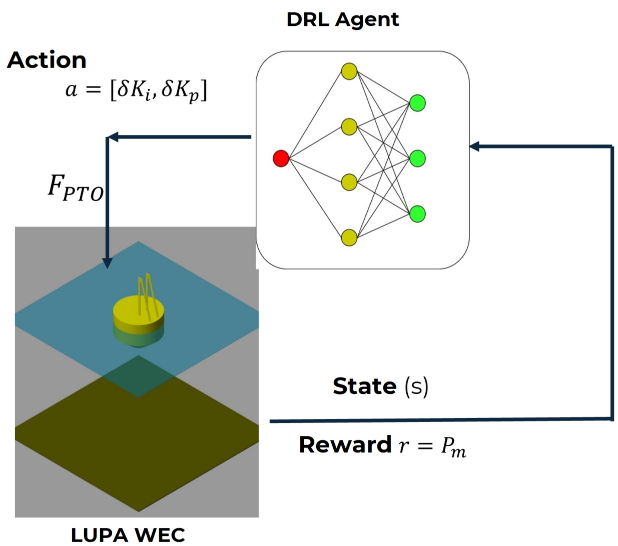

Figure 2: Framework of the DRL control.

3 Results

The testing results are presented in this section, including the initial testing and improvements, regular and irregular wave testing results, and results from testing with a reward function designed to consider PTO losses.

3.1 Initial testing results and improvement

The proposed DRL control is trained with the LUPA numerical model offline and then deployed in real time. As far as the real-time deployment of the control is concerned, the control (essentially a trained neural network) is implemented in a SpeedGOAT machine via a MATLAB/SIMULINK block called "generatePolicyBlock" (which can load the trained neural network). The control receives real-time measurements from different sensors (e.g., the motor drive, encoder, load cell, and draw wire) that are synchronized in SpeedGOAT, and calculates the desired PTO force that will be converted to a command current signal sent to the motor drive. In general, the real-time deployment process of the proposed DRL control is straightforward, and one reason for this is that the control is developed using the MATLAB Deep Learning Toolbox, which has a good interface with the real-time machine. In addition, the processing time of the control is also very small (around 0.008 ms), which is far less than the sampling rate of the sensors (1 ms). Finally, with regard to the sensor measurements, the signal obtained is relatively clean, which poses no further need for filtering.

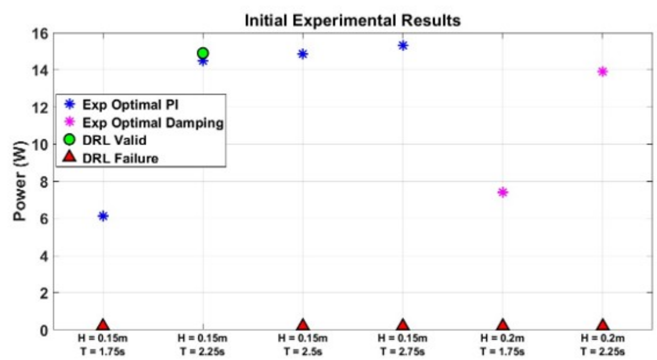

Figure 3: DRL control performance compared with OFC control when initially applied.

Despite the ease of real-time control implementation, the practical performance of the pre-trained control is very poor. Fig. 3 presents the testing results of the pre-trained DRL control. In this figure, the blue star and pink stars represent the testing results of the OFC control [12]. It is noted that the OFC control applies a PI control law when tested under less energetic wave conditions ($H = 0.15$m), while it applies only a damping control law under more energetic wave conditions ($H = 0.2$m) to protect the device from violating the stroke constraint (i.e., the device physically hitting the end stop, which may cause damage). It is clearly observed from this figure that among the controls trained for different wave conditions, only one functions properly in practice, while all the others fail. Here, the definition of control failure refers to both control gains reaching their boundary limits (a requirement in practice to prevent the control from becoming unstable). After investigating the testing results, it is found that the control failure is due to the following reasons: (1) nonlinear PTO loss, (2) random initial conditions, (3) nonlinear events, and (4) process noise.

