

令和元年(第19回)海上技術安全研究所研究発表会

波力発電装置の最適制御への 深層強化学習の適用

Sli.doのイベントコード
Event code : NMRI_AI_03



海洋先端技術系 再生エネルギー研究グループ
梅田 隼



国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所

1

発表内容

2

- 研究背景と目的
- 波力発電装置の最適制御
 - ・制御対象と最適制御問題
 - ・強化学習による制御
- 発電／学習シミュレーション
 - ・規則波中の計算結果
 - ・不規則波中の計算結果
- 結言



国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所

2

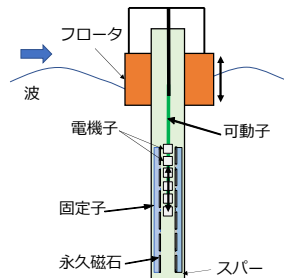
研究背景

3

- ポイントアブソーバー型の波力発電装置（WEC）
 - ・制御により広範な波浪条件下で発電が可能となる
 - ・数多くの制御法が考案
- 発電機構(PTO)に永久磁石式リニア発電機
 - ・利点：上下運動を直接電力に変換，機構が簡便



© 2008 AQUARET


国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所

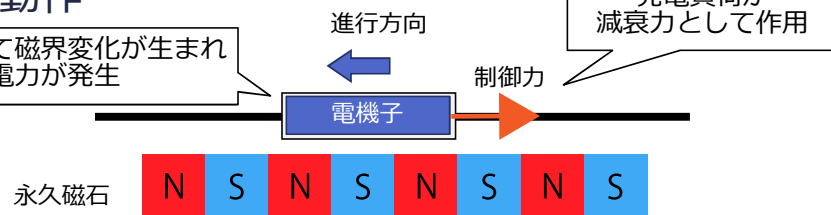
3

リニア発電機の発電と制御

4

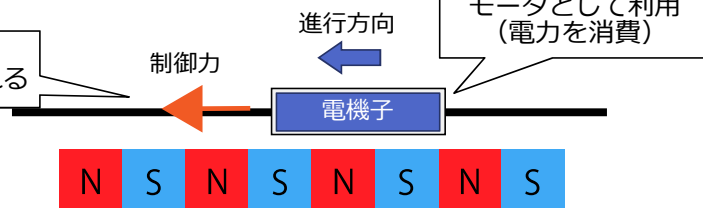
●発電機動作

運動によって磁界変化が生まれ
起電力が発生



●モータ動作

電力の一部は
内部抵抗により失われる


国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所

4

WECの最適制御

5

◆モデルベース

海技研で実施

最適ゲインによるPI制御
モデル予測制御

✓ 最適制御の有効性を確認

✓ 制御モデルの精度が制御性能に大きく影響
> 考慮していない物理現象や経年劣化



モデルの高精度化

H^∞ 制御, 適応制御, ロバストモデル予測制御

別のアプローチとして・・・

◆モデルフリー

ファジィ制御

ニューラルネットワーク制御 (強化学習)

●目的：モデルフリーの制御法での発電性能の把握

- ・強化学習の一つであるDQNをWECの最適制御に適用し,
制御モデルに基づいた最適制御と性能を比較する



国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所

5

発表内容

6

●研究背景と目的

●波力発電装置の最適制御

- ・制御対象と最適制御問題
- ・強化学習による制御

●発電／学習シミュレーション

- ・規則波中の計算結果
- ・不規則波中の計算結果

●結言



国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所

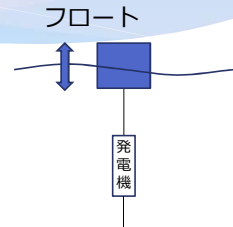
6

制御対象

7

- WECの仕様
 - フロート(円柱)のみ考慮
 - 水槽模型サイズ
 - 変位, 出力などの制限なし

項目	値
直径	0.35 m
喫水	0.150 m
質量	12.9 kg
固有周期	1.2 s



- 上下運動のみを仮定 (1自由度)

