

# ごみ焼却ストーカ炉における燃料供給の安定化に向けた 強化学習による制御の有効性検証

## Feasibility Study of the Effectiveness of Control Using Reinforcement Learning for Stabilizing Fuel Supplying in Stoker Furnace



池田 孝\*<sup>1</sup>  
Takashi Ikeda

佐々木 俊也\*<sup>2</sup>  
Shunya Sasaki

馬場 由羽貴\*<sup>2</sup>  
Yuki Baba

瀬戸口 稔彦\*<sup>3</sup>  
Toshihiko Setoguchi

エネルギーミックス実現に向けて有望視されているごみ焼却プラントでは、プラントの収益力強化に向けた自動運転技術が求められている。ごみ焼却プラントの燃料供給装置は、燃料性状のばらつきの影響が大きいために、有望な制御技術が確立されていない。本報では、燃料性状の変化を踏まえながら、柔軟に運転状態に応じた適切な運転操作を選択できると期待される、強化学習に基づく燃料供給制御技術を開発し、燃料供給装置を模擬したシミュレータ上において、既存運転条件に比べて、過剰供給回数を56%、過少供給回数を54%低減させることができた。

### 1. はじめに

ごみ焼却プラントは、エネルギーミックス実現に向けて、エネルギーの多様化と分散化、自給率の向上に大きく寄与する製品であり、作業員の負荷低減と高効率化によるランニングコストの低減によって、プラントの収益力強化を向上させる自動運転技術が求められている。一方で、ごみの燃料性状のばらつきによる燃料の供給量の変動によって、炉内の異常燃焼や消炎が発生し、意図せず発生する一酸化炭素や窒素酸化物などの有害排出物対策に対して、作業員の手動介入が必要となる課題が生じている。また、燃料の供給量の変動により、発電のための蒸発量の変動するため、定格負荷を維持するという観点でも作業員の手動介入が必要となり、自動化運転適用の妨げとなっている。

本報では、ごみ焼却プラントの燃料供給装置に対して、気候などの短期的、季節などの長期的な燃料性状の変化を踏まえながら、運転状態に応じた適切な運転操作を柔軟に選択できると期待される、強化学習に基づく燃料供給制御技術を開発し、燃料の供給量の変動を抑制できることを検証した結果を報告する。

### 2. ごみ焼却ストーカ炉と燃料供給装置

ごみ焼却ストーカ炉は、世界で最も普及しているごみ処理方式であり、ごみはテーブル状の往復動する燃料供給装置によって、焼却炉(ストーカ)に供給される。ごみは、ストーカ上で、移動床によって攪拌されながら、後流(炉尻)に搬送され、その搬送過程で床底部から吹き込まれる空気と接触させられることで燃焼する(図1)。ストーカは上流から、乾燥段、燃焼段、後燃焼段に分かれ、それぞれの場所で、乾燥、熱分解反応、チャー燃焼反応が直列的に起こる。ごみ焼却炉の燃料供給装置は、ストーカの上流部、ホップの下部に設置され、油圧によって往復動し、間欠的に、燃料であるごみをストーカに供給する。燃料供給装置の操作パラメータは、ストローク、速度、前進/後退の間の停止時間である。ホップ内のごみは、その自重によって圧密されるが、季節、気

\*1 技術戦略推進室 ビジネスインテリジェンス&イノベーション部 主席部員 技術士(機械部門)

\*2 デジタルイノベーション本部 EPI 部

\*3 総合研究所 燃焼研究部 主幹研究員 工学博士

候、地域の差による燃料の水分量のばらつき、また燃料の種類による形状や付着力のばらつきによって、同じ操作パラメータであっても燃料の供給量にばらつきが生じる。従って、所定のルールに基づく制御は困難であり、短期的・長期的な燃料性状の変化を伴う運転状態に対して、柔軟に追従できる燃料供給装置の制御技術が求められている。