Table 2: PTO loss value at different wave conditions

Wave Conditions	c_{pto} value (N.s/m)
H = 0.15m T = 1.75s	500
H = 0.15m T = 2.25s	700
H = 0.15m T = 2.5s	800
H = 0.15m T = 2.75s	900
H = 0.2m T = 1.75s	1200
H = 0.2m T = 2.25s	800

Nonlinear PTO loss: It is found during the test that the numerical model overpredicts the motion of LUPA under both controlled and uncontrolled conditions. Physically, this is mainly due to significant drivetrain losses, especially on the mechanical side (e.g., the belt connection with the motor). In addition, it is also found that this additional loss varies under different operating conditions. However, this nonlinear PTO loss was not properly considered in the initial training of the control. To address this issue, a linear damping-based loss term is added:

$$F_{loss} = -c_l \dot{z} \tag{8}$$

where this damping coefficient is defined as a random number selected from a certain range for different wave conditions:

$$c_l \in [c_{pto} - 300, c_{pto} + 100] \tag{9}$$

In this equation, the c_{pto} is identified for each testing wave condition by matching the system responses. The range defined in this equation is intended to cover the variability in the system responses. The c_{pto} values identified for different wave conditions are summarized in Table 2. In addition, Fig. 4 presents the numerically predicted versus experimentally measured heave displacement before (Fig. 4 (a)) and after (Fig. 4 (b)) adding this loss. It is clear from the figure that the simulated system response shows good agreement with the measured data after accounting for this PTO loss.

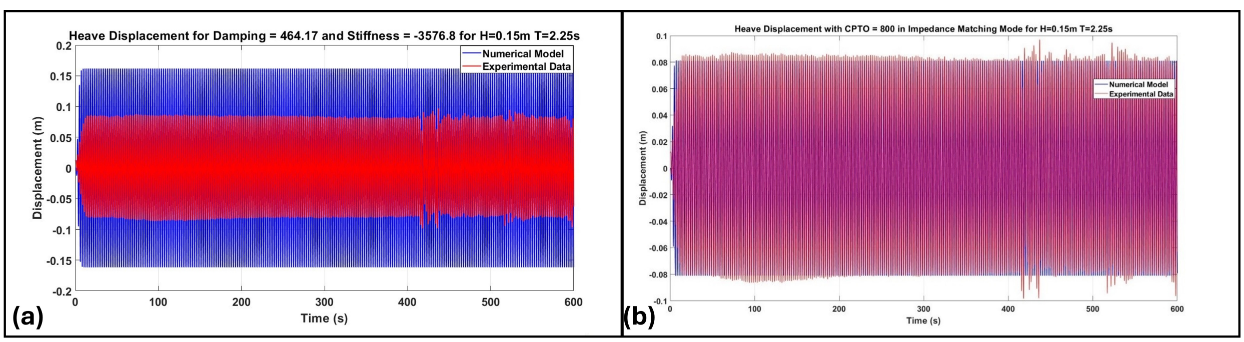

Figure 4: Comparison between the simulated and measured heave displacement (a) before fixing the PTO loss and (b) after.

Random initial conditions: Unlike numerical simulations where initial states (e.g., displacement and velocity) were always zero, real testing involved random initial conditions due to the testing setup. More specifically, during the test, the control is enabled manually by the user after a few waves have passed the WEC (to prevent the device from experiencing large transients). It was found during testing that this random initial condition has a significant impact on control performance.

Fig. 5 presents an example of control failure due to the initial conditions. In this example, the control is enabled at 9.4s, where the corresponding initial displacement and velocity are 0.021m and −0.168m, respectively. This initial condition is "unseen" by the DRL control, as the control was always trained with zero initial conditions. It is clear from the figure that the control quickly reaches the boundary due to this random initial condition and resulting transient.

Nonlinear events: It is also found that various nonlinear events have a significant impact on control performance. These so-called nonlinear events are characterized by unexpected positive or negative impulses in motion amplitude, which can easily disrupt the control. Physically, these events may be caused by nonlinear wave behavior (e.g., due to the wave paddle) or nonlinearities in the mechanical drivetrain or wave-structure interaction.

Fig. 6 shows an example illustrating the impact of nonlinear events. In this figure, a dramatic low peak occurred at around 323s in the heave displacement of the WEC, increasing by 21.38% compared to prior low peaks as shown. This event suddenly caused the integral control gain to decrease and shortly become unstable.

Process noise: Finally, the last observed potential contributor to control failure is process noise (e.g., fluctuations in the system responses). Physically, this process noise may result from various sources such as wave behavior, uncertainties, and nonlinearities in the drivetrain and wave-buoy interaction. While such noise is expected in practice, it was not considered in our model.

