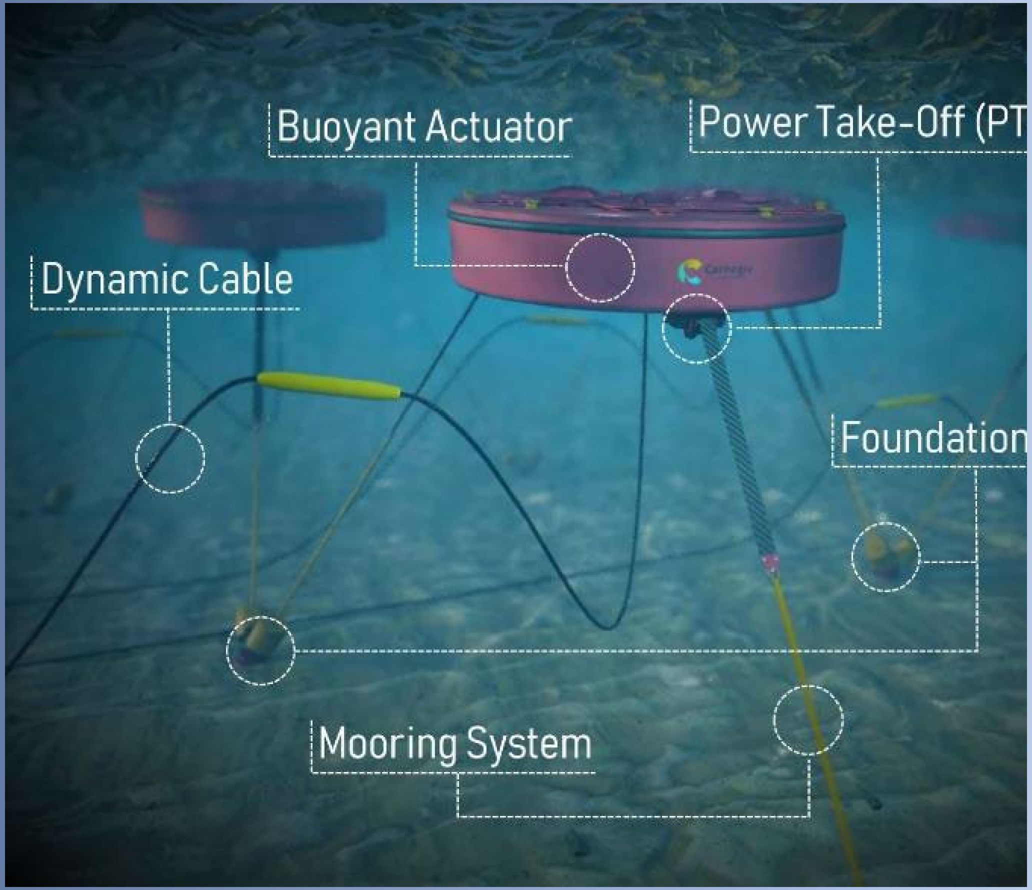
Ocean Wave Energy: Optimizing Reinforcement Learning Agents for Effective Deployment

Vineet Gundecha[1], Sahand Ghorbanpour[1], Ashwin Ramesh Babu[1], Avisek Naug[1], Alexandre Pichard[2], Mathieu Cocho[2], Soumyendu Sarkar*[1]

[1]Hewlett Packard Enterprise
[2]Carnegie Clean Energy

Climate Change AI

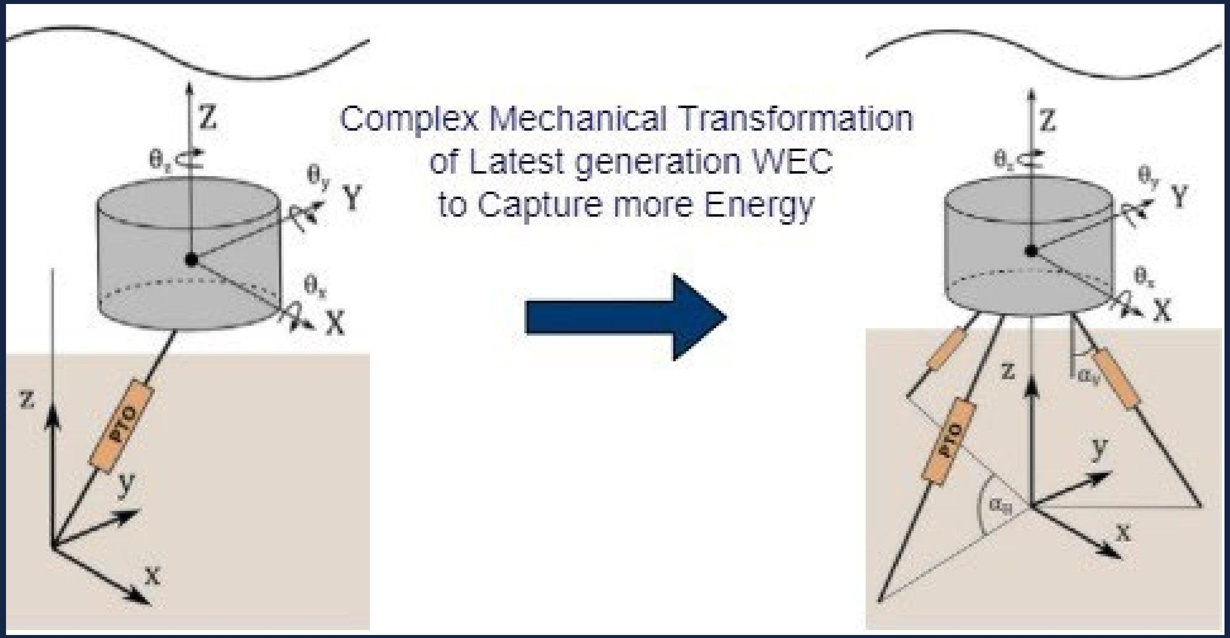
NEURAL INFORMATION PROCESSING SYSTEMS



- Increase in energy efficiency ↑Revenue
- Reduce structural stress ↓ Maintenance
- Protect from acute weather events

Carnegie Clean Energy's CETO 6 Wave Energy Platform won the EuropeWave Phase 3 contract for deployment in BiMEP in Spain after evaluation for cost effective design validation in Phase 2.