流体力はパネル法により別途計算

メモリー影響を考慮した時間領域の運動方程式

$$(m_f + m_\infty)a + \int_0^t k_{33}(t-\tau)v d\tau + c_{33}z = f_{ext} + \overset{\text{制御力}}{f_g}$$



国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所

7

WECの最適制御問題

8

- 目的 : **平均電力の最大化**

電力 $P_n = (\text{制御力}) \times (\text{速度}) - (\text{銅損})$ → 発電機の内部抵抗の損失
流れる電流 (制御力) の2乗に比例

$$\text{Maximize: } \bar{P}_n = \frac{1}{T} \int_0^T f_t(t)v(t) - \delta f_t^2(t) dt$$

- 入力 : リニア発電機による制御力

$$f_t(t) = C_g v(t) + K_g z(t)$$

最適応答を与える C_g, K_g を求める

規則波中では浮体の流体力から最適 C_g, K_g を一意に決定

(制御法 : ACL)

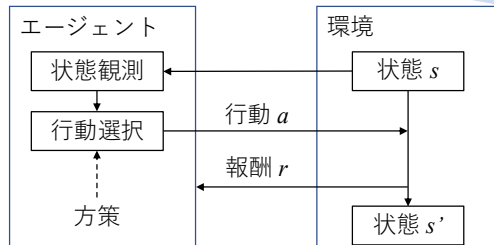


国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所

8

強化学習 (Q-Learning)の概要

9



将来にわたって得られる報酬を最大化する最適行動方策を経験で学習

最適行動方策を行動価値関数 $Q(s,a)$ を用いて評価
 $Q(s,a)$: 状態 s で行動 a を取った後の期待値

どの状態で、どう行動したら、どういう報酬が得られるのかを経験で学習

$$Q(s, a) \leftarrow \underbrace{Q(s, a)}_{\text{行動価値の推定値}} + \alpha \left[\underbrace{r + \gamma \max_{a'} [Q(s', a')]}_{\text{試行サンプリングによる価値推定}} - Q(s, a) \right]$$

行動価値の推定値 試行サンプリングによる価値推定

Qの値が最も高い行動を取り続けることで最適行動となる



国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所

9

WECへの適用

10

- s が連続値で定義する環境だと、状態と行動に対して定義される Q を保存するための領域が無限に必要となる
- Q をニューラルネットワーク(NN)で関数近似
 - ・そのままではうまく学習できない
- 多層NNでも上手く学習できるテクニックを組み合わせたものをDeep Q-network (DQN) と呼ぶ
- DQNを用いることでWECの最適制御を可能とする



国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所

10

DQNによる制御の設計方針

11

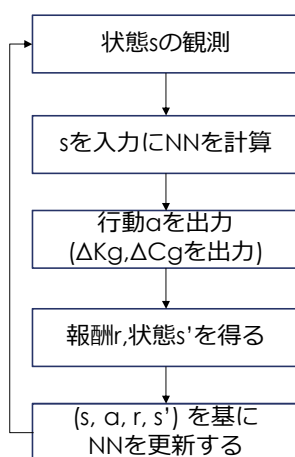
- 設定に必要な条件
 - 入力データ (状態 s)、行動 a 、報酬関数 (制御モデルが不要)
- 用いた学習効率を上げる手法
 - Experience replay
 - 報酬のクリッピング
 - Fixed-target Q
 - Dueling DQN
 - Huber損失 etc.


