

強化学習を用いた活性汚泥法における運転条件の自律最適化手法の検討

長岡工業高等専門学校 非会員 ○井戸村 幸毅, 非会員 徳嵩 友志
正会員 川上 周司, 非会員 土田 勝範

1. 背景と目的

下水道事業は国内の年間消費電力量の約 0.7%を占め¹⁾, 活性汚泥法の処理場では反応タンクへ空気を送る送風機が焼却設備を除いて最も消費電力の大きな設備である²⁾. 曝気量の制御には溶存酸素 (DO) 濃度を一定に保つ DO 制御が一般的に用いられているが, 硝化反応の進行状況に基づく制御ではないため必要以上に空気を供給する時間帯が生じる³⁾. 必要量を超えた曝気に使われる送風機の電力は処理に寄与しないため, 処理水質を保ちながら曝気を必要な分に近づけられれば, 処理場の消費電力と CO₂排出量を削減できる.

強化学習は, 試行錯誤を通じて状態に応じた行動を学習する手法であり, 水質と曝気の間関係をあらかじめ数式で表さなくても, 計測値とその結果から曝気の方法を学習できる. このため, 負荷変動に応じた曝気制御への応用が期待される. 下水処理では, 曝気や薬品注入などの運転制御に強化学習を適用した研究が報告されており^{4,5)}, 従来の DO 制御より運転コストを削減できた例もある⁵⁾. しかし, これらはシミュレーション上での評価であり, 水資源管理分野のレビューでも学習は主にシミュレーション環境で行われ実環境との差を埋めることが課題とされている⁶⁾. 現状では, 実規模の排水処理技術において曝気制御を強化学習から運用した例は報告されていない.

そこで本研究では, 活性汚泥法に対し強化学習を用いた運転条件の自律最適化を検討する. 具体的には, 実験室規模のリアクターと, 水質センサーから PC へ直接データを取得して曝気を制御する実験系を構築し, リアクターの処理水を実処理水と対応づける条件を整理して実環境に基づく状態, 行動を定義したうえで, 強化学習モデルを構築することを目的とした. 本報では, 実験系の構築状況と状態, 行動の定義に向けた検討結果を報告する.

2. 実験方法

2.1 活性汚泥リアクター

曝気槽と沈殿槽を一体化した連続式リアクターに人工培地を連続供給し, 越流した混合液を沈殿槽で固液分離し, 汚泥は適宜曝気槽へ返送した. 現在は常時曝気で運転しており, これを比較の基準とした (表 1).

表 1 リアクターの運転条件

容積	曝気槽約6.6 L, 沈殿槽約0.5 L
流量・HRT	1 L/d, 曝気槽HRT 約6.6 d
培地COD	1,800 mg/L (COD _{Cr})
培地組成	グルコース, 酢酸Na, ペプトン, 酵母エキス, 肉エキス, 尿素, 無機塩類
運転	常時曝気, 汚泥返送は約10日に1回
水温	24 °C (恒温機内)

2.2 計測・制御システム

システム構成を図 1 に示す. 水質の計測には, 試薬を用いずに COD や硝酸などを測定できる UV-VIS 分光式の水質センサーである NiCaVis 705 IQ NI を用いた. 沈殿槽の上澄みをポンプで送って計測し, 沈殿槽へ返送した. 計測値はコントローラの内蔵 Web サーバーから自作プログラムで取得した. データ取得と曝気ポンプの制御はエッジ PC が担い, 学習・推論はサーバー PC が担った, 両者は FastAPI を介し通信した. 通信断などの異常時は曝気を最大に固定するように設定した.

