

12 波力発電装置の最適制御への深層強化学習の適用

海洋先端技術系 * 梅田 隼, 藤原 敏文

1. 諸言

これまでに海技研では、発電機構に永久磁石式リニア発電機(Permanent Magnet Linear Generator: PMLG)を採用したリニア式波力発電装置(Wave Energy Converter: WEC)の開発に取り組んできた。リニア発電機は制御応答が良く、往復運動を回転運動へ変換なしに直接発電できる特徴をもつ。リニア式WECは可動部であるフロートの同調点では発電電力量が大きい、それ以外では低下する。そのため、広範な波浪条件下で発電性能を高めるために最適制御が行われる。これまでに提案されている最適制御法^{2,3)}はWECの力学モデルが用いられるが、様々な要因により実現現象と力学モデルの誤差が生じて、最適制御とならないことが想定される。そこで、本論では力学モデルを用いない手法として強化学習の1つであるDeep Q-Network(DQN)^{4,5)}を適用する。数値シミュレーション上でDQNによる制御時の平均電力、電力変動を示し、DQNによる制御の発電性能を確認する。

2. 数値シミュレーション

2.1 計算モデル

計算対象としたWECモデルの主要目を表-1に示す。将来的に検証が可能であることを念頭に置いて水槽模型サイズで計算した。本論では運動モデルを簡略化するために、フロート上下揺れの1自由度の方程式と仮定する。時間領域でのフロート上下揺れの運動方程式は式(2.1)から(2.3)で表す。変位や出力制限などの機械的制約を考慮していない。

表-1 計算対象の主要目

Part	Item	Value
Float	Outer diameter	0.350 m
	Inner diameter	0.114 m
	Depth	0.300 m
	Water-plane area	0.0860 m
	Draft	0.150 m
	Mass	12.9 kg
	Added mass	8.00 kg
PMLG	Thrust coefficient	37.93 N/A
	Winding resistance	2.115 Ω

$$(m_f + m_a)\ddot{z} + \int_0^t k_{33}(t-\tau)\dot{z}d\tau + c_{33}z = f_{ext} + f_{PTO} \quad (2.1)$$

$$k_{33}(\tau) = \frac{2}{\pi} \int_0^\infty N(\omega) \cos \omega \tau d\omega \quad (2.2)$$

$$f_{ext} = \sum_{i=1}^M E_i(\omega_i) \sqrt{2S(\omega)\Delta\omega} \sin(\omega_i t + \phi_i) \quad (2.3)$$

ここで、 m_f はフロートの質量、 m_a は周波数無限大での付加

質量、 z はフロート変位、上付きドットは時間微分、 k_{33} は造波減衰力係数 $N(\omega)$ を用いて式(2.2)で表されるメモリー影響関数、 c_{33} は復原力係数、 f_{ext} は波強制力、 f_{PTO} はリニア発電機の制御力である。 f_{ext} は波強制力係数 E_1 と波スペクトル $S(\omega)$ で表される M 個の規則波成分を重ね合わせることで得られる。 ϕ は各成分波の位相差である。フロートの造波減衰力、付加質量、波強制力は3次元パネル法の計算コードWAMIT⁶⁾で計算した。図-1、2にフロートの造波減衰力係数、波強制力係数の計算結果を示す。