Complex Multi-generator structure ► Complex Optimal Control



Simple Earlier Generation Wave Energy Converter

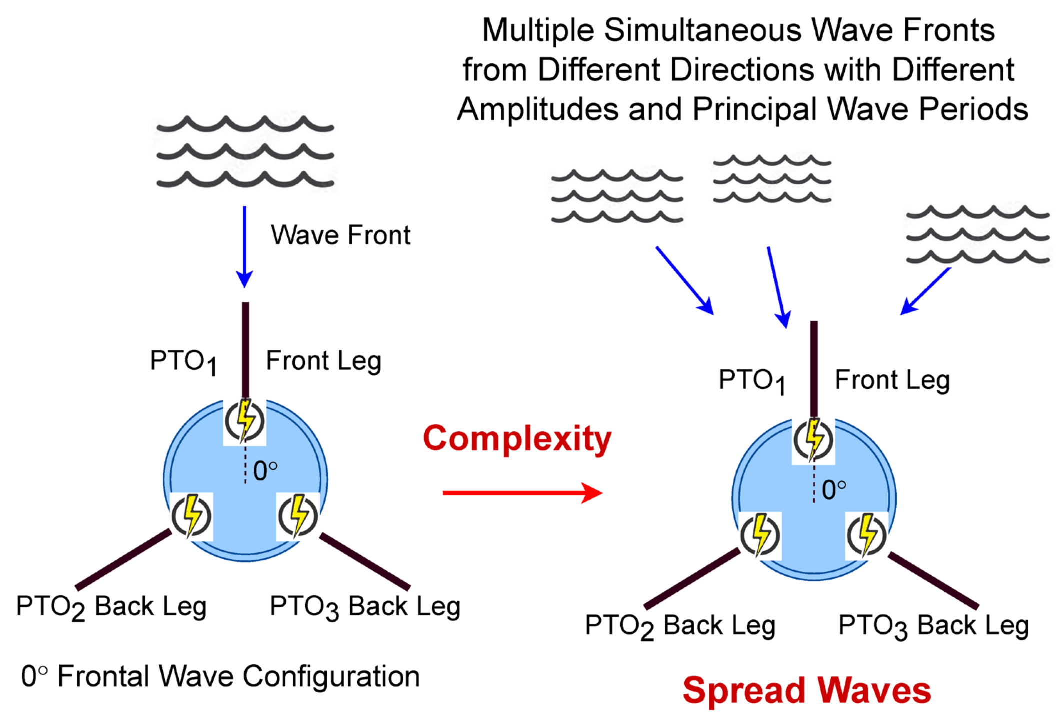
Complex CETO 6 Multi-Generator Wave Energy Converter

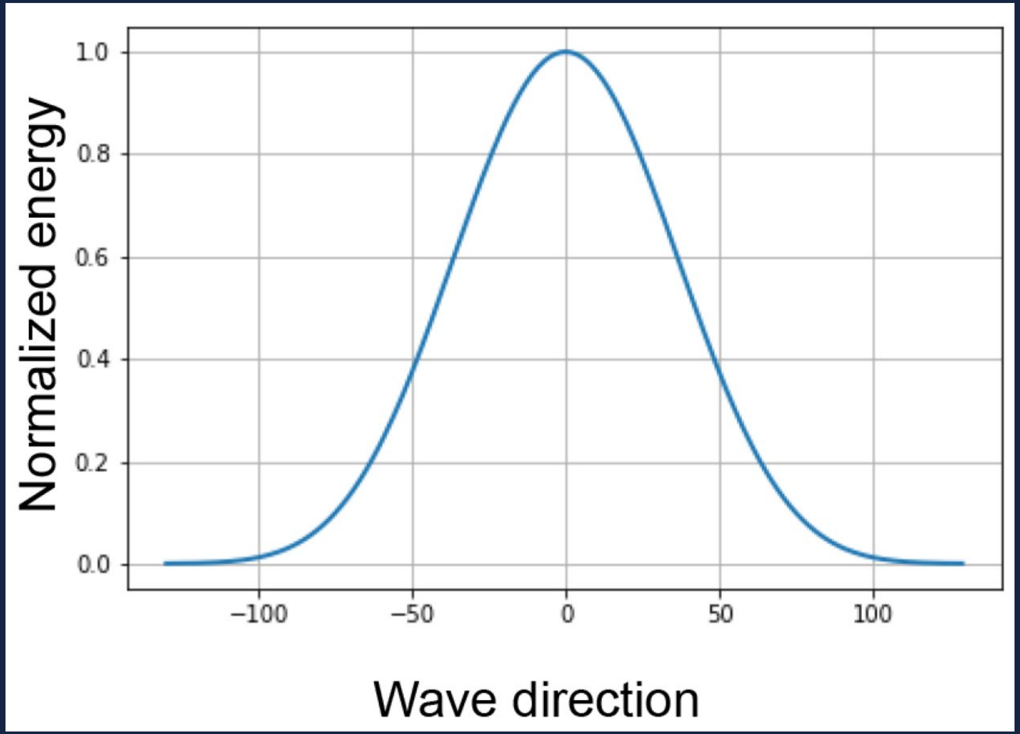
Multi-Agent

Reinforcement Learning

Controller

Spread waves ▶ Complexity





➡ Spread waves significantly increase complexity for WEC with simultaneous waves from different angles based on atmospheric events at different parts of the ocean.