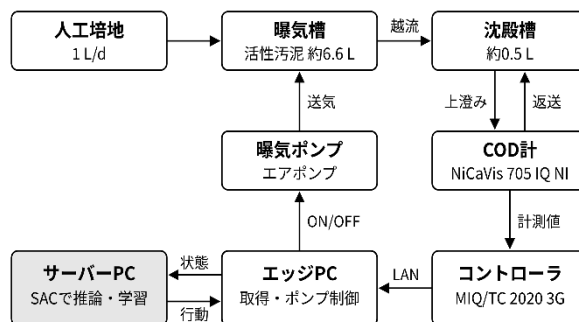


図 1 システム構成

2.3 強化学習モデルの構成

実装には Gymnasium と Stable-Baselines3 を使い、連続値の行動を扱える Soft Actor-Critic (SAC) ⁷⁾を暫定的に採用した。状態を処理水の COD 計測値、行動を一定時間内の曝気ポンプの稼働時間、報酬を処理水 COD と曝気時間に基づく値とする暫定案とし、アルゴリズムの選定を含め実データを踏まえて定義することとした。

3. 結果と考察

3.1 実験系の構築状況

沈殿槽の上澄みを水質センサーに通す機構と曝気を制御するプログラムを作成し、計測値を状態として取得して曝気時間を出力する学習用の実験系を構築した。PC から計測値を 1 分間隔で自動取得できること、試験環境でエッジ PC-サーバーPC 間のデータ送受信ができることを確認した。

3.2 リアクターの処理状況

手分析の COD_{Cr} は、HACH 社の試薬である TNT822 を用いて吸光光度法で測定し、測定範囲の 20~1,500 mg/L を超える試料は希釈した。ろ過を行わない全 COD は懸濁した汚泥に左右され、汚泥が浮上した沈殿槽上澄みでは約 1,460mg/L となった。このため、曝気による処理の変化は 0.22μm でろ過した溶解性 COD で評価することとした。手分析による沈殿槽上澄みの溶解性 COD は汚泥返送前の 810mg/L、返送後が 126mg/L であった。この結果は返送後の曝気槽の 130mg/L とほぼ等しく、培地 1,774mg/L に対する除去率は約 93%であった。返送前は沈殿槽に滞留した汚泥から有機物が溶出した可能性がある。なお、超純水をろ過したブランクが 98mg/L と高かったため、ろ過を見直したうえで、溶解性 COD の測定モードとしたセンサーを手分析値で校正した。

3.3 状態と行動の定義に向けた検討

実処理場では最終沈殿池の上澄みを消毒して放流するため、沈殿槽の上澄みを処理水とみなした。ただし培地は実下水より濃く測定法も異なるため、処理水質は除去率でも評価することとした。また本リアクターの HRT は約 6.6day であり、反応タンクの検討例である約 8h³⁾の約 20 倍と長く処理に余裕があり、日変動もないため、曝気の削減効果は常時曝気に対する曝気時間の削減率で示し、実処理場での削減量とは区別して扱った。

曝気の変化が処理水質に表れるまでには時間を要するため、そこで、曝気を一時的に変えたときの COD 計測値の遅れと変化幅を確認して制御の時間間隔を決め、応答が遅い場合は、曝気に速く応答する DO とその変化量を状態に加えることで評価を行うことができると考えた。

さらに、汚泥返送の前後で上澄みの溶解性 COD が大きく異なったため、返送後の状態から次の返送までを 1 回の学習期間とした。

4. まとめ

強化学習を行うための実験系を構築して COD 計を校正し、溶解性 COD により処理状況と汚泥返送の影響を把握した。今後は曝気に対する応答を確認し、DO の要否を含めて状態を定義したうえで学習を進める。

参考文献

- 1) 国土交通省国土技術政策総合研究所：B-DASH プロジェクト No.27 温室効果ガス削減を考慮した発電型汚泥焼却技術導入ガイドライン（案），国総研資料第 1093 号，2019.
- 2) 国土交通省水管理・国土保全局下水道部：下水処理場のエネルギー最適化に向けた省エネ技術導入マニュアル（案），2019.
- 3) 吉田航ほか：下水処理プロセスの曝気量制御技術，三菱電機技報，Vol.94，No.7，2020.
- 4) Chen, K. et al.: Chemosphere, Vol.279, 130498, 2021.
- 5) Hernández-del-Olmo, F. et al.: Applied Sciences, Vol.13, No.8, 4752, 2023.
- 6) Kåge, L. et al.: Frontiers in Water, Vol.7, 1537868, 2025.
- 7) Haarnoja, T. et al.: Soft actor-critic algorithms and applications, arXiv:1812.05905, 2018.