

II. 厚生労働科学研究費補助金（健康安全・危機管理対策総合研究事業）  
「水道情報の活用等による技術水準の確保及び技術継承のための研究」  
分担研究報告書

ビッグデータに基づく浄水場運転管理予測手法の検討

研究分担者 山村 寛 中央大学理工学部 教授

**研究要旨**

砂ろ過は、表面吸着によって凝集沈殿で処理されなかった微細なフロックを除去する。砂ろ過の過程でフロックが砂の表面に蓄積することで、時間の経過と共に損失水頭が増加する。損失水頭が一定の値に達すると、破過によりフロックが剥離し、濁質が処理水へ流出するため、一定時間毎にろ過層の洗浄を実施している。しかし、水道事業の広域化に伴う浄水場の統廃合により、1つの浄水場にかかる負荷が増大することが予想されるため、原水水質と凝集条件に応じた適切な洗浄時間の再検討が必要である。これまで、砂ろ過を表す物理モデルとして **Kozeny-Carman** 式が提案されたが、パラメーターの決定に時間を要することから、実施への導入に至っていない。本研究では、数式モデルで定義された急速ろ過過程の入力と出力の関係を機械学習によってモデル化することで、損失水頭を予測するろ過池閉塞シミュレーションモデルを開発した。具体的には、重回帰モデル（MLR）、ランダムフォレスト（RF）、ディープニューラルネットワーク（DNN）、長期短期記憶（LSTM）の精度を比較し、最適モデルを決定した。また、ろ過速度を変動させることで、水道事業の広域化に伴う損失水頭上昇速度のシミュレーションを実施した。

モデルの学習には新潟県長岡市水道局の妙見浄水場と気象庁で蓄積された2年分のデータ（2021年4月-2023年3月）を利用した。損失水頭予測モデルの入力値を検討するため、急速ろ過過程の数式モデルから、フィッシュボーン図を作成した。モデル性能の評価指標として、真値と予測値の誤差を測る二乗平均平方根誤差（RMSE）を使用した。特に、テストデータにおけるろ過の稼働を打ち切る直前で損失水頭がピークに達した時間のRMSEで評価することで、最適モデルを決定した。ろ過速度の増加が損失水頭に及ぼす影響をシミュレーションするため、**Kozeny-Carman** 式を用いた数値実験を行った。ろ過速度を現状の1、3、5および10割増加させたときの損失水頭の数値実験値を算出し、それぞれ最適だったモデルに学習させた。シミュレーションでは、損失水頭の上昇速度が高かった2023年1月23日14時から2023年1月27日2時までの期間における予測値を使用した。妙見浄水場におけるろ過経過時間の上限である84時間以内に、もう一つのろ過洗浄条件である損失水頭が2mに達するまでの時間を比較し、ろ過速度の増加が損失水頭に与える影響を評価した。

作成したフィッシュボーン図より、ろ過速度、水温、凝集剤注入量、ろ過池流入水濁度、ろ過継続時間が入力値として必要な因子であることが明らかになった。LSTMは時系列モデルであるため、ろ過継続時間を入力値に含めずに学習し、MLR、RF、DNNはろ過継続時間を含めて学習した。MLR、RF、DNNおよびLSTMのRMSEを比較した結果、それぞれ0.26m、0.26m、0.38m、0.15mとなり、LSTMが損失水頭のピークを最も正確に予測した。この結果より、LSTMを損失水頭の最適予測モデルとした。LSTMを用いてろ過速度を1、3、5、10割増加させたシミュレーションを行った結果、1割増加の場合は、84時間以内に損失水頭が2mに到達せず、3割増加では69時間、5割増加では57時間、10割増加では43時間で2mに達すると予測された。これらのシナリオにおけるRMSEはそれぞれ、0.14m、0.20m、0.22m、0.32mであり、ろ過速度が大きくなるにつれて予測誤差が増加する傾向が明らかになった。

4つのモデルの精度を比較し、LSTMが最も高精度であったのは、LSTMが時系列変化を学習するモデルであり、ろ過層内の経時的な変化を学習可能だったためだと考えられる。より高精度なモデルの構築には、フィッシュボーン図より、砂とフロックの球形度および流路形状情報の入力値への追加が挙げられる。ろ過速度の増加シミュレーションの結果、ろ過速度が3割以上増加すると、洗浄頻度の見直しが必要であることが示唆された。洗浄時間を3割増加では15時間、5割増加で27時間、10割増加で41時間程度、早める必要があると予測された。一方で、1割の増加では現状との差が見られなかったが、ろ過継続時間を延長するに従い、差が生じると考えられる。

本研究では、ろ過池閉塞を予測するモデルとして、LSTMが最適であると示された。また、本モデルは、浄水場で容易に観測可能なデータから損失水頭の予測ができたため、実務で応用でき

る簡略さを持つと示唆された。本モデルの活用によって、浄水場の統廃合前に洗浄時間の決定が行えることから、広域化に伴う処理水量の増加にもシミュレーションが可能となると考えられる。

## A. 研究目的

急速ろ過は、広範な水質に対応できる浄水処理プロセスであり、凝集・沈殿によるコロイド粒子の粗大化および沈殿に加えて、凝集・沈殿で除去仕切れなかった微小なフロックを砂ろ過により除去することで、極めて清澄な水を得ることができる。特に砂ろ過プロセスは、クリプトスポリジウムなどの病原性生物の除去に関して重要な役割を担っており、適切な維持管理が求められている。