国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所

11

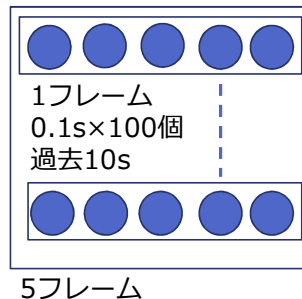
DQNによる制御

12

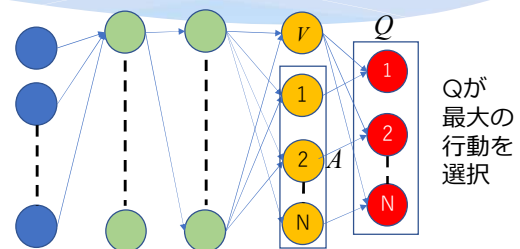


状態 S

5×100×5次元ベクトル
変位, 速度, K_g , C_g , 水位



Q関数：ニューラルネットワーク



$$\text{制御行動： } C_g \leftarrow C_g + \Delta C_g$$

$$K_g \leftarrow K_g + \Delta K_g$$

$$\Delta C_g = -1, 0, 1$$

$$\Delta K_g = -5, 0, 5$$

3×3の9通りから選択



データ2048個でバッチ学習


国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所

12

報酬設定

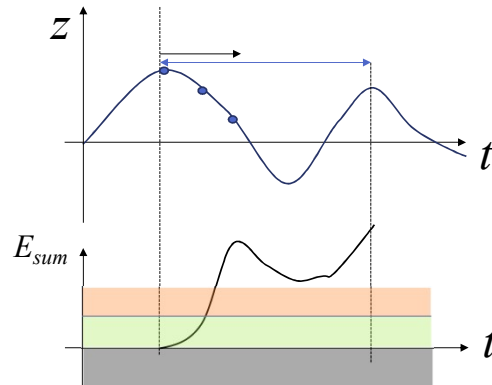
13

- 変位の極大値から次の極大値までの電力から決定

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

$$E_{sum} = \sum_{i=1}^m P_n^{(i)}$$

P_a はスケーリング係数で10とした
状態での報酬の重みづけはできないが、学習は早く進む
性能向上に向けて、さらなる検討が必要



国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所

13

発表内容

14

- 研究背景と目的
- 波力発電装置の最適制御
 - 制御対象と最適制御問題
 - 強化学習による制御
- 発電／学習シミュレーション
 - 規則波中の計算結果
 - 不規則波中の計算結果
- 結言



国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所

14

学習環境条件

15

- 規則波の条件
 - 波高 0.1 m
 - 波周期 0.8 s ~ 2.6 s
- 1条件の試行回数
 - 750 回
- 1試行の計算時間
 - (波周期) × 20 s
- 不規則波の条件
 - 有義波高 0.1 m
 - 有義波周期 1.0 ~ 2.4 s
 - Jonswap型スペクトル
- 1条件の試行回数
 - 3500 回
- 1試行の計算時間
 - (有義波周期) × 120 s

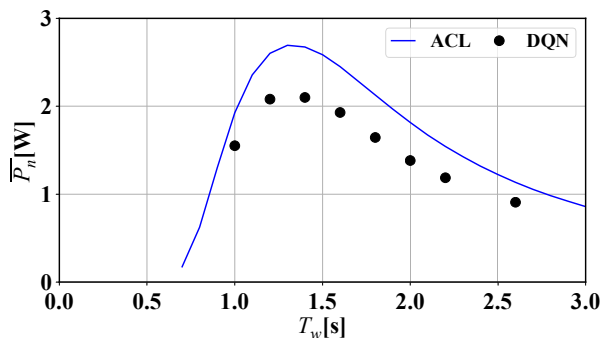


国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所

15

規則波中の平均電力

16



- 入射波が同じなど環境条件が同じになる場合が多いため、過学習によりNNが破綻する場合があった
- DQNの平均電力はいずれの波周期でも規則波最適解であるACLの約80%

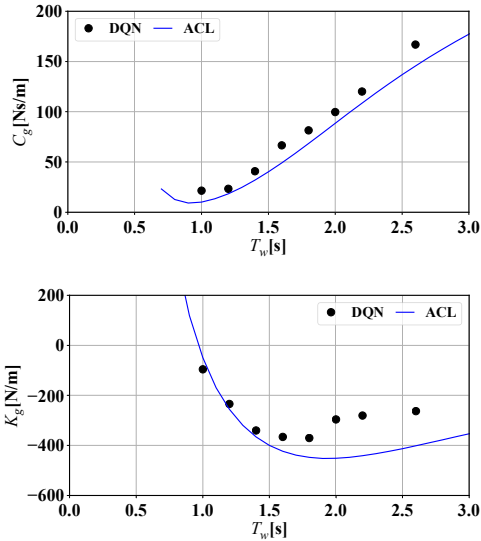


国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所

16

制御変数の比較

17



- C_g はACLより大きいですが、変化の傾向はACLと一致
- K_g は1.4 sまではACLに近い傾向を示すが、1.6 s以降では差が大きい
- DQNでは報酬のクリッピングやNNなど多くの近似を行っているため、最適解と差が生じる