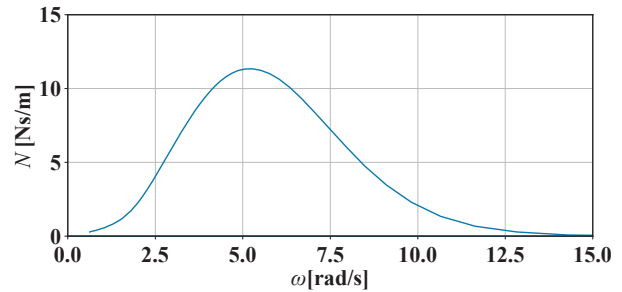


図-1 フロートの造波減衰力係数

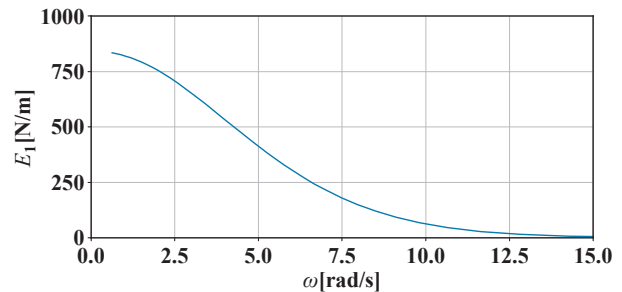


図-2 フロートの波強制力係数

PMLGの制御力および瞬時電力は式(2.4)から(2.7)で表す。制御力はフロート速度、変位に比例する力の和であり、制御パラメータ C_g, K_g を変化させて制御を行う。

$$f_{PTO} = -C_g \dot{z} - K_g z \quad (2.4)$$

$$P = f_{PTO} \dot{z} \quad (2.5)$$

$$P_c = R \left(\frac{f_{PTO}}{K_t} \right)^2 \quad (2.6)$$

$$P_n = P - P_c \quad (2.7)$$

ここで、 P は波からフロートに入力するパワー、 P_c は制御時生じる銅損、 R は巻線抵抗、 K_t は推力定数、 P_n は瞬時電力である。 P_n が正となる時はフロートの運動エネルギーの一部を電力に変換する発電機動作、負のときは電力を消費してフ

ロートを動かすモータ動作である。本論の最適化は P_n の時間平均値を最大化することである。

2.2 制御方法

使用したニューラルネットワーク構造を図-3 に示す。変位 z , 速度 $\dot{z}$, C_g, K_g , 波変位 ζ (5 変数) をサンプリング周期 0.1 s で 10 s 間のデータ 100 点を 1 フレームとすると, 入力データは過去 5 フレーム分を合わせた $5 \times 100 \times 5$ 次元のベクトルである。3 層の畳み込みニューラルネットワーク, 3 層の全結合型ニューラルネットワークの後, Q 値を出力する。結合層の後, 今回は畳み込み層の後にプーリング層は使用していない。活性化関数 $f(u)$ は, 隠れ層には式(2.8)の Rectified Linear Unit (ReLU), 出力層では線形関数 $f(u) = u$ を用いた。

$$f(u) = \begin{cases} 0 & (u < 0) \\ u & (u \geq 0) \end{cases} \quad (2.8)$$

ここで, u はノードからの総入力である。ニューラルネットワークの畳み込みニューラルネットワークのカーネルサイズおよびストライドは図-3 に示す値とした。また, 行動決定には確率 ε でランダム行動 $1-\varepsilon$ の確率で行動価値関数に基づいた行動選択を取る ε -greedy を採用した。 ε は 0.4 から試行回数を重ねることに 0.01 まで減少する。

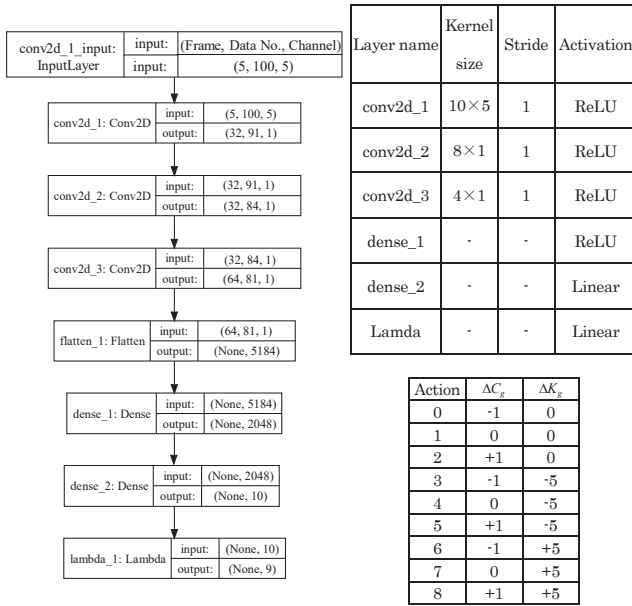


図-3 ニューラルネットワーク構造

DQN による制御では, 図-3 中の 9 つの行動 0~8 から行動価値関数が最大となる C_g, K_g の変化量の組を決定し, 式(2.9), (2.10)により更新する。