Improvement: To improve the performance of the control, the above-mentioned contributors to control failure are now considered during the training process. More specifically: (1) the control is numerically enabled at a random time to introduce random initial conditions; (2) the system responses simulated by the LUPA WEC-Sim model are directly modified (e.g., scaled between 70% and 130%) within a random time window; (3) normally distributed white noise is added to the system dynamics; and (4) an additional PTO loss term is included with a random damping coefficient selected from a predefined range. After introducing this significant randomness into the training of the DRL control, the control performance is significantly improved.

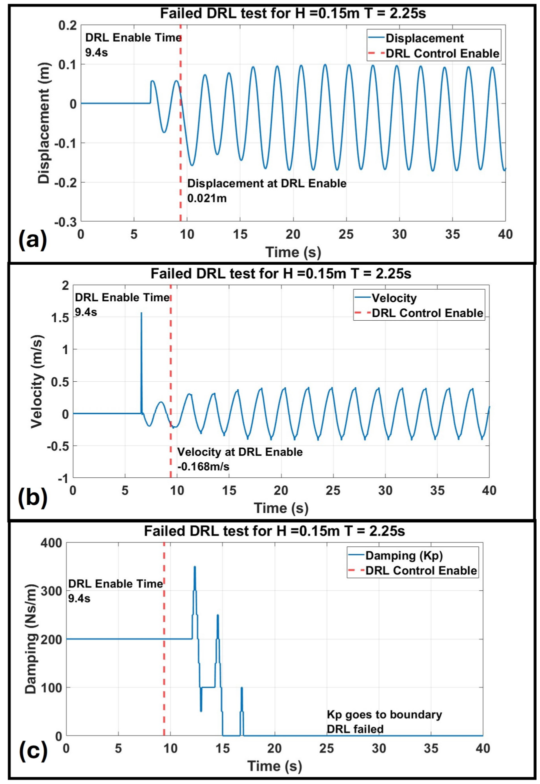

Figure 5: Details of the impact of initial conditions.

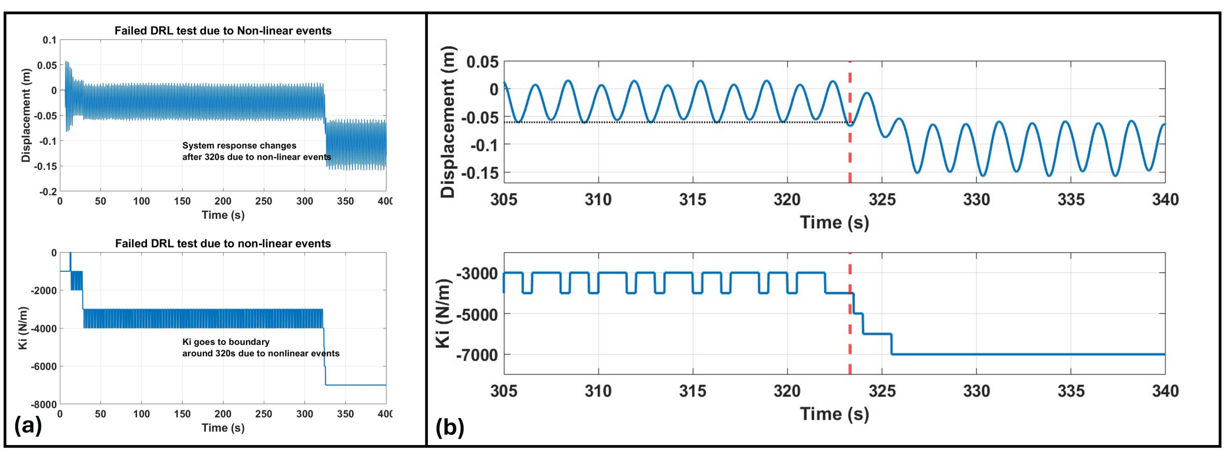

Figure 6: Details of the impact of nonlinear events.