Variations of waves differ with locations ▶ Complexity

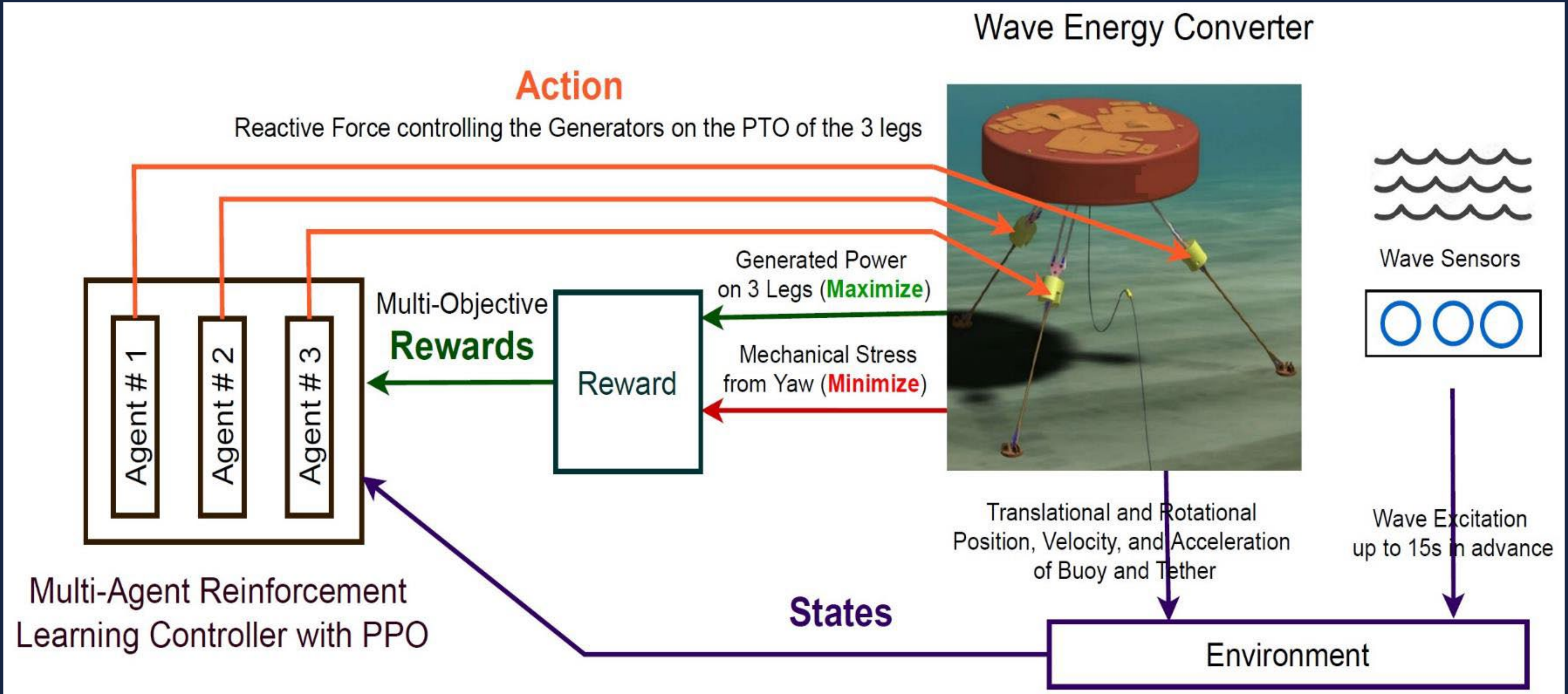
Occurance (%) of Wave Types at BiMEP deployment site

Height (m) \ TPeriod(s)	6	8	10	12	14	16
1	13.89	16.88	13.99	6.23	1.57	0.22
2	0.72	4.74	7.42	9.14	5.73	1.71
3	0.00	0.42	1.88	3.01	3.37	2.37
4		0.01	0.55	1.61	2.08	2.46

Normalized Power for Different Wave Types

Height (m) \ TPeriod(s)	6	8	10	12	14	16
1	1.16	1.21	1.01	0.79	0.61	0.48
2	3.64	3.68	3.27	2.68	2.15	1.70
3	6.50	6.55	5.86	4.98	4.12	3.37
4	9.47	9.42	8.44	7.37	6.28	5.26

Reinforcement Learning System Architecture



Climate Change AI NEURAL INFORMATION PROCESSING SYSTEMS

State Design

position	position of the buoy with velocity and acceleration for the translational and rotational motion
yaw	rotational yaw motion to monitor stress
tether	extension and velocity of tether
wave	wave elevation and rate of change for present and 10s ahead in time from sensors

$$s = [\ e\ \dot{e}\ \ddot{e}\ g\ \dot{g}\ z\ \dot{z}\ ]^T$$

where,

e - the buoy position,

g - the tether extension

z - wave excitation

All RL agents share the continuous observation space of position and wave.

Action Space

The continuous action space for the individual RL agent is defined by the reactive force $f_{gen(i)}$ for the controlled generator, where "*i*" represents the index for the agent.

Reward

$$Reward_i = \alpha.(P_{own(i)} + \eta_i.P_{others})$$

where,

P - generated power defined by, $-f_{gen} * \dot{e}$.

η - the hyperparameter for the team coefficient,

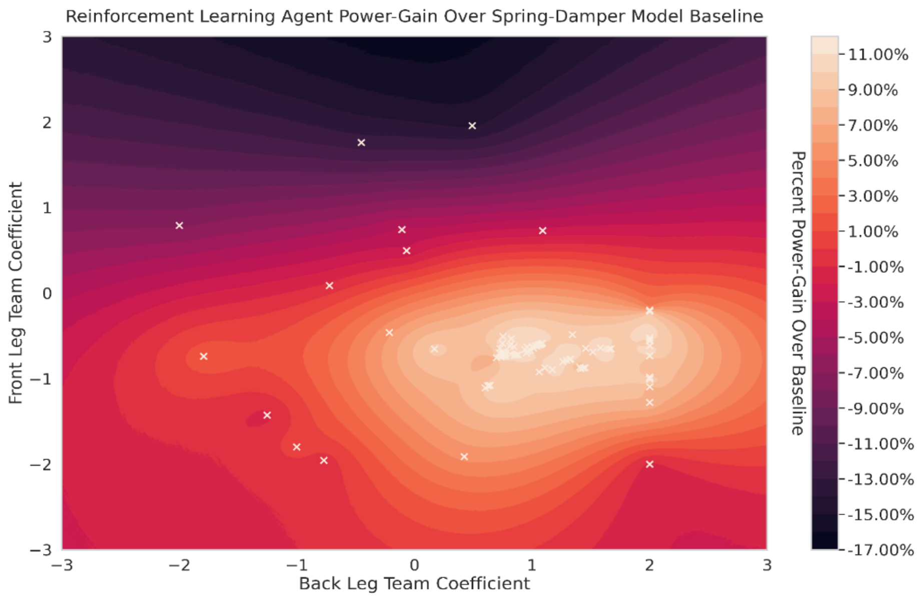
Co-opetition between multiple RL agents and Hyper-parameter optimizations