砂ろ過は、砂層表面への吸着によって、砂層の間隙よりも小さな微小フロックの除去が可能となる。ろ過の継続に伴って表面に堆積した微小フロックが蓄積することで間隙が閉塞し、間隙流速の増加と損失水頭の増加を引き起こす。また、微小フロックの砂層表面への吸着が飽和する破過に到達した際には、ろ過池から微小フロックや病原性微生物が漏れ出し、処理水の安全性が著しく低下するため、破過に達する前に、通常は、砂に付着する微小フロックを剥離するためのろ層洗浄を定期的に行うことが多い。ろ層洗浄には、水流ジェットを砂層表面に吹き付ける表面洗浄と、下部から上部に向けて流水する逆流洗浄を実施する人が多い。

日本の人口は、2008年にピークに達して以降、急激な減少傾向にあり、2065年には現在の人口から35%減の約8800万人にまで減少することが予測されている。人口減少に伴って、水需要や料金収入も大幅な減少が見込まれており、持続的な経営に向けて、複数水道事業の統廃合を含む広域化の推進が不可欠とされる。広域化事業では、既存浄水場の負荷増大により、複数の浄水場を統廃合する計画も検討されている一方で、砂ろ過における限界流速の推定が難しいことが、ダウンサイジング計画の策定にあたっての課題になっている。浄水場では一定間隔毎にろ層洗浄を実施しているが、ろ過池の負荷を増大した際には、長期間かつ大規模のろ過池を使用した、洗浄頻度の最適化が必要となる。安価かつ迅速に広域化を推進するためには、ろ過池の洗浄頻度等をシミュレーション可能なモデルの構築が求められる。

これまで、様々な研究において、砂ろ過池の損失水頭を予測するモデルが構築されている。松井ら<sup>1)</sup>は、Kozeny-Carman式を利用し、損失水頭を予測するシミュレーションを実施した。ろ層内の比表面積と圧力損失の関係を以下に示す。

$$\rho g \frac{\partial H}{\partial Z} = K_K \mu \frac{s^2}{\varepsilon^3} V \quad (1)$$

ここで、 $\rho$  : 水の密度(=1.0)[g/cm<sup>3</sup>]、 $g$  : 重力加速度(=980)[cm/s<sup>2</sup>]、 $H$  : 損失水頭[cm]、 $Z$  : ろ層深さ[cm]、

$K_K$  : 流路形状に依存する係数(-)、 $\mu$  : 水の粘性係数[g/cm<sup>2</sup>・s]、 $s$  : ろ層内比表面積[cm<sup>2</sup>/cm<sup>3</sup>]、 $\varepsilon$  : ろ層の空隙率(-)、 $V$  : 空塔ろ過速度[cm/s]である。

本研究において提案された数式モデルは、従来から研究されていたろ層の状態に関するパラメーターに加えて、フロック特性を新たにパラメーターとして導入したモデルであり、カラム試験において高い精度が確認されている。一方で、シミュレーションに必要となる各種パラメーターは、繰り返し実験により推定する必要があるため、値の決定に時間を要することが課題となっている。また、流入フロックの性質が経時的に変化する場合や不均一なフロックの場合への対応について課題が残ることから、実施への導入には至っていない。

今回は、統計モデルおよび機械学習モデルにより、ろ過方程式のパラメーターを推定する手法を検討する。統計モデルの代表として重回帰モデル(Multiple linear regression : MLR)が挙げられる。MLRは、複数の説明変数と目的変数を用いて回帰分析を行い、最小二乗法を用いて回帰係数が決定される。Subinら<sup>2)</sup>は、韓国の浄水場を対象として、原水流量、濁度、pH、アルカリ度、水温および電気伝導率のデータを使用し、凝集剤注入量を予測するMLRモデルを構築した。その結果、 $R^2 = 0.63$ の精度で予測することに成功している。

機械学習モデルの代表として、ランダムフォレスト(Random forest : RF)およびディープニューラルネットワーク(deep neural networks : DNN)が挙げられる。RFはアンサンブル学習の一種であり、複数の決定木を構築する手法である。個々の決定木の性能は高くないが、複数用いることで予測精度を向上させることができる。RFの特徴として、過学習を起こしにくく、正規化の処理が不要な簡便さを持つことが挙げられる。また、比較的外れ値に影響されにくく、入力変数間の複雑な関係を記述可能である。Subinら<sup>2)</sup>は、韓国の浄水場を対象として、原水の流量、濁度、pH、アルカリ度、水温および電気伝導率のデータを使用し、RFで凝集剤注入量予測モデルを構築した結果、 $R^2 = 0.59$ の精度での予測に成功している。ニューラルネットワークは、人間の脳の構造を模したアルゴリズムである。入力層、隠れ層、出力層から成り、これらがノードにより並列につながり合った構造を持つ。DNNは、より深い層構造を有するニューラルネットワークである。モデル構築にあたっては、教師あり学習と教師なし学習が存在する。教師あり学習とは、入力とそれに対する答え(教師データ)を読み込ませる方法である。与えられた入力パターンに対応する出力を得るために、最適な重みを見つけ出し、各ニューロンに与えられ