Figure 7 illustrates the DRL control's detailed performance under a regular wave condition with a wave height of 0.15m and a period of 2.25s. In particular, Figure 7(a) shows the wave elevation recorded by a wave gauge located approximately 28.7m from the buoy, near the wavemaker. Two distinct wave events are identified in this signal, highlighted by the red and yellow boxes. The first event, occurring around 276s, resulted in a noticeable disturbance in the system responses, as shown in Figures 7(b) and (c). Despite this strong nonlinear disturbance, the control maintained robustness. As depicted in Figures 7(d) and (e), the control gains initially deviated from their optimal values but gradually returned to optimal levels by approximately 348s, after the system responses return to a normal pattern. A second nonlinear event is observed at around 506s, which also caused a clear change in the system behavior, highlighted by the green boxes in Figures 7(b) and (c). Interestingly, no corresponding irregularity in the wave input was detected at that time. The control exhibited a similar response to this disturbance as in the first case. Lastly, the power output under this wave condition is shown in Figure 7(f), with the average wave power produced measured at approximately 11.47W.

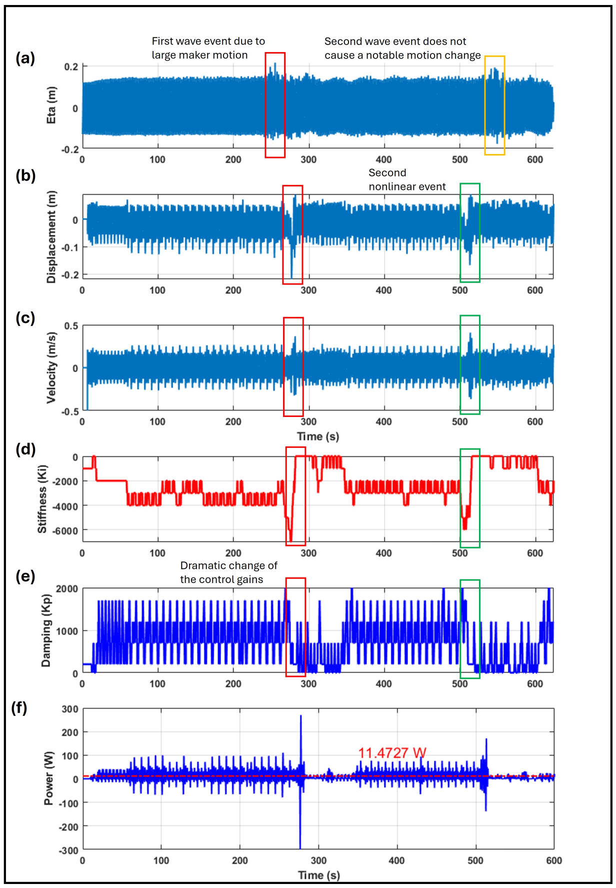

Figure 7: Detailed performance of DRL control under a regular wave with height of 0.15m and period of 2.25s under S1 observation states

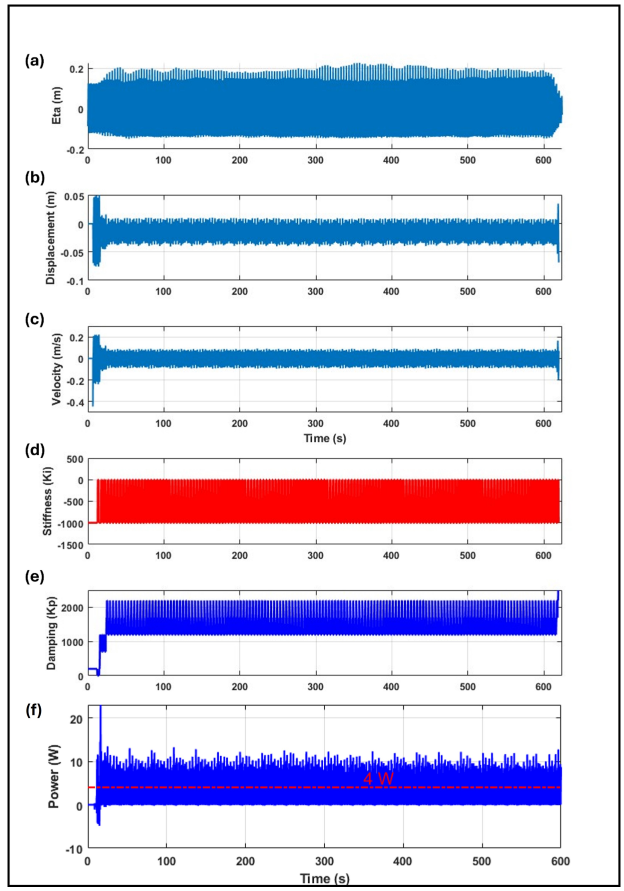

Figure 8: Detailed performance of DRL control under a regular wave with height of 0.15m and period of 1.75s under S6 observation states