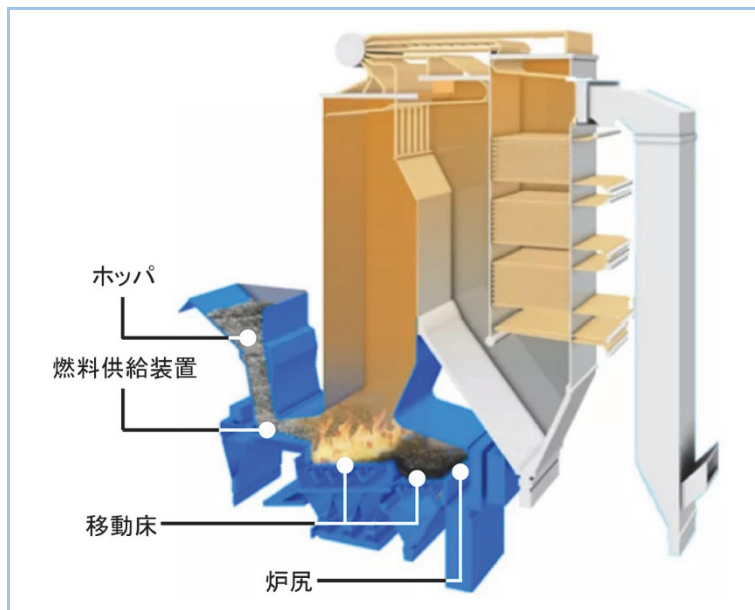


図1 ごみ焼却ストーカ炉

### 3. 強化学習を用いた燃料供給装置の制御技術の開発

強化学習の最も基本的なアルゴリズムは Q-learning である。Q-learning では、現在の状態に対して、学習によって得られた報酬の期待値である Q テーブルに基づき取るべき行動が決定される。Q-learning によって適切な行動を決定するには、Q テーブルの状態と行動の組合せの網羅性を高める必要がある。そこで、今後の実機展開を見据え、本報では、燃料供給装置の制御技術に、Q テーブルをニューラルネットワークで予測する強化学習 Deep-Q-Network<sup>(1)</sup>を適用した。Deep-Q-Network は、都度、制御した結果をニューラルネットワークで学習し、Q テーブルに反映することによって、Q テーブルを構築するデータの過不足を補うことができ、また常に Q テーブルを、最新の環境条件に適応するように更新できると期待される。

本報では、燃料供給装置の挙動特性を再現したシミュレータを用いて、Deep-Q-Network の燃料供給装置の制御技術への適用性を検証した。シミュレータは二次元であり、ホッパから燃料供給口に至る空間を離散化し、離散化した検査体積の間の物質収支を解くことで、燃料の供給量を計算している。またシミュレータでは、燃料供給装置による圧密、せん断を考慮しており、性状のばらつきによる付着力のばらつき、またその結果として、本報で強化学習が制御対象とする燃料供給量のばらつきが定性的に考慮できる。

強化学習の各設定項目を表1に示す。強化学習を用いて適切な制御を実現するには、制御対象の現在の状態を定義する観測項目と、任意の行動の結果得られた評価値に対する報酬及び罰則の設定を行う必要がある。観測項目については、今後の実機への適用を見据えて、画像分析やセンサの設置などで、直接的/間接的に把握可能な項目とした。観測項目、また報酬と罰則によって制御のための行動が変化するため、検証では、観測項目と報酬及び罰則について表2に示す3ケースで検証を行った。表2の3ケースは、事前の検討によって良好な制御が期待されたケースである。

表1 強化学習の各設定項目

項目	値
NN 中間層	3
NN 中間層ノード数	16
NN 活性化関数	Relu(中間層), 恒等関数(出力層)
NN 最適化関数	Adam 法
$\epsilon$ -greedy ランダム率	1.0(開始時) $\rightarrow$ 0.1(終了時)
割引率	0.99

表2 検証ケースに対する観測項目と報酬設計

case No.	観測情報	供給量による報酬 ※1	超過供給による罰則 ※2	超過供給による罰則 ※3
case1	燃料供給量 現在のストローク	+5	-1	0
case2	↑	+1	0	0
case3	↑	+1	0	-1

※1:1 ストロークあたりの供給量が目標値 $\pm$ 10%以内の場合に報酬付与

※2:5 ストロークあたりの過少供給回数(目標値 $\times$ 50%未満)率 $\times$ 罰則値  
(5 ストローク中, 5 ストロークの過少供給で罰則-1)を付与

※3:1 ストロークあたりの供給量が目標値+20%より大きい場合に罰則値を付与

#### 4. 燃料供給装置制御システムの有効性の検証結果と考察

本報では、小型ごみ焼却ストーカ炉を対象に、燃料供給装置の有効性の検証を行った。検証では、燃料の過剰及び過少供給を評価するため、目標とした燃料供給量は 90t/d に対し、燃料供給量が目標供給量の 10%以上となった過剰供給の回数及び、燃料供給が目標値の 50%未満となる過少供給の回数を評価した。また検証では、燃料供給量への影響が最も大きいと考えられる、燃料供給装置のストロークを制御対象とした。比較対象は、既存運転条件相当として、ストロークは一定値とした。燃料性状の変化として、燃料であるごみの種類の違いを模擬した密度及び、ごみの水分量の違いを模擬したホッパ内の付着力を時間経過とともに図2のように、燃料供給装置が 10 往復する度に変更した。ごみ焼却炉では、ピットからクレーンで燃料がホッパ内に投入されており、クレーンによる燃料の投入の度に密度及び付着力が変わるとした。計算初期の影響を除くため、燃料供給装置が 100 往復した時点から 150 往復する時点までを評価期間とした。