た重みの大きさによってニューロンの信号強度を変化させることで非線形関係における予測を可能にする。Zangら<sup>3)</sup>は、中国北西部を対象とし、灌漑用水量、蒸発量、降雨量、気温から FFNN (feed-forward neural network) で地下水面の深さを出力し、 $R^2 = 0.004-0.495$ で予測することに成功している。

時系列データの学習には、RF および DNN を時系列モデルに発展させたりカレントニューラルネットワーク (Recurrent neural networks : RNN) の適用が有効となる。RNN は、 $t$  および  $t+1$  間の隠れ層をノードでそれぞれ接続することで、時系列でパラメータを更新・保持することが可能となる。ただし、長期間になるほど、パラメータ値が消失することが課題となっており、RNN では長期の予測には正確な予測が困難とされる。RNN の課題を解決し、長期の時間依存性と短期の時間依存性の双方とも学習可能となったのが長期短期記憶 (Long Short-Term Memory : LSTM) である。LSTM は、入力、出力、忘却の 3 ゲートと CEC (Constant Error Carousel) セルを持つ。CEC セルには活性化関数がないため、逆伝播の際に勾配が消失せず、学習を進めることができる。Zangら<sup>3)</sup>は、中国北西部を対象とし、灌漑用水量、蒸発量、降雨量、気温を用いて地下水面の深さを予測した。提案された LSTM モデルの精度は、 $R^2 = 0.789-0.952$ であった。

物理・化学モデルで定義された急速ろ過に関する方程式中のパラメータを、上述した統計モデルおよび機械学習モデルにより推定することが可能になれば、ろ過速度の上昇に伴う損失水頭上昇速度などを机上でシミュレーション予測することが可能になる。以上の背景から、本研究では、浄水場の運転データを用いて、ろ過方程式中のパラメータを統計モデルおよび機械学習モデルにより算出する手法を確立する。さらに、ろ過速度を変動させた際の損失水頭上昇速度を予測するためのシミュレーションモデルを構築する。

## B. 研究方法

### 損失水頭予測モデルの開発

#### 入力値の検討

急速ろ過過程の数式モデルをもとに、フィッシュボーン図を作成し、損失水頭を予測するための入力値を検討した。作成したフィッシュボーン図を以下に示す (図 1)。

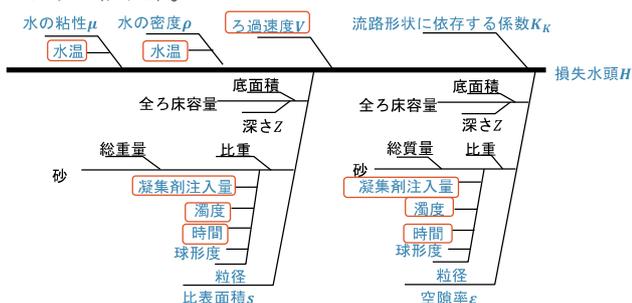


図 1. フィッシュボーン図 (損失水頭)

定数を黒字、変数を青字、モデルに含む因子を赤字で表現した。松井ら<sup>1)</sup>は、流路形状に係る係数、球形度、水の密度を定数としていたが、実現象を想定し、本研究では変数として扱った。Kozeny-Carman 式に直接含まれる因子は、重力加速度とろ層深さを除き、水の粘性、密度、ろ過速度、流路形状に依存する係数、比表面積、空隙率である。フィッシュボーン図により、これらの因子を浄水場で測定可能なデータで説明できるか確かめた。

水の粘性および密度は水温によって決定される。粘性は粒子の沈降速度に影響するため、衝突合一確率に関わると考えられる。空隙率は、砂と砂の隙間の割合を表しており、全ろ床の容量と砂の容量から求めることができる。また、比表面積は全ろ床容量あたりの表面積であり、全ろ床の容量と砂の表面積の関係から求めることができる。ろ床の容量は変化しないため、両因子においても、砂の表面にフロックが蓄積され、時間と共に変化する砂の粒径が特に重要となる。そのため、付着力と衝突合一確率に影響する凝集剤注入量、濁度、ろ過継続時間が関連因子であると考えた。

作成したフィッシュボーン図より、水温、ろ過速度、凝集剤注入量、ろ過池流入水濁度、時間が損失水頭予測モデルに必要な入力値と判断した。ただし、本研究で使用するモデルのうち LSTM は時系列モデルであることから、LSTM を使用する場合に限り、ろ過継続時間を入力値から除外した。

#### 使用したデータ

本研究では、新潟県長岡市に所在を置く妙見浄水場のデータを使用した。妙見浄水場は信濃川表流水を水源としており、施設能力は  $118,000 \text{ m}^3/\text{日}$  である。平均ろ速は  $128 \text{ m/日}$ 、処理方式は急速ろ過方式を採用しており、着水井、沈殿池、急速ろ過池から構成されている。沈殿池までは 2 ルートに分かれており、沈殿後の処理水は、10 池の急速ろ過池で処理される。全てのろ過池が重力開放形単層ろ過池であり、ろ過砂はマンガン砂 5 cm、ケイ砂 65 cm が使用されている。