In addition to incorporating significant randomness during training to enhance control performance, the impact of different state selections is also investigated. Representative state configurations, drawn from relevant literature [8, 13, 14], are summarized in Table 3. These state sets were evaluated under the same regular wave condition ($H = 0.15$m and $T = 1.75$s). Based on the results in Table 3, it is evident that states $S1$ and $S6$ yield the best performance in terms of wave power production. The performance of the DRL control using $S1$ has already been discussed (see Fig. 7); here, we further examine

Table 3: Different observation states and their performance

Number	Expression	Power
S1:	$s = (z, \dot{z})$	4W
S2:	$s = (z, \dot{z}, H_s, T_p)$	Unstable, No Power
S3:	$s = (z, \dot{z}, F_{PTO})$	Unstable, No Power
S4:	$s = (z, \dot{z}, P_{PTO})$	Unstable, No Power
S5:	$s = (z, v, P_{PTO}, H_s, T_p, a_{t-1})$	2.5W
S6:	$s = (z, \dot{z}, H_s, T_p, a_{t-1})$	4W
S7:	$s = (z, \dot{z}, \frac{\sum P_{PTO}}{N})$	2.7W
S8:	$s = (z, \dot{z}, H_s, T_p, \frac{\sum P_{PTO}}{N})$	Unstable, No Power

the control performance using $S6$.

Figure 8 presents the detailed control performance when $S6$ is used as the state under the same regular wave condition (H = 0.15m, T = 1.75s). As shown in Figures 8(a), (b), and (c), no distinct nonlinear wave events—such as large impulses—are observed during the test period. However, the wave elevation exhibits noticeable fluctuations, which influence the system responses. Despite this, the control demonstrates strong robustness throughout the testing period and maintains a steady-state, time-varying behavior, as seen in Figures 8(d) and (e). Specifically, the integral gain oscillates around −625N/m, while the proportional gain oscillates around 1450N·s/m. Regarding power production, the average output is approximately 4W, as shown in Figure 8(f).

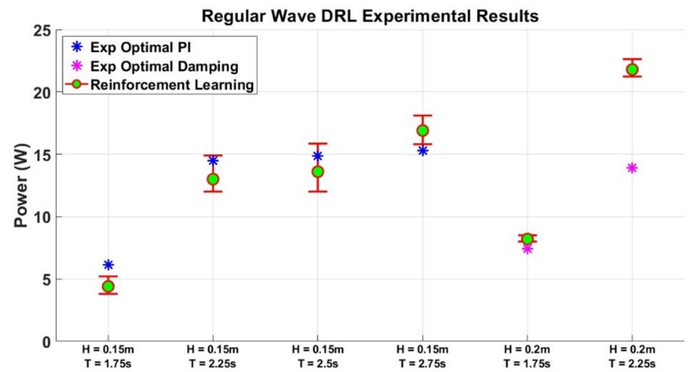

Figure 9: DRL control performance compared with OFC control after improvement.

When comparing the DRL control performance using $S1$ and $S6$, we observe notable differences. In the case of $S1$, where the state comprises only displacement and velocity, the controller demonstrates high adaptivity to nonlinear disturbances. However, this adaptivity comes at the cost of noticeable deviations from optimal performance during such events, along with considerable recovery times. On the other hand, $S6$ incorporates additional inputs such as the sea state and previous control action, resulting in a more robust response. This setup shows minimal deviation during disturbances but reduces the controller's adaptivity. While this robustness enhances performance in regular wave scenarios, it may lead to too rigid behavior under irregular wave conditions, which will be further discussed later.

3.2 Testing results under regular waves

Given that S6 provides better performance in terms of energy harvesting, this observation state is selected for all regular wave conditions. The performance of the DRL control across various regular wave conditions is presented in Fig. 9. In this figure, again, the blue and pink stars represent the tank testing results of the OFC control. The green dots represent the performance of the DRL control. It is noted that for each wave condition, the DRL control was executed three times, the red bars indicate the maximum and minimum power production from the three tests. As illustrated in the figure, the proposed control achieves power output close to the OFC control across most wave conditions. A significant difference can be noticed at a wave height of 0.2 m and period of 2.25 s, where the DRL control produces approximately 56.8% more power. This is due to the DRL control incorporating a spring component, whereas the OFC control under this condition only applies damping.