$$C_g \leftarrow C_g + \Delta C_g \quad (2.9)$$

$$K_g \leftarrow K_g + \Delta K_g \quad (2.10)$$

報酬は変位の極大値から次の極大値までの 1 周期間で得た電力量から, 式(2.11), (2.12)のように 4 段階にクリッピングした。報酬の重み付けはできなくなるが, 学習が進みやすくなる。

$$E_{sum} = \sum_{i=1}^m P_n^{(i)} \Delta T \quad (2.11)$$

$$r = \begin{cases} -1.0 & (E_{sum} / E_a < 0) \\ -0.2 & (0 \leq E_{sum} / E_a \leq 0.1) \\ 0.3 & (0.1 < E_{sum} / E_a < 0.5) \\ 1.0 & (E_{sum} / E_a \geq 0.5) \end{cases} \quad (2.12)$$

ただし, m は変位のゼロクロス点間の最大値から次の最大値までのステップ数(1 ステップ $\Delta T=0.1$ s), E_a はスケール係数である。模型スケールの PMLG の定格出力 20 W を想定し, E_a は 50% の 10 とした。報酬は性能および学習速度に影響するため, その設計方法については今後の検討が必要である。最後に DQN の制御と学習のフローチャートを図-4 に示す。10 s 間のデータから新しいフレームを作成後, 0.5 s 前の古いフレームを捨てる。ランダム行動か行動価値関数に基づいた行動選択どちらかを選択し, 行動後に報酬, 遷移先の状態を得る。学習は十分サンプルを保存してから行う。行動価値関数の更新量がなくなれば, 学習完了で終了する。

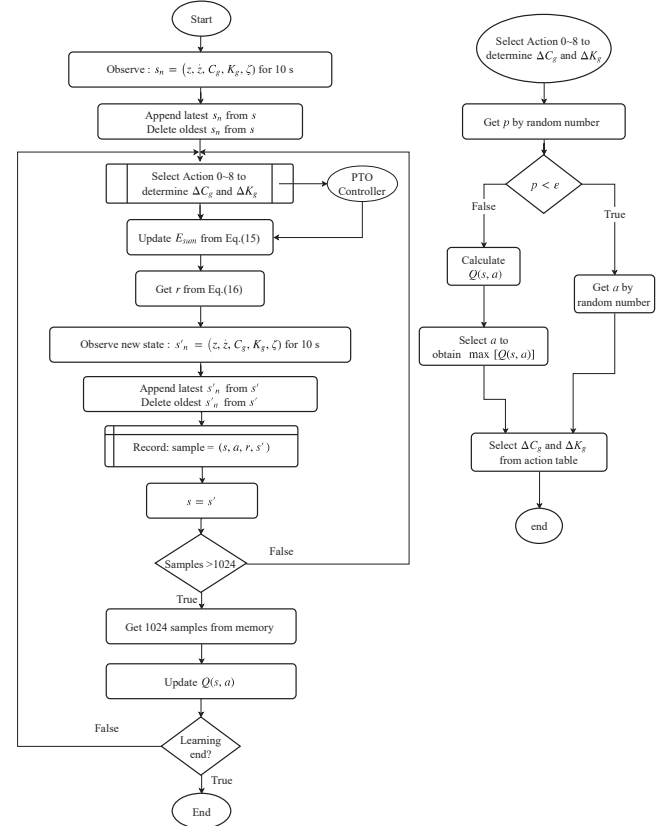


図-4 DQN 制御と学習のフローチャート

本論では DQN との比較に, Approximate Complex-Conjugate Control with Considering Copper Loss (ACL)²⁾を用いた。ACL は

正弦波が入射時、発電電力量を最大化する最適制御であると言える。ACLの制御パラメータは入射波の周波数に応じてフロートの造波減衰力、付加質量、発電機の推力係数、巻線抵抗から一意に決まる。不規則波の場合、入射波の角周波数を有義波周期から求めた角周波数に置き換えて計算し、一定値とした。