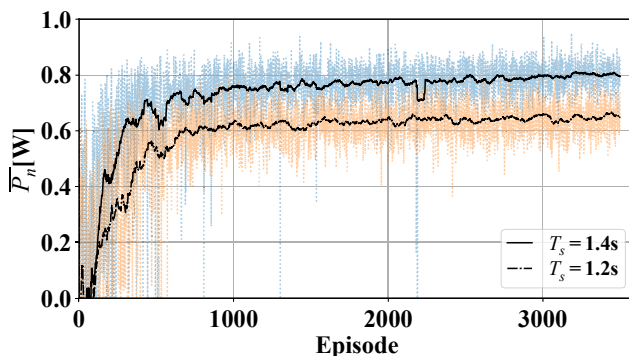


国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所

17

試行毎の平均電力（有義波周期：1.2, 1.4 s）

18



黒線は過去50試行毎の平均値

- 試行を重ねるごとに学習していき、平均電力が向上している
- 試行毎に平均電力のばらつきがある
- ACLとの比較では50試行の平均値を使用

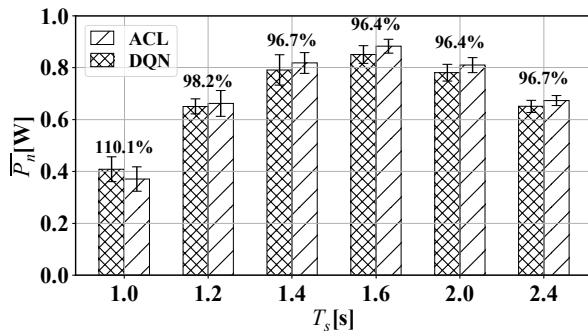


国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所

18

不規則波中の平均電力

19



エラーバーは50試行の標準偏差を表す



国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所

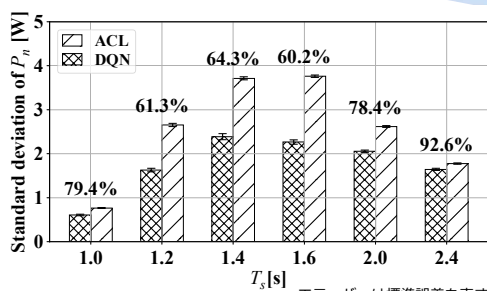
DQNは・・・

- 不規則波中ではACLに対して約96～110%
- 波周期1.0sではACLより大きい。その他の周期ではわずかに劣る。
- 50 試行の標準偏差はACLと同程度。試行毎のばらつきはDQNによるものではなく波エネルギー変動に由来

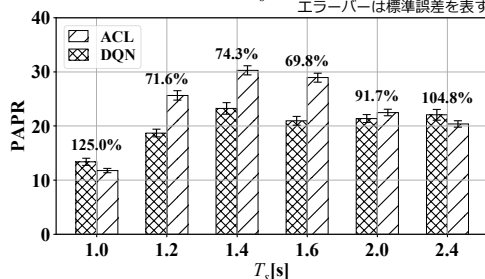
19

不規則波中の電力変動

20



エラーバーは標準誤差を表す



国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所

- 瞬時電力の標準偏差はACL > DQN
- ピーク値と平均値との比 (PAPR)は最も平均電力が高い1.6 sでACLより約30%減少
- 長周期では差は小さい
- DQNはACLと比べて電力変動は小さい制御

20

結言

21

- 規則波中では, DQNの平均電力は最適理論解のACLの約80%
- 不規則波中では, DQNの平均電力はACLの約96~110%. DQNはACLと比べて平均発電性能はわずかに劣るが, WECの制御・物理力学モデルを使うことなく制御できる
- DQNは, ACLと比べて電力変動を小さく抑えられるため有用である



国立研究開発法人 海上・港湾・航空技術研究所 海上技術安全研究所