3.3 Testing results under irregular waves

The control performance is further validated under irregular wave conditions. The applied wave profile has a significant wave height of 0.21m and a peak period of 3.09s. During testing with regular waves, $S6$ was selected as the observation state due to its demonstrated robustness. This same state was initially used for the irregular wave test. However, it was observed that under irregular conditions, the PI gains of the DRL control were consistently adjusted in response to peak system responses, rather than adapting to the wave profile. As a result, the gains frequently converged to their boundary values, resulting in suboptimal energy production. Given that $S1$ (comprising displacement and velocity) exhibited better adaptivity in earlier tests, it was reselected for use under irregular wave conditions to improve the control performance.

Fig. 10 presents the detailed performance of the DRL control under irregular wave condition. The figure indicates that the integral gain in the control strategy coverges at its lower bound of −6000N/m, which is due to the relatively long wave peak period of 3.09s. Under such conditions, a more negative stiffness is needed to bring the system's natural frequency closer to the wave frequency. Meanwhile, the proportional gain in the PI controller shows strong adaptivity to the system's behavior. Notably, the variation of the proportional gain in the DRL control follows the fluctuations in system responses. As illustrated in Fig. 10 (b) and (c), the proportional gain decreases when the WEC exhibits low motion (indicated by the red boxes), and increases as the motion amplitude increases (indicated by the green boxes). Finally, the power harvested by the DRL control is approximately 11.6W, which is significantly higher than the energy captured by an optimal damping control (5.4W). The irregular wave test was repeated three times under the same wave condition but with different random seeds. The power output remained consistent across trials, generating 14W and 15.5W power in the other two tests, demonstrating the control's reliability and effectiveness under irregular wave conditions.

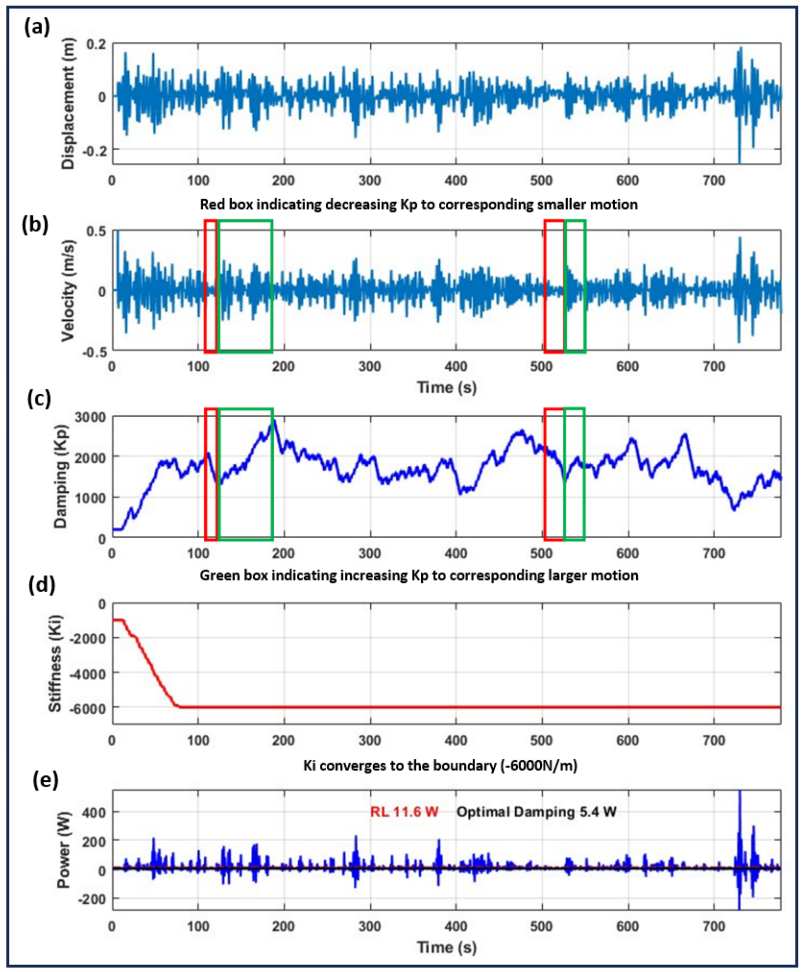

Figure 10: Detailed performance of the DRL control under an irregular wave.