2.3 計算条件

規則波および不規則中のシミュレーションを実施し、平均電力を評価した。規則波の条件は波高 $H_w=0.1$ m、波周期 $T_w=1.0\sim 2.6$ sである。1 Episodeの評価時間は $T_w \times 20$ sで10波以上の時間とした。刻み時間 Δt は0.01 s、1 Episode中に最大20回の学習を行う。ただし、シミュレーション中に E_{sum} が-30を下回ると、失敗と判断し、Episodeを終了し、次のEpisodeに移る。 C_g, K_g の初期値は5 Ns/m、-100 N/mとし、正常にEpisodeが終了した場合の初期値は前Episode終了時の値を受け継いでスタートする。1条件につき750 Episodeの計算を行った。初期10 Episodeはランダム行動を選択し、Experience replay用にメモリーへ保存する。メモリー内から1024個のサンプルを取り出し、バッチ学習する。

不規則波の波変位はJONSWAP型波スペクトル⁷⁾を用いて式(2.13)から(2.18)で与えた。

$$\zeta(t) = \sum \sqrt{2S(f)\Delta f} \sin(2\pi ft + \phi) \quad (2.13)$$

$$S(f) = aH_s^2 T_p^{-4} f^{-5} \exp[-1.25(T_p f)^{-4}] \gamma^b \quad (2.14)$$

$$a = \frac{0.0624}{0.23 + 0.0336\gamma - 0.185(1.9 + \gamma)^{-1}} \quad (2.15)$$

$$b = \exp[-0.5(T_p f - 1)^2 \sigma^{-2}] \quad (2.16)$$

$$T_p = 1.05 T_s \quad (2.17)$$

$$\sigma = \begin{cases} 0.07 & (f \leq f_p) \\ 0.09 & (f > f_p) \end{cases} \quad (2.18)$$

ここで、 ϕ は成分波の位相差、 f は成分波の周波数、 γ はスペクトルの形状パラメータ、 H_s は有義波高、 T_s は有義波周期である。不規則波の波条件は $H_s=0.1$ m、 $T_s=1.0\sim 2.4$ s、 $\gamma=1.0$ とした。不規則波の初期位相はEpisode毎に乱数で与えて、異なる時系列を持つ波変位データを作成した。1 Episodeの評価時間は $T_s \times 120$ sで100波以上を含む時間とした。1つの有義波周期につき3500 Episodeの計算を行った。初期50 Episodeはランダム行動を選択し、Experience replay用にメモリーへ保存する。メモリー内から2048個のサンプルを取り出し、バッチ学習する。その他の条件は規則波と同様である。

3. シミュレーション結果

3.1 規則波中の平均電力

規則波中の平均電力、制御パラメータの時間平均値を図-5、6に示す。横軸は波周期 T_w 、DQNの平均電力は十分学習が進んだ700 Episode後の50 Episode間の時間平均値を平均化した値である。規則波ではACLに対して、DQNは全波周期で、

約80%程度の平均電力となった。 C_g はACLより大きい、傾向はACLと一致する。 K_g は $T_w=1.4$ sまではACLに近い傾向を示すが、 $T_w=1.6$ s以降では、ACLより大きい。DQNでは報酬のクリッピングにより、ある一定以上の発電電力量は同じ評価となる。そのため、 C_g や K_g の平均値は概ね近い値を示すが、平均値付近を変動するため、規則中では最大であるACLの平均電力に到達しなかったと考えられる。

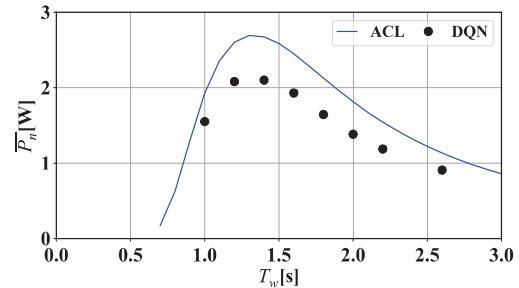


図-5 規則波中での平均発電電力の比較($H_w=0.1$ m)