- Multi-agent Co-opetition: Disparity in power and optimum trade-off by individual legs, to get most combined power in all legs, make the best solution a combination of co-operation and competition for different agents.
- Multi-dimensional hyper-parameter optimization was key for RL agent performance gains

Reward Shaping

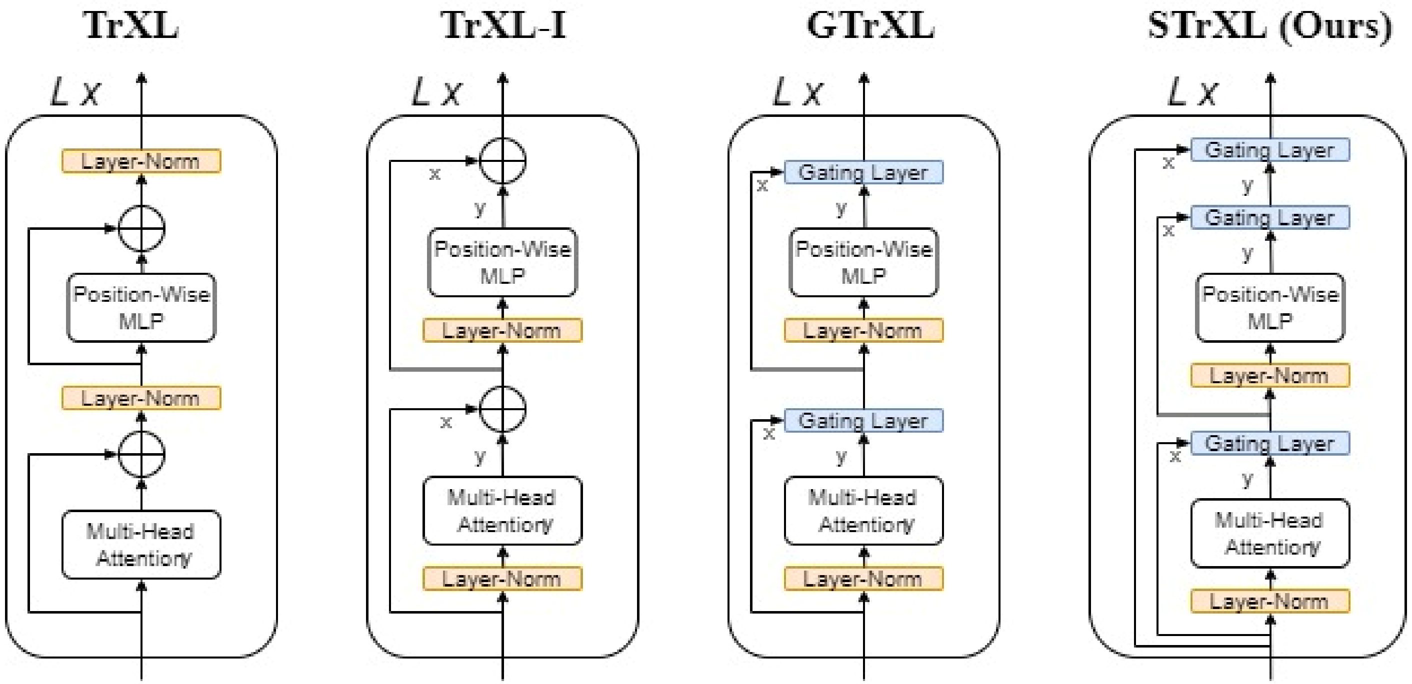
$$Reward = P_{own} + \eta \, . P_{others}$$

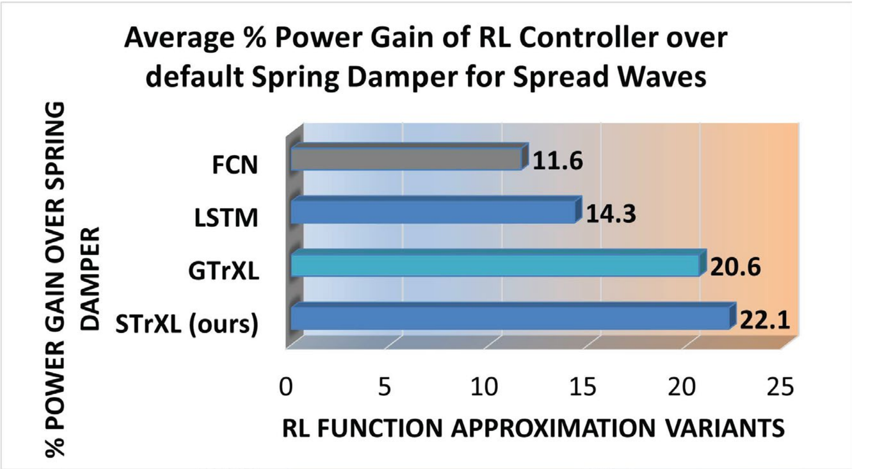
where, η = team coefficient, P_{own} is the power of the generator being controlled and P_{other} is power from other generators