3.4 Testing results with new reward design

One of the significant benefits of the proposed DRL control is its ability to optimize WEC performance from wave to wire, due to its model-free nature. However, in this study, the reward collected from the environment is mechanical power rather than electrical power, due to two main reasons: (1) the current constraints of the experimental setup, specifically the absence of a dedicated power analyzer to measure both AC and DC power output; and (2) the lack of a dedicated PTO model that can calculate LUPA's electrical power output for use in DRL training.

As mentioned previously, developing the control strategy only based on mechanical power does not necessarily lead to optimal electrical power output, especially when factoring in losses within the PTO drivetrain. To better evaluate electrical power output under the limitations of the current test setup, a reward function that considers these losses is proposed:

$$r = -\frac{\sum(F_{PTO}\dot{z})}{N} - \alpha F_{PTO}^2 \tag{10}$$

where the first term represents the moving average of harvested mechanical power, and the second term applies a penalty on the PTO force to account for drivetrain losses. The parameter α serves as a tunable weight that adjusts the impact of this penalty. This formulation allows the DRL control to balance energy production with drivetrain efficiency, providing a more realistic approximation to optimize electrical power output under the given testing constraints.

Figure 11 presents the performance of the DRL control using the penalized reward function, where the penalty weight is set to 0.0002. The test is conducted under a regular wave condition with a height of 0.2m and a period of 1.75s. Compared to the unpenalized case, the system responses clearly show that both motor torque and current are significantly reduced—by approximately 35% (as shown in Figures 11(a) and (b)), due to the constraint imposed on the PTO force. As a result, the heave displacement of the WEC is also reduced, as shown in Figure 11(c). Overall, the mechanical power output is reduced by 35.8% (Figure 11(d)) relative to the unpenalized case.

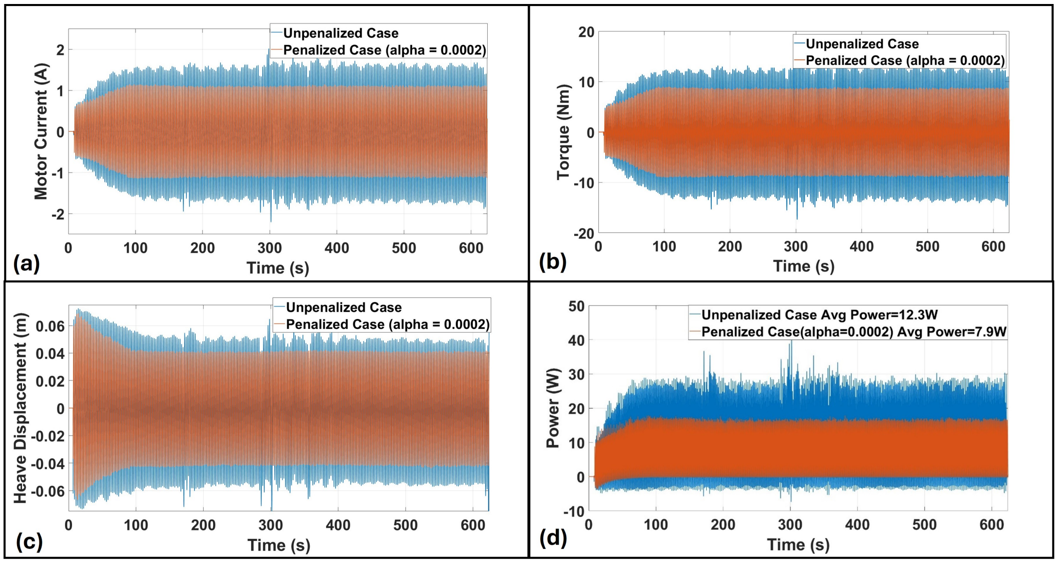

Figure 11: Detailed performance of the DRL control with penalty (α = 0.0002) in the reward function compared with unpenalized case.

Regarding losses, it is important to note that the motor copper loss (one of the significant components of the overall PTO loss) is proportional to the square of the current. Therefore, a 35% reduction in current implies a potential reduction in copper loss of approximately 58.1%, which exceeds the reduction in mechanical power. This indicates a potential improvement in electrical power output.