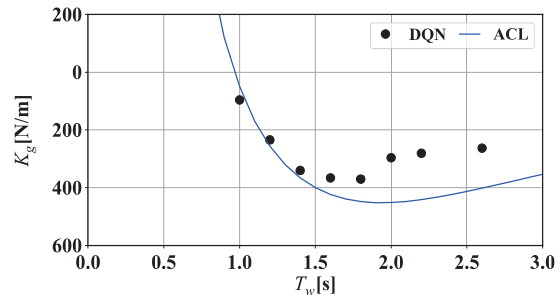
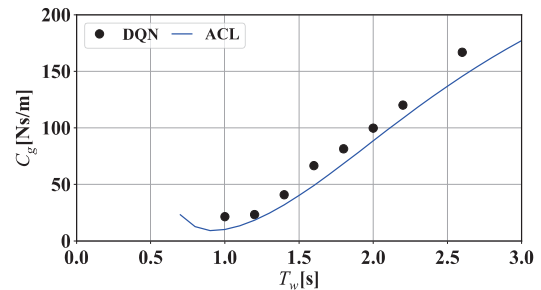


図-6 規則波中の制御パラメータの時間平均値の比較($H_w=0.1$ m)

3.2 不規則波中の平均電力

図-7は $T_s=1.2, 1.4$ sでのEpisode毎の平均電力を示す。薄い点線はEpisode毎の結果、黒線は50 Episodeの平均値を示す。Episodeを重ねるごとに学習していき、平均電力が向上している。次に波周期ごとの平均電力を図-8に示す。1 Episodeの評価時間は約120~336 sで、平均電力評価に十分な長さでない。十分学習が進んだ3450 Episode後の50 Episode中で P_n の全区間の時間平均値を、平均化した値を平均電力とした。ACLも同様に50 Episodeの平均電力である。エラーバーは50 Episode中の平均電力の標準偏差、図中の数字はACLの平均電力に対するDQNの平均電力の比を表す。 $T_s=1.0$ sではDQNの平均電力はACLより大きい、 $T_s=1.2$ s以上ではDQNの平均電力の方がわずかに小さい。規則波中ではDQNの平均電力はACLの平均電力の約80%であったが、不規則波中ではACLに対して約96~110%と近づく。50 Episodeの標準偏

差は ACL と同程度である。

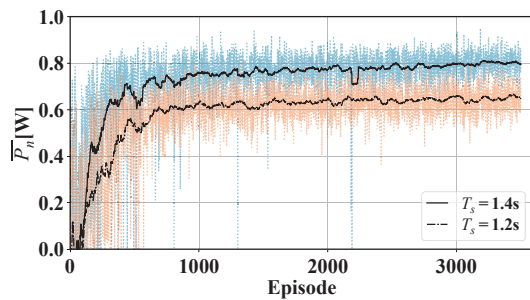


図-7 Episode 毎の平均発電電力の変化

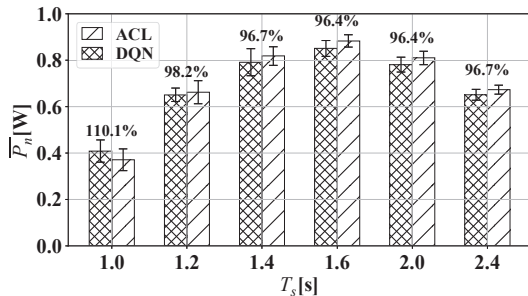


図-8 不規則波中での平均電力 (有義波高 0.1 m, エラーバー: 標準偏差)