データは 2021 年 4 月～2023 年 3 月までの期間の 1 時間ごとのデータを使用した。説明変数には、3 号池のろ過速度、水温、凝集剤注入量、1 系と 2 系のろ過池流入水濁度を用いた。目的変数には、3 号池の損失水頭を使用した。

ろ過速度に関しては、浄水場で観測しているろ過量から変換して求めた。以下に変換式を示す。なお、ろ過池底面積は  $102.5 \text{ m}^2$  である。

$$\text{ろ過速度} [\text{m/h}] = \frac{\text{ろ過量} [\text{m}^3/\text{h}]}{\text{ろ過池底面積} [\text{m}^2]} \quad (2)$$

また例外として、水温に関して、妙見浄水場では

1 週間に 1 回の観測でありデータ数が不足していたことから、生田ら<sup>4)</sup>の重回帰式を参考にし、気温と降雪深より水温に変換して求めた。気象庁の過去の気象データ<sup>5)</sup>から新潟県長岡市の 1 時間ごとの同期間における気温と降雪深のデータを取得した。重回帰式を以下に示す。

$$y = 3.41 + 0.35t_1 + 0.27m_2 - 0.037s_1 \quad (3)$$

ここで、 $y$ : 水温予測値、 $t_1$ : 長岡市当日気温、 $m_2$ : 長岡市気温 816 時間移動平均、 $s_1$ : 長岡市 24 時間前降雪深である。

ただし、長岡市気温 816 時間移動平均のうち、気温の欠損が含まれている場合は、その時刻の気温を除いた平均値を算出した。また、降雪深が欠損している時刻は降雪深を 0 として計算した。

得られたデータの特徴を把握するために記述統計を行い、経時変化グラフと相関を算出した。前処理前のデータの概要を以下に示す (表 1)。

表 1. 浄水場データの概要(前処理前)

	度数	平均	標準偏差	最小値	中央値	最大値
損失水頭[m]	17520	0.50	0.48	0.00	0.39	3.65
ろ過速度[m/h]	17520	2.99	1.78	0.00	3.83	5.28
水温[°C]	17444	12.11	5.61	1.98	12.07	23.72
凝集剤注入量[mg/L]	17520	65.18	23.08	0.00	60.00	290.00
流入水滴度1系[mg/L]	17520	0.24	0.09	0.03	0.23	1.59
流入水滴度2系[mg/L]	17520	0.17	0.09	0.02	0.15	0.99
ろ過継続時間[h]	17520	28.66	27.08	0.00	23.00	116.00

## データの前処理

データの前処理は、洗浄時間の補正、線形補間、移動平均、欠損値除去、正規化、データの分割の順で行った。ろ過池は、定期的に洗浄が実施されているため、洗浄と稼働の時間が交互に繰り返されている。そのため、損失水頭、ろ過速度のいずれかが 0 のときを、ろ過池の洗浄時間とした。ろ過池洗浄中は、他の因子が作用していないとみなし、その時刻における全ての入力値、出力値を 0 とした。

欠損がある部分を補うために、欠損値前後の数値を線形に繋ぐ線形補間を行った。また、ノイズを減らし、長期のトレンドを追いやすくするために、本研究では移動平均による平滑化を行った。移動平均は 7 時間分のデータを使用し、計算結果は 7 時間のうちの最も後ろに入力された。移動平均の式を以下に示す。

$$S_n = \frac{a_{n-6} + a_{n-5} + \dots + a_n}{N} \quad (4)$$

ここで、 $S_n$ : 移動平均値、 $a_n$ : 区間  $N$  のデータ、 $N$ : 区間のデータ数である。

1 つでもデータに欠損がある時刻は、その時刻すべてのデータを削除した。欠損値の除去後は値のスケールの違いによる説明変数間での重みの不公平をなくすために、正規化を行なった。これにより、0 から 1 の範囲にスケールを揃えることができる。本

研究で使用した正規化の式は以下の通りである。

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (5)$$

ここで、 $x'$ : 正規化されたデータ、 $x$ : 元のデータ、 $min$ : 最小値、 $max$ : 最大値である。

前処理の結果、データセットは 17,444 に減少した。最後に、モデルを学習させるための訓練データ 80% と、モデルの汎用性を評価するためのテストデータ 20% に前後で分割した。

## モデルの実装

MLR は、ライブラリとして scikit-learn を使用した。また、学習は 1 回のみ行った。RF も同様に scikit-learn を使用し構築した。決定木の数を 300 本、木の最大の深さを 5 とした。学習は 5 回行った。

DNN では Keras を用いた。アーキテクチャは、中間層数 10 層、ノード数 256 個、ドロップアウトは 0.3 を 11 層、活性化関数を ReLU とし、さらにバッチ正規化を行う層を 11 層追加した。また、学習条件はエポック数 100 回、学習率はエポック数に応じて変動させた ( $1 \leq \text{epochs} < 25 : 0.001$ ,  $25 \leq \text{epochs} < 50 : 0.0005$ ,  $50 \leq \text{epochs} < 75 : 0.0001$ ,  $75 \leq \text{epochs} < 100 : 0.00005$ )。最適化関数は Adam、損失関数は平均二乗誤差 (MSE) を使用した。モデルは 5 回学習させ、過学習を防ぐための Early Stopping を導入したため、エポック数は 1 回目のみ 100 回であり、2 回目、3 回目、4 回目では 79 回、5 回目は 91 回で学習が停止した。