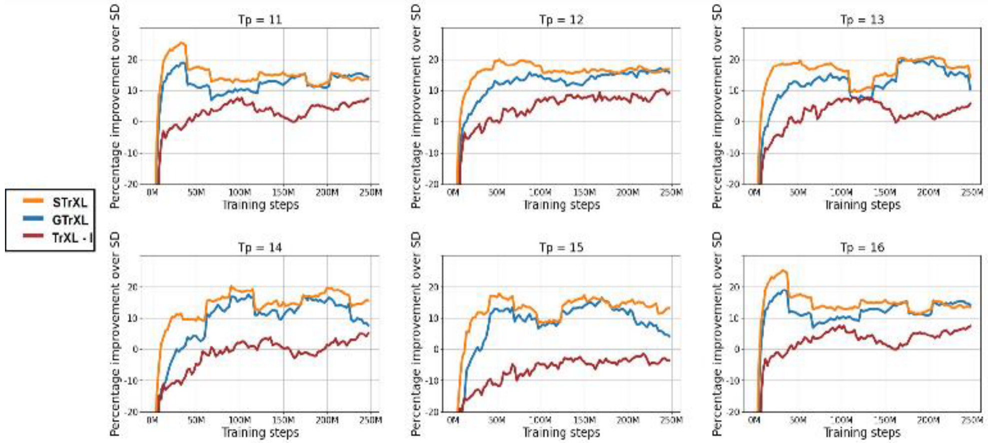


Novel STrXL Transformer architecture for RL beats SOTA

Climate Change AI
NEURAL INFORMATION PROCESSING SYSTEMS







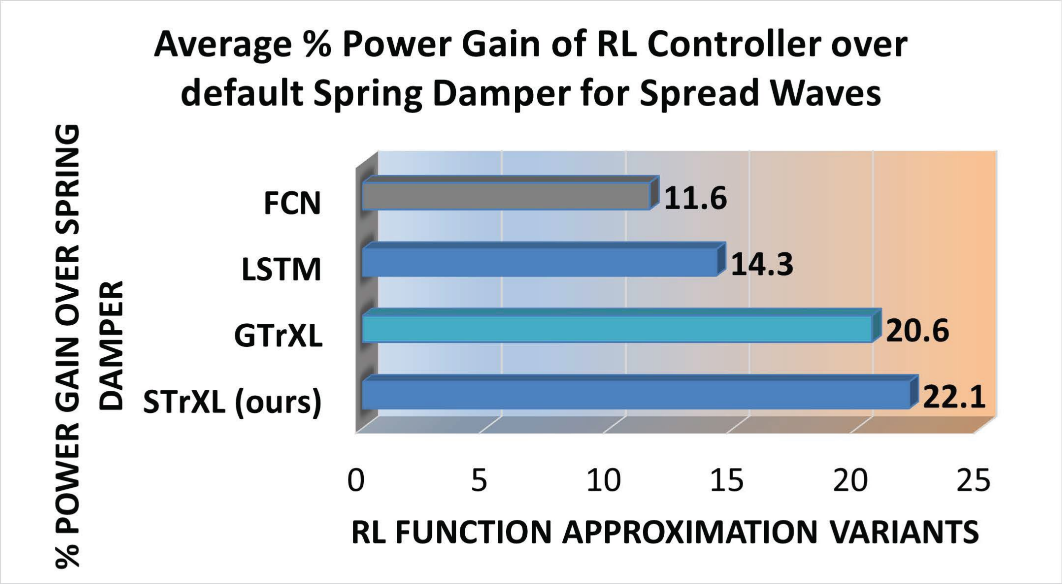
- Our Novel STrXL **beats State-of-the-art** GTrXL transformer on training speed and performance for multi-agent RL for CETO 6 WEC

- Transformers are hard to train for multi-agent RL

Our Novel STrXL beats SOTA GTrXL on Performance

RL Controller Power gain (%) over default spring damper (SD) controller for different Function Approximations

Spread Waves: RL % Gain of Energy Capture over default (SD controller) % Gain for Wave Height = 2m												
Wave Time Period(s)	6	7	8	9	10	11	12	13	14	15	16	Avg
FCN	15.2	15.4	12.0	11.7	12.2	10.2	13.5	8.4	9.2	10.1	9.4	11.6
LSTM	18.2	19.2	15.2	14.2	15.2	13.2	11	11	12.5	15.1	12.1	14.3
GTrXL	22.2	24.1	25.4	23.9	19.3	14.9	23.2	15.1	17.4	19.9	21	20.6
STrXL (ours)	23.1	25.2	24.2	25.2	21.4	22.3	25.4	17.2	20.2	20.5	18.2	**22.1**



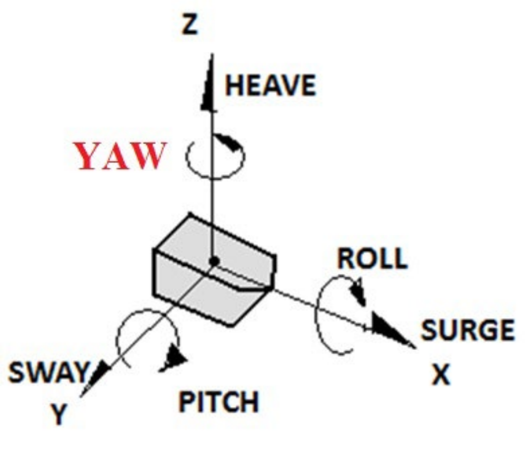
Performance : our Novel STrXL beats State-of-the-art GTrXL transformer for multi-agent RL for CETO 6 WEC for **Spread Waves**

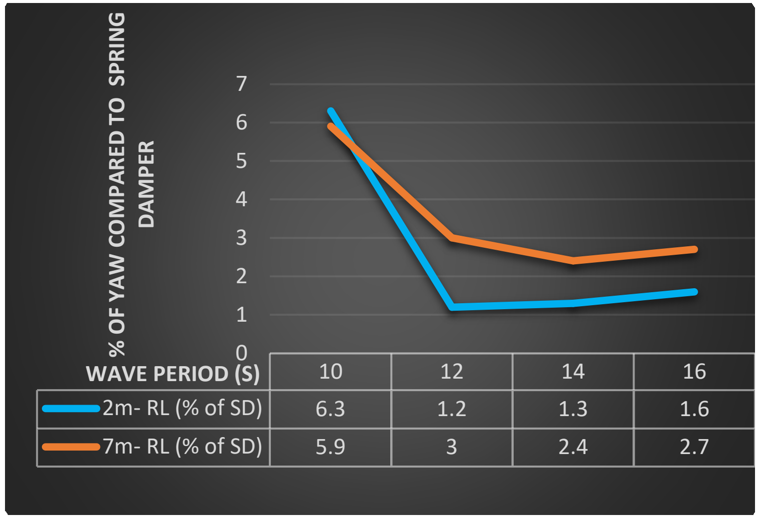
Mechanical stress and Maintenance mitigation with ML Trust