波力発電装置では制御によって平均発電電力が大きくなるとともに、電力変動は小さいことが望ましい。図-9, 10 に瞬時電力の標準偏差およびピーク値と平均値との比 (Peak-to-Average Power Ratio: PAPR), それぞれの 3450 Episode 後の 50 Episode 間の平均値を示す。図中のエラーバーは標準誤差を表し、図中の数字は ACL に対する DQN の値の比である。どの波周期でも ACL より DQN の標準偏差は小さいため、ACL に比べて DQN のほうが電力変動は小さいといえる。PAPR は最も平均電力が高い $T_s=1.6$ s で約 30%減少し、 T_s が長くなると、ACL との差は小さくなる。長周期では ACL より最大電力は低いが、平均発電量は小さいため、PAPR では概ね同じになった。 $T_s=1.0$ s では ACL のほうが低いが、PAPR は他の周期に比べて相対的に小さいため、DQN の方が優れていると言える。PAPR は小さいほど、電力変換機のサイズが小さくなり、発電設備へのコストが低くなる。

4. 結 言

発電機構にリニア発電機を採用した波力発電装置の制御方法に、強化学習の 1 つである Deep Q-Network(DQN)を適用した。規則波中および不規則波中の平均電力について従来の制御法である ACL と比較し、以下の結果を得た。

- ・規則波中において力学モデルを用いることなく、平均発電電力は規則波中の最適理論解の ACL の約 80%となった。
- ・不規則波中では力学モデルを用いることなく、DQN の平均発電電力は ACL の約 96~110%となった。
- ・瞬時電力の標準偏差は ACL より小さい。ピーク値と平均値の比 (PAPR) を比較すると、ACL と同等かそれ以下と

なった。

DQN は ACL と比べて発電性能は劣るが、制御のための力学モデルが必要なく、電力変動は小さく抑えられる制御法といえる。

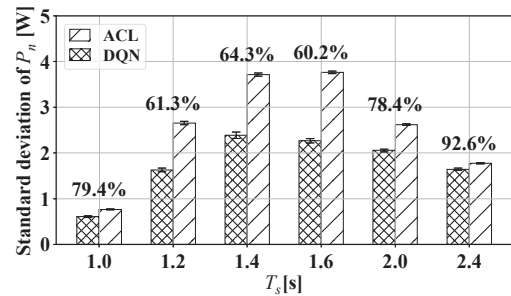


図-9 瞬時電力の標準偏差 (有義波高 0.1 m, エラーバー: 標準誤差)

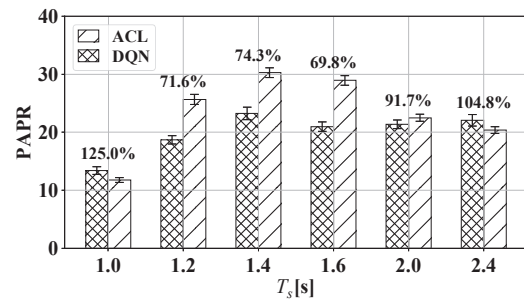


図-10 瞬時電力のピーク値と平均値との比 (有義波高 0.1 m, エラーバー: 標準誤差)

参 考 文 献

- 1) 国立研究開発法人新エネルギー・産業技術総合開発機構: 平成 26 年度~平成 29 年度成果報告書 海洋エネルギー技術研究開発 次世代海洋エネルギー発電技術開発 リニア式波力発電, (2018), 20180000000866.
- 2) De La Villa Jaén, A., García-Santana, A. and Montoya-Andrade, D.E.: Maximizing Output Power of Linear Generators for Wave Energy Conversion, International Transactions on Electrical Energy Systems, Vol. 24, No. 6(2014), pp. 875-890.
- 3) Hals, J., Falnes, J. and Moan, T.: Constrained optimal control of a heaving buoy wave-energy converter, Journal of Offshore Mechanics and Arctic Engineering, Vol.133, No.1(2011), 011401.
- 4) 牧野貴樹ほか: これからの強化学習, 森北出版(2016).
- 5) Mnih, V., et al.: Human-level control through deep reinforcement learning, Nature 518.7540(2015) 529.
- 6) Lee, C.-H.: WAMIT Theory Manual., Report No. 95-2, Dept. of Ocean Engineering, Massachusetts Institute of Technology (1995).
- 7) 合田良実: 数値シミュレーションによる波浪の標準スペクトルと統計的性質, 海岸工学講演会論文集, 第 34 巻(1987), pp. 131-135.