LSTM では Keras を用いて構築を行った。アーキテクチャは、活性化関数 tanh と sigmoid を使用し、LSTM ブロック数 32 個とした。また、学習条件はニューラルネットワークと同様で、エポック数 100 回、学習率はエポック数が増えるごとに学習率を下げ ( $1 \leq \text{epochs} < 25 : 0.001$ ,  $25 \leq \text{epochs} < 50 : 0.0005$ ,  $50 \leq \text{epochs} < 75 : 0.0001$ ,  $75 \leq \text{epochs} < 100 : 0.00005$ )、最適化関数は Adam、損失関数は平均二乗誤差 (MSE) を使用した。また、LSTM 入力の高さを 72 時間に設定した。LSTM においても学習は 5 回行った。

## モデル性能の評価

各モデルの性能を評価する指標として、決定係数 ( $R^2$ ) と二乗平均平方根誤差 (RMSE) を利用した。 $R^2$  は、1 に近いほど高精度とみなした。一方、RMSE は、値が 0 に近いほど予測が適合していると判断した。モデル全体の性能は、 $R^2$  と RMSE から評価し、LSTM のテストデータの数に合わせて行われた。

住谷ら<sup>6)</sup>を参考に、テストデータにおけるろ過の稼働を打ち切る直前で損失水頭がピークに達した時の RMSE で評価し、最適モデルを決定した。このときの誤差は MRMSE と呼ぶ。MLR を除くモデルは 5 回の学習を行ったが、最も高い精度で比較した。

$R^2$ 、RMSE の式は以下の通りである。

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (6)$$

$$RMSE(y, \hat{y}) = \sqrt{\frac{1}{n} \sum_{i=0}^{n-1} (y_i - \hat{y}_i)^2} \quad (7)$$

ここで、 $n$  : サンプル数、 $i$  : サンプル番号、 $y_i$  : 正解値、 $\hat{y}_i$  : 予測値、 $\bar{y}$  : 正解値の平均である。

### ろ過速度を変動させたシミュレーション 数値実験

浄水場から取得したろ過速度は 0 m から 5.28 m の範囲内である (表 1)。そのため、前節までで本研究で作成したモデルは、この範囲内のデータを学習しているため、学習範囲を超えろ過速度でのシミュレーションが難しい。しかしながら、本研究で作成したモデルは広域化によりろ過速度が増加した時の洗浄頻度をシミュレーションすることを想定して作成している。そのため、以下の Kozeny-Carman 式を用いた数値実験を行った。ろ過速度を現状の 1、3、5 および 10 割増加させたときの損失水頭の数値実験値を求めた。

$$\frac{H}{Z} = \frac{k \cdot \mu \cdot s^2 \cdot v}{\rho \cdot g \cdot \varepsilon^3} \quad (8)$$

ここで、 $\rho$ :水の密度[kg/m<sup>3</sup>]、 $g$ :重力加速度[m/s<sup>2</sup>]、 $H$ :損失水頭[m]、 $Z$ :ろ層深[m]、 $k$ :コゼニー係数、 $\mu$ :水の粘性係数[kg/m・s]、 $s$ :ろ層内比表面積[m<sup>2</sup>/m<sup>3</sup>]、 $\varepsilon$ :ろ層の空隙率、 $v$ :ろ過速度[m/s]である。

(8)式に、重力加速度 9.8 m/s<sup>2</sup>と、浄水場の実際の値である損失水頭、ろ過速度、深さ 0.7 [m]、粘性、密度を代入し、未知のパラメータ $\frac{k \cdot \mu \cdot s^2}{\varepsilon^3}$ を求めた。粘性及び、密度は、以下の(9)から(16)式を用いて水温から換算することで求めた。

・水温から密度への換算式

$$\rho = \frac{999.83952 + 16.945176(T) - 7.9870401 \times 10^{-3}(T)^2 - 46.170461 \times 10^{-6}(T)^3 + 105.56302 \times 10^{-9}(T)^4 - 280.54253 \times 10^{-12}(T)^5}{1 + 16.879850 \times 10^{-3}(T)} \quad (9)$$

水の粘性の換算式は水温によって異なる。

・水温が 0 °Cより高く 20 °Cより低い場合の式

$$\mu = 10^{-3}(10^A) \quad (10)$$

$$A = \frac{1301}{998.333 + 8.1855(T - 20) + 0.00585(T - 20)^2} - 1.30223 \quad (11)$$

・水温が 20 °Cより高く 100 °Cより低い場合の式

$$\mu = (1.002 \times 10^{-3})(10^B) \quad (12)$$

$$B = \frac{1.3272(20 - T) - 0.001053(T - 20)^2}{T + 105} \quad (13)$$

ここで、 $\rho$  : 密度[kg/m<sup>3</sup>]、 $T$  : 水温[°C]、 $\mu$  : 粘性[kg/m/s]である。未知のパラメータ $\frac{k \cdot \mu \cdot s^2}{\varepsilon^3}$ の算出後、再び式に $\frac{k \cdot \mu \cdot s^2}{\varepsilon^3}$ 、重力加速度、深さ、密度、粘性、増速させたる過速度を代入し、損失水頭を算出した。