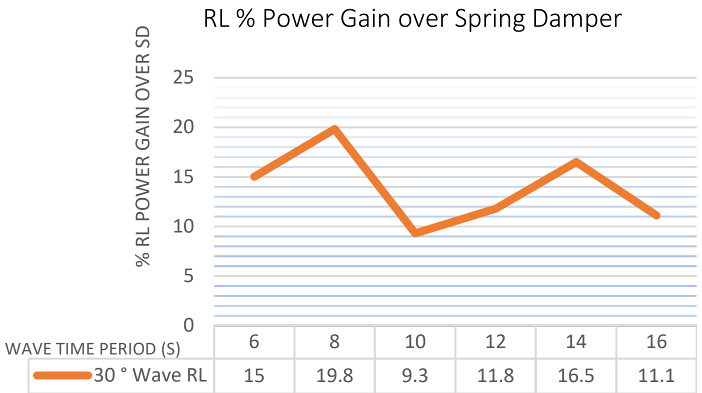
Climate Change AI

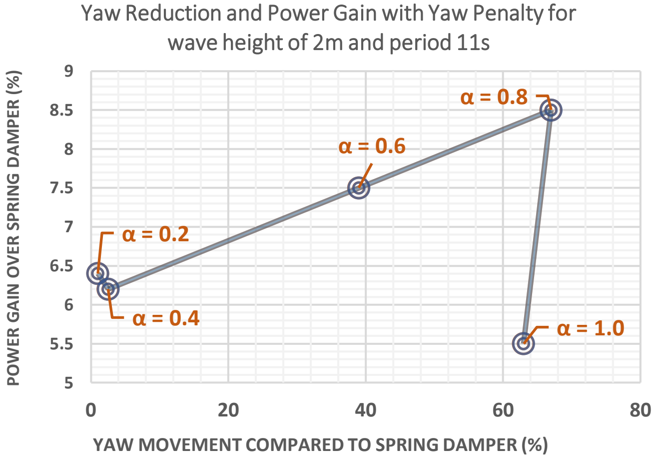
NEURAL INFORMATION PROCESSING SYSTEMS

- Yaw rotational motion of the voluminous buoy causes damaging mechanical stress and is higher for angled waves.
- Penalty for yaw: $\boldsymbol{Reward = (\alpha)\ power + (1-\alpha)\ yaw}$,
  where α is tunable, lesser α => stronger penalty.
- Result: Yaw is almost eliminated resulting in less maintenance compared to spring damper (SD).
- Reward shaping with yaw minimization improved power generation, as more power directed to the generator
- Combined reward maximizes energy capture maximization and minimizes stress.









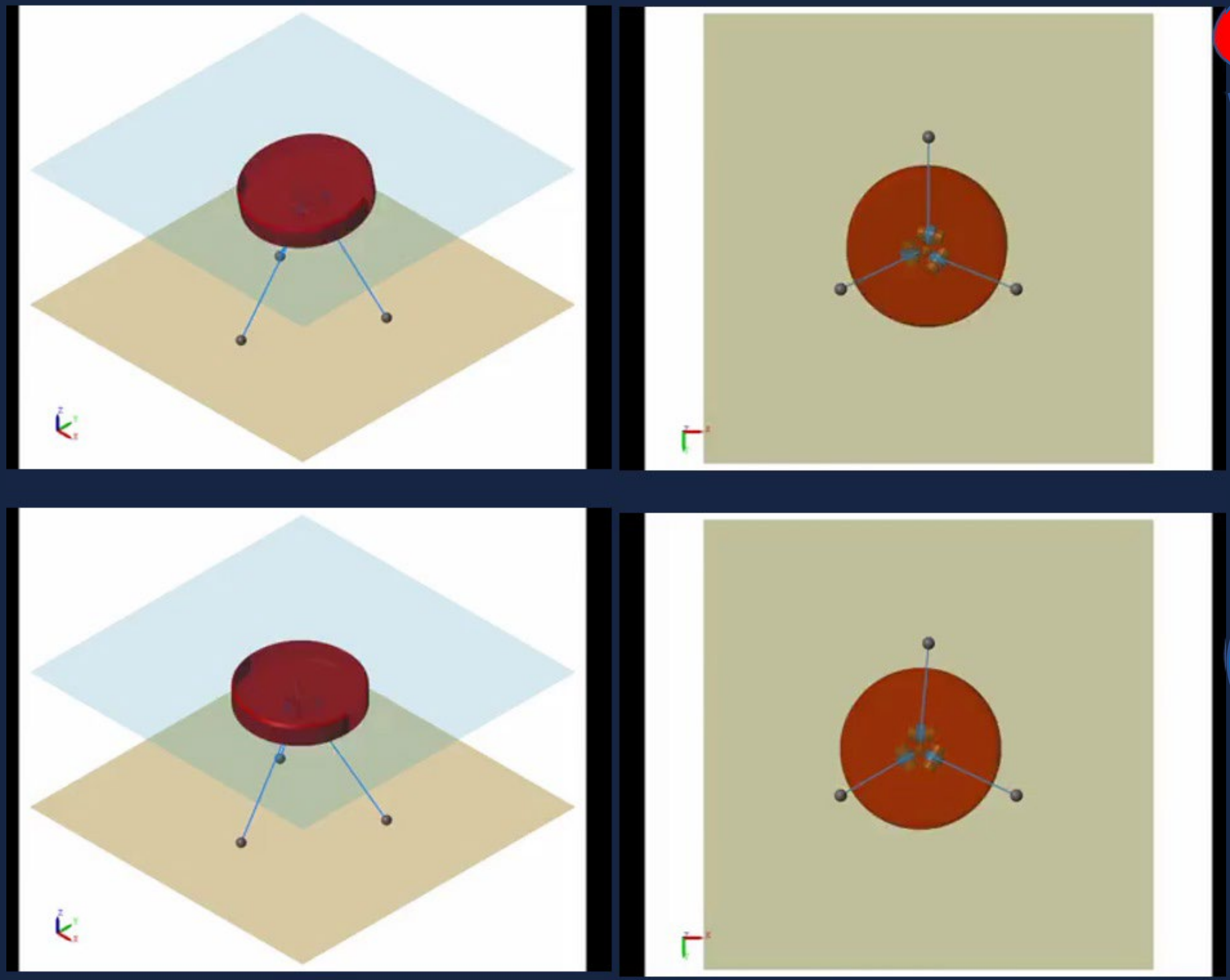
RL controller reduces stress : Yaw minimization