4 Conclusion

This study presents the practical implementation and validation of a DRL-based control for WECs. The control is tested under varied wave conditions and generates a meaningful dataset, which is now publicly available in [15]. The key findings of this project are: (1) The DRL control is trained by using MATLAB Deep Learning Toolbox, which is found to be very easy to integrate with the SpeedGOAT real-time machine, and the computational cost is also very small. (2) The robustness of the control is a critical challenge in practical application. Introducing significant randomness in the training of the DRL control can significantly improve control robustness. (3) The observation state of the control also has a significant impact on control performance. It is found that the control is adaptive and optimal but slightly less robust when the displacement and velocity are applied as the state. In contrast, further adding wave states and prior action into the state can improve robustness but reduce adaptivity. (4) Including a penalty term for the PTO control force in the reward function is an effective way to minimize the drivetrain losses. In the future, the team plans to develop a motor model for LUPA and further train and test the control directly with electrical power output.

Acknowledgments

The authors gratefully acknowledge the US Department of Energy, the Pacific Ocean Energy Trust (POET), and the TEAMER program for supporting and funding this work. In addition, the authors would like to thank Dr. Gordon Parker for his valuable discussions and for providing access to high-performance computing resources used for DRL training. Appreciation is also extended to Dr. Tim Maddux for his insightful discussions and on-site support with processing the wave gauge data.

References

[1] Benjamin Drew, Andrew R Plummer, and M Necip Sahinkaya. *A review of wave energy converter technology*. 2009.

[2] Tunde Aderinto and Hua Li. "Ocean wave energy converters: Status and challenges". In: *Energies* 11.5 (2018), p. 1250.

[3] Bo Yang et al. "A critical survey of power take-off systems based wave energy converters: Summaries, advances, and perspectives". In: *Ocean Engineering* 298 (2024), p. 117149.

[4] Johannes Falnes and Adi Kurniawan. *Ocean waves and oscillating systems: linear interactions including wave-energy extraction*. Vol. 8. Cambridge university press, 2020.

[5] Shangyan Zou et al. "Optimal control of wave energy converters". In: *Renewable energy* 103 (2017), pp. 217–225.

[6] Shangyan Zou and Ossama Abdelkhalik. "Control of wave energy converters with discrete displacement hydraulic power take-off units". In: *Journal of Marine Science and Engineering* 6.2 (2018), p. 31.

[7] Xiang Zhou et al. "Assessment of electrical power generation of wave energy converters with wave-to-wire modeling". In: *IEEE Transactions on Sustainable Energy* 13.3 (2022), pp. 1654–1665.

[8] Enrico Anderlini et al. "Towards real-time reinforcement learning control of a wave energy converter". In: *Journal of Marine Science and Engineering* 8.11 (2020), p. 845.

[9] Shangyan Zou et al. "Deep reinforcement learning control of wave energy converters". In: *IFAC-PapersOnLine* 55.27 (2022), pp. 305–310.

[10] Shangyan Zou et al. "Optimization of the electricity generation of a wave energy converter using deep reinforcement learning". In: *Ocean Engineering* 244 (2022), p. 110363.

[11] B Bosma et al. "Design and modeling of a laboratory scale WEC point absorber". In: *14th European Wave and Tidal Energy Conference (EWTEC 2021)*. 2021.

[12] Bret Bosma, Courtney Beringer, and Bryson Robertson. "Experimental passive and reactive control of a Laboratory Scale WEC Point Absorber". In: *Proceedings of the European Wave and Tidal Energy Conference*. Vol. 15. 2023.

[13] Enrico Anderlini, Gordon G Parker, and Giles Thomas. "Docking control of an autonomous underwater vehicle using reinforcement learning". In: *Applied Sciences* 9.17 (2019), p. 3456.

[14] Kemeng Chen et al. "Design and Tank Testing of Reinforcement Learning Control for Wave Energy Converters". In: *IEEE Transactions on Sustainable Energy* (2024).

[15] Shangyan Zou et al. *TEAMER: Experimental Validation and Analysis of Deep Reinforcement Learning Control for Wave Energy Converters*. Marine and Hydrokinetic Data Repository, Michigan Technological University, https://mhkdr.openei.org/submissions/628. Accessed: 2025-07-13. 2025. URL: `https://mhkdr.openei.org/submissions/628`.