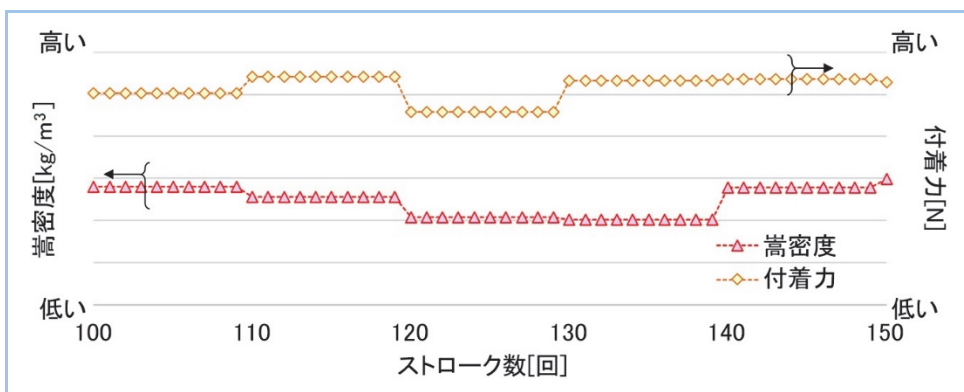


図2 計算で与えた燃料性状の変化

ストロークごとの供給量の case3 の結果を図3に示す。また過剰供給の回数及び過少供給の回数の、評価期間における積算結果を表3に示すが、表2の 3 ケースの結果は、過剰/過少供給回数の低減の観点で、いずれも同様の良好な制御効果を示した。図3では3 ケースの中で最も良好な結果を示した case3 を例として示す。比較対象とした、ストローク一定の既存運転条件相当では、図3に示すように、過剰供給と過少供給が繰り返されていたが、Deep-Q-Network を適用した case3 では、密度及び付着力が変わる 10 往復毎にストロークが Deep-Q-Network による制御に

よって変更された。Deep-Q-Network による制御の結果、既存運転条件に比べ過剰供給回数が最大 56%低減し、また過少供給回数が最大 54%低減した。

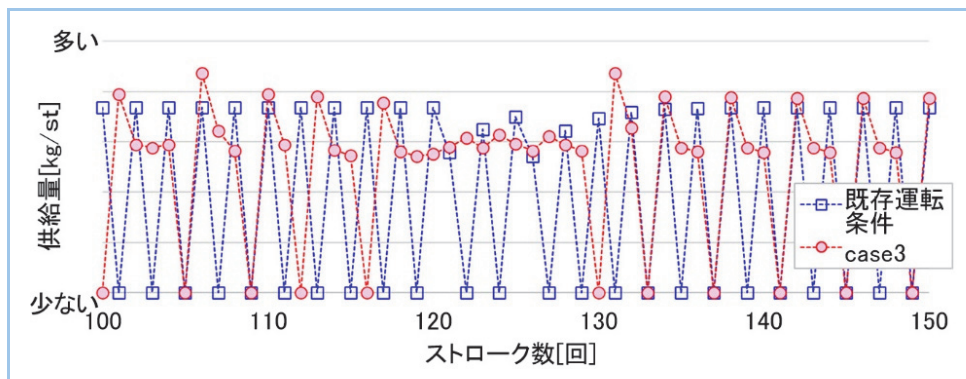


図3 ストローク毎の供給量

表3 過剰供給の回数及び過少供給の回数の積算結果

	過少供給回数	過剰供給回数
既存運転条件	25	22
case1	13 (▲48%)	13 (▲41%)
case2	13 (▲48%)	14 (▲36%)
case3	11 (▲56%)	10 (▲54%)

カッコ内数字は、既存運転条件を 100%とした際の低減割合

密度及び付着力の変化に対し、case3 において Deep-Q-Network による制御で変更された 10 往復毎の燃料供給装置のストロークの変化量を評価した結果、運転員による操作の考え方に対する矛盾はなく、単純化された条件ではあるものの、Deep-Q-Network によって適切な燃料供給装置の制御が実現できたと考える。

実機では、より多くの複雑な要因が燃料供給量のばらつきに影響しているが、製品化に向けてその一つ一つの検証及び評価を網羅的に実施することは、膨大な時間と労力を必要とし、また市場の確保の観点からも望ましくない。また、本報で示した制御技術の高度化を図るには、実環境における運用に基づくデータ取得が必要不可欠である。従って、今後、本技術の製品化に向けては、製品として求められる最小限の要求仕様を開発目標として早急に実機に展開を行い、実環境における運用の中でデータを蓄積しながら、本報の技術の更新を行っていく。

## 5. まとめ

ごみ焼却プラントで求められている、作業員の負荷低減と高効率な運用のための自動運転技術の確立に向けて、燃料の供給量の変動を抑制できる燃料供給装置の制御技術の開発を推進中である。ルールに基づく制御が困難であり、短期的・長期的な燃料性状の変化を伴う運転状態に対して、柔軟に追従できる燃料供給装置の制御技術として有望視されている、強化学習を用いた制御技術について開発を行い、シミュレータを用いた検証によって、既存運転条件に比べ過剰供給回数が最大 56%低減し、また過少供給回数が最大 54%低減し、強化学習の有効性を確認した。

ごみ焼却プラントは、エネルギーミックス実現に向けた、エネルギーの多様化と分散化、自給率の向上に大きく寄与する製品であり、競合他社も積極的な技術開発を続けていく。

## 参考文献

(1) Volodymyr Mnih, et al., Human-level control through deep reinforcement learning, Nature 518, 529-533.