Isometric view

Bottom view

Spring-damper controller

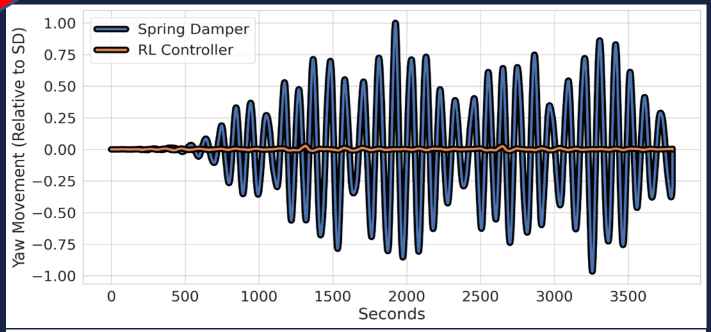
RL controller



8x Speed

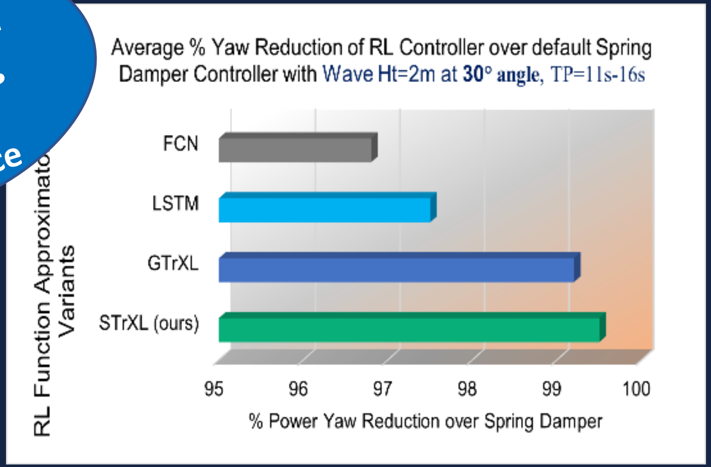
Twisting Yaw higher for Spring Damper

Yaw v/s time

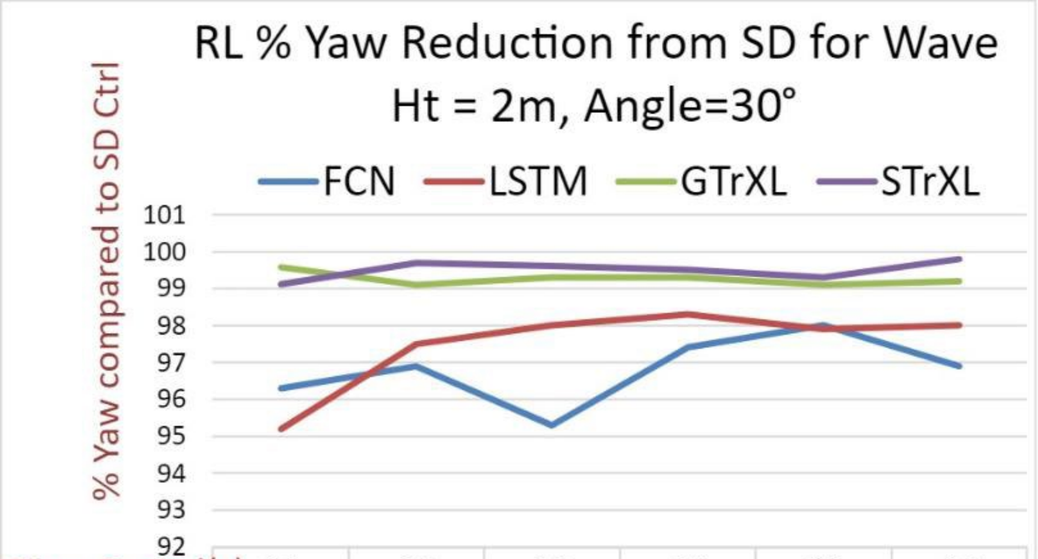


Comparison of Yaw movement between RL and the spring damper (SD) controllers for an episode with wave height of 2m and principal wave period of 12s. Values are relative to maximum SD yaw.

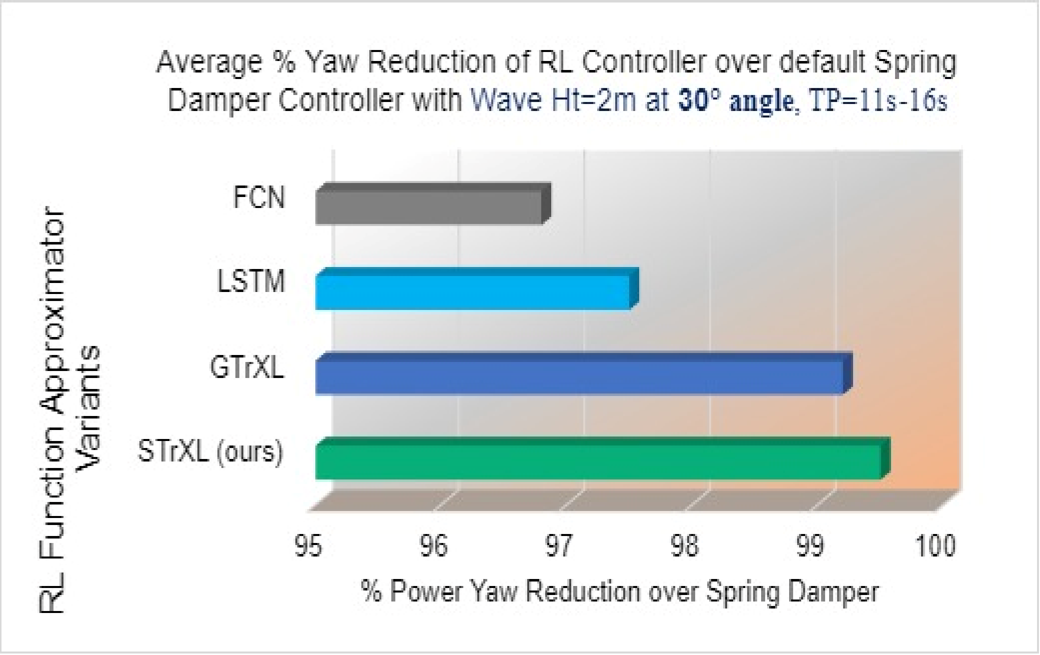
Twisting Yaw minimized with RL minimizing stress and maintenance