### モデルの実装

シミュレーションは、LSTM を用いて行った。前処理、モデルの構造、学習条件は前節と同様の設定で行った。モデルの構築は、ろ過速度毎に行ったため、ろ過速度と損失水頭が異なる 4 つのシミュレーションモデルが構築された。また、各モデルは 5 回学習をさせ、最も高精度の結果を用いてシミュレーションを行った。入力値として、水温、凝集剤注入量、ろ過池流入水濁度は現在の浄水場のデータを用い、ろ過速度は任意の値 (1 割増、3 割増、5 割増、10 割増) を各モデルに入力した。出力値は、Kozeny-Carman 式から求めた損失水頭の数値実験値を使用した。

洗浄時間の提案には、LSTM より算出された値を使用した。洗浄頻度の提案を行うため、損失水頭上昇速度が大きい 2023 年 1 月 23 日 14 時から 2023 年 1 月 27 日 2 時までの期間におけるテストデータの予測値を使用した。これは、処理水への濁質混入や細菌の繁殖を防ぎ、安全な水道水を供給するためである。妙見浄水場におけるろ過経過時間の上限である 84 時間以内に、損失水頭が 2 m に達するまでの時間を比較し、ろ過速度の増加が損失水頭に与える影響を評価した。

## C. 研究結果

### 3.1 損失水頭予測モデルの開発

#### 3.1.1 データの分析

損失水頭とろ過継続時間の相関関係は 0.86 と高かった (表 2)。また、損失水頭とろ過速度の相関が 0.58 であることから、相関があることが示唆された (表 2)。

表 2. 変数の相関

	損失水頭	ろ過速度	水温	凝集剤注入量	流入水濁度1系	流入水濁度2系	ろ過継続時間
損失水頭	1.00						
ろ過速度	0.58	1.00					
水温	-0.25	-0.02	1.00				
凝集剤注入量	-0.12	0.10	0.28	1.00			
流入水濁度1系	-0.01	0.00	0.13	0.11	1.00		
流入水濁度2系	-0.06	0.00	0.34	0.21	0.54	1.00	
ろ過継続時間	0.86	0.50	-0.07	-0.04	0.03	0.03	1.00

経時変化のグラフから、損失水頭の高さは期間によって差があることが確認された (図 2)。ろ過速度は一定の値の範囲内で推移しており、大きな変化は見られなかった (図 3)。水温は、先行研究の式より変換されたが、0 °Cから 25 °Cの範囲で推移しており、

実際の水温と同程度であると言える（図 4）。凝集剤注入量は水温が高い時期に多く投入されていることがわかった（図 5）。ろ過池流入水濁度において、1系は 2 系と比較し、やや高濁度であることが示された（図 6, 7）。ろ過継続時間から、妙見浄水場ではおおよそ同程度の時間でろ過の稼働を打ち切り、ろ過層内の洗浄を行っていることが明らかとなった。

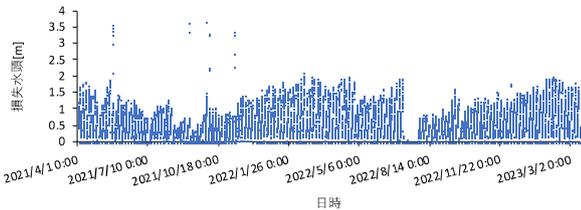


図 2 損失水頭の経時変化

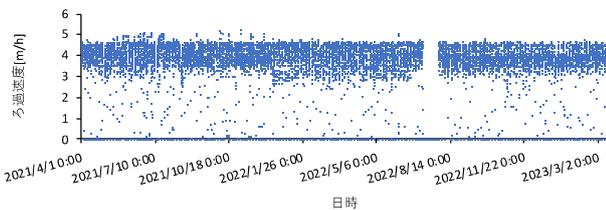


図 3 ろ過速度の経時変化

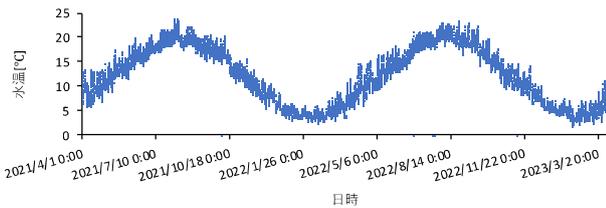


図 4 水温の経時変化

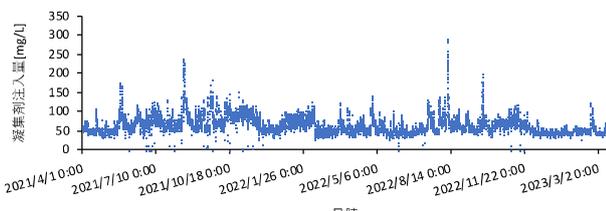


図 5 凝集剤注入量の経時変化

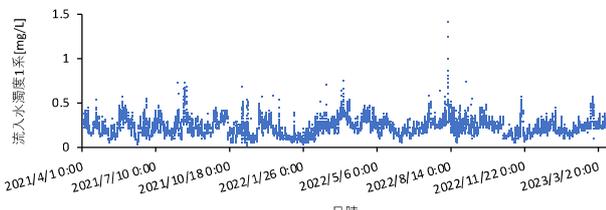


図 6 ろ過池流入水濁度 1 系の経時変化

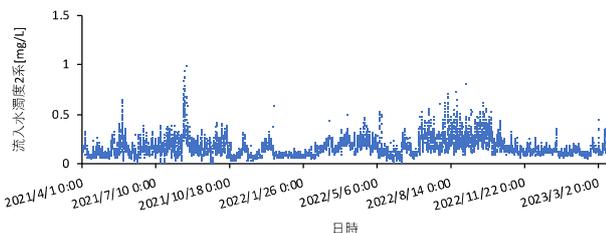


図 7 ろ過池流入水濁度 2 系の経時変化

### モデルの検討

重回帰モデルから算出された重回帰式は以下の通りである。

$$\begin{aligned}
 y_{\text{損失水頭}} = & 0.01 + 0.689161 * x_{\text{ろ過速度}} - 0.49769 * x_{\text{水温}} - \\
 & 0.538196 * x_{\text{凝集剤注入量}} \\
 & - 0.151343 * x_{\text{流入水濁度 1 系}} + 0.007803 * x_{\text{流入水濁度 2 系}} \\
 & + 1.494537 * x_{\text{ろ過継続時間}} \quad (14)
 \end{aligned}$$

偏回帰係数より、ろ過継続時間次いでろ過速度の重みが大いことがわかった。また、 $R^2$ が 0.92、RMSE が 0.14 m であり、RF と同程度の精度で予測された（表 3）。しかし、損失水頭がピークに達した時点では、正解値より低く予測していることがわかった（図 16, 20）。全体を通して、同じ高さに予測のピークがあることから、おおまかな傾向を捉えることはできたものの、細かな変動には対応できていない。

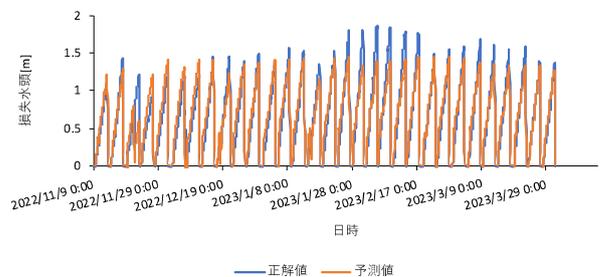


図 8. MLR で予測した損失水頭の経時変化

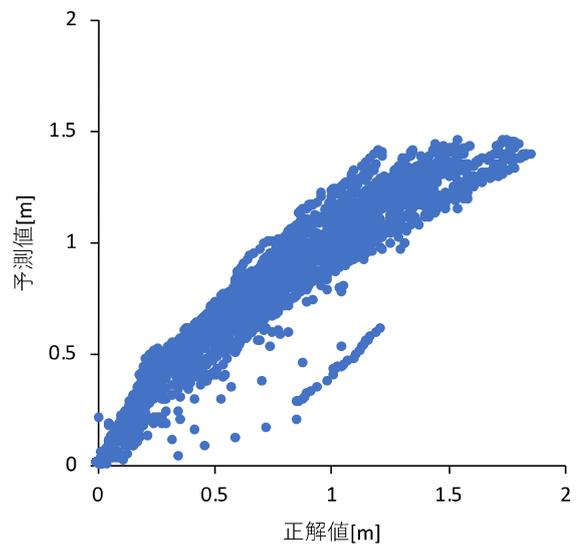


図 9. MLR 損失水頭散布図

RF では、 $R^2$ が 0.93、RMSE が 0.13 m であった（表 3）。しかしながら、RF が予測した値は、損失水頭が高い時点で正解値と大きく離れており、重回帰モデルと似た挙動を示した（図 10, 11）。

DNN は、モデル精度において MLR、RF、LSTM の 3 モデルと比較し、劣る結果となった（表 3）。損失水頭が高い時間と低い時間において特に予測の誤差が生じており、中央値前後をとる挙動を示した（図

12, 13)。

LSTM は、全体を通して $R^2$ が最も高く、RMSE が最も低いことが示唆された (表 3)。また、損失水頭の細かい挙動を捉えられていることがわかった (図 14, 15)。特に損失水頭が最大時の予測は、他のモデルと比較し、正解値に近い挙動を示した。

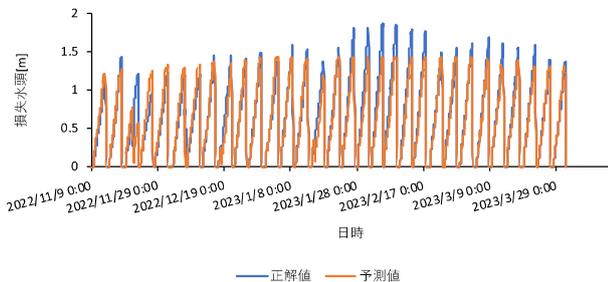


図 10. RF で予測した損失水頭の経時変化

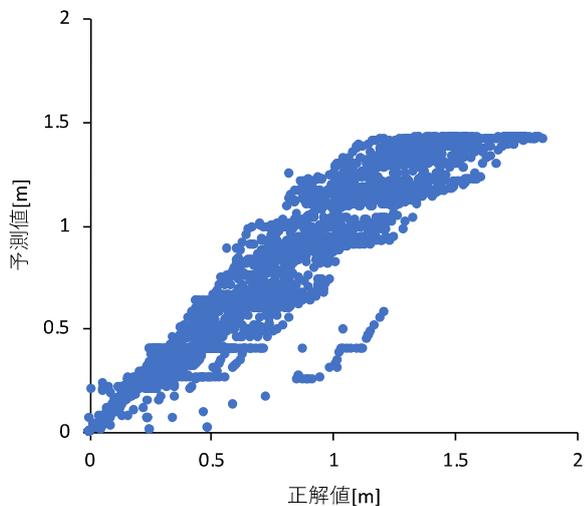


図 11. RF 損失水頭散布図

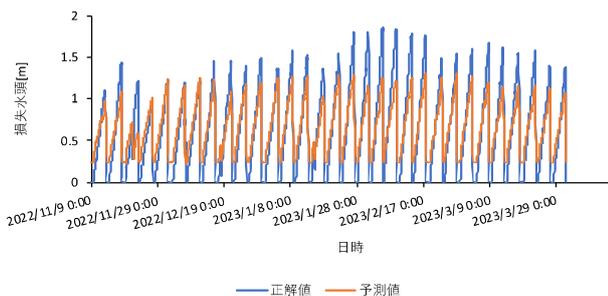


図 12. DNN で予測した損失水頭の経時変化

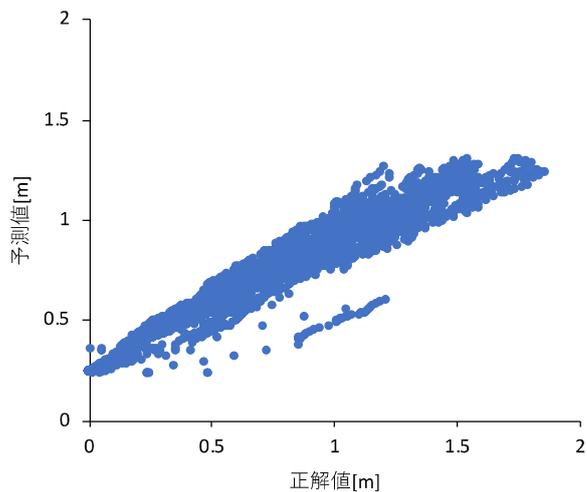


図 13. DNN 損失水頭散布図

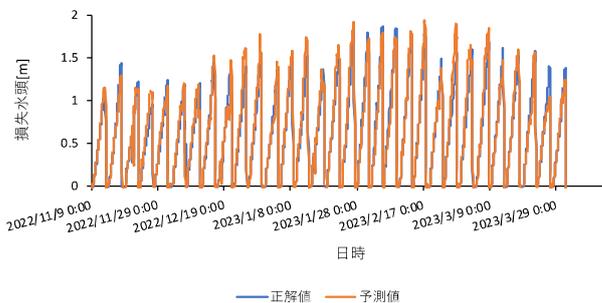


図 14. LSTM で予測した損失水頭の経時変化

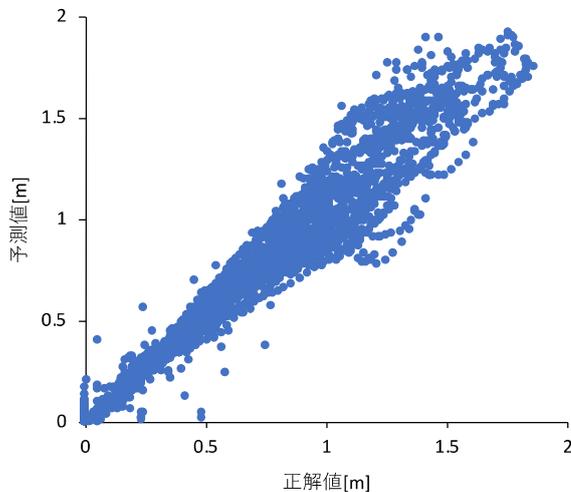


図 15. LSTM 損失水頭散布図

表 3. モデル精度の比較

	MLR	RF	DNN	LSTM
R2	0.92	0.93	0.82	0.95
RMSE	0.14	0.13	0.21	0.11
MRMSE	0.26	0.26	0.38	0.15

### ろ過速度を変動させたシミュレーション

損失水頭上昇速度は必ずしも一定ではなく、時期によって差があった (図 16~18)。ここでは、洗浄頻度の提案を行うため、損失水頭上昇速度が大きい期間に合わせて洗浄時間の提案を行った。

2023年1月23日14時～2023年1月27日2時の期間を抽出した結果を以下に示す(図19)。ろ過速度を3割、5割、10割増加させた場合で、妙見浄水場のろ過経過時上限の84時間以内に、損失水頭上限の2mに到達した。また、1割の増加では、現状との差が見られなかった。

物理式より算出した損失水頭と、モデルから予測された損失水頭の結果から算出したR<sup>2</sup>、RMSE、MRMSEをまとめたものを示す(表4)。結果より、ろ過速度を増加させるにつれ、損失水頭がピークに達した時点での予測に誤差が生じていることがわかった。