RL Controller % Yaw Reduction over default controller



Time Period(s)	11	12	13	14	15	16
FCN	96.3	96.9	95.3	97.4	98	96.9
LSTM	95.2	97.5	98	98.3	97.9	98
GTrXL	99.58	99.1	99.3	99.3	99.1	99.2
STrXL	99.12	99.7	99.6	99.5	99.3	99.8



Deployment specific RL controller tuning with Spread Waves

- Some sea states carry higher energy than others
- Some sea states have higher occurrence than others

we normalize the generated power by the optimal power for every sea state

$$r(t) = \frac{power(t)}{power\_opt(tp, ht)} - \alpha * f(yaw(t)) * yaw(t)$$

- power generated from different wave types is normalized by the max theoretical power for the wave types, and wave types are sampled based on occurrence
- ά is the weighting factor for the yaw penalty, $yaw(t)$ is the yaw at time t, and $f\ (yaw(t))$ is a non-linear yaw factor that penalizes higher yaw more than lower yaw

% Power Gain of RL controller over default Spring Damper for CETO 6 WEC

Height (m) \ TPeriod(s)	6	8	10	12	14	16
1	9.60	22.80	36.30	45.50	50.50	54.70
2	3.80	10.60	15.20	24.60	23.90	31.60
3	0.80	4.20	8.20	13.60	15.20	18.40
4	-0.70	1.90	3.50	6.94	11.20	15.10

% Power Distribution across different wave types (weighted by occurance)

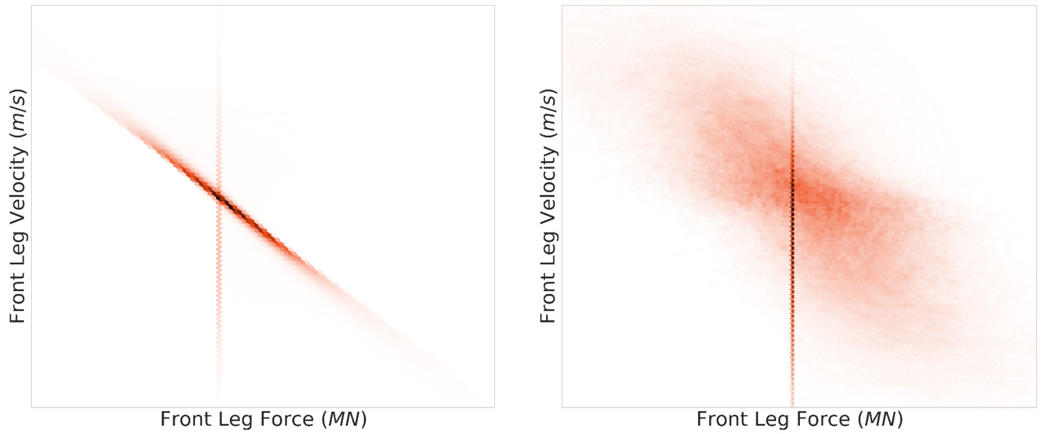
Height (m) \ TPeriod(s)	6	8	10	12	14	16
1	6.87	8.75	6.04	2.11	0.41	0.04
2	1.12	7.46	10.36	10.49	5.26	1.24
3	0.00	1.18	4.71	6.40	5.93	3.42
4		0.04	1.98	5.07	5.58	5.53

Intuition behind Reinforcement Learning controller performance

Climate Change AI

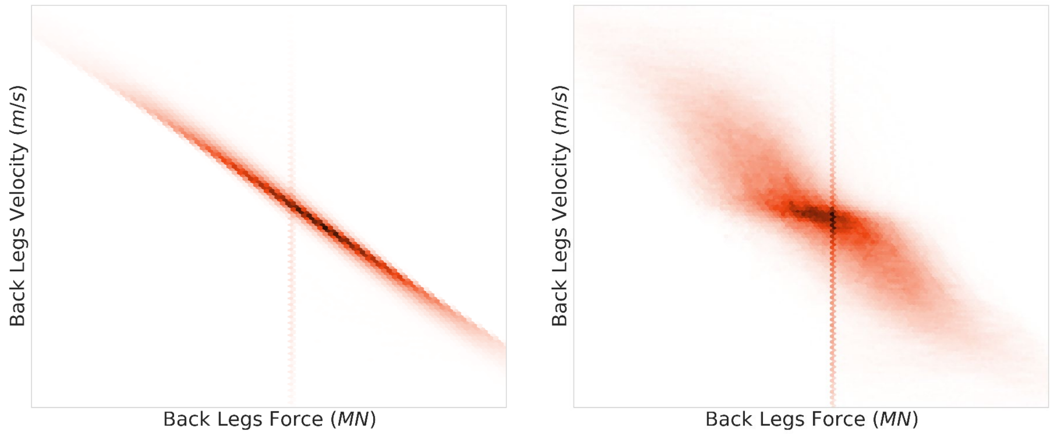
NEURAL INFORMATION PROCESSING SYSTEMS

- Spring Damper is more greedy and reactive forces for the generators on the legs are almost proportional to the instantaneous velocity of the tether as energy is captured working against this motion.
- RL controller is fuzzy about the proportionality of reactive force and tether (leg) velocity, as it compromises short-term objectives for greater gains on energy capture at the more opportune segments of the wave cycles with discounted returns.



Spring Damper
Front Leg

Reinforcement Learning
Front Leg



Spring Damper
Back Leg

Reinforcement Learning
Back Leg

- Better co-ordination between the multiple generators and legs with varying waves and 6 degrees of motion which the existing state of the art controllers fail to do

- 22% power gains boosting revenue opportunities
- Reduced mechanical stress, which impacts maintenance and operating costs
- Actively mitigated survival conditions, helping to preserve capital investment
- This MARL architecture applies to other clean energy problems like wind energy, both for individual wind power generators and wind farms
- STrXL can help faster training of Transformers for RL with better performance
- Carnegie CETO 6 platform won the Phase 3 contract of the EuropeWave project to be deployed in BiMEP in Spain. This followed their win at the Phase 2 testing.

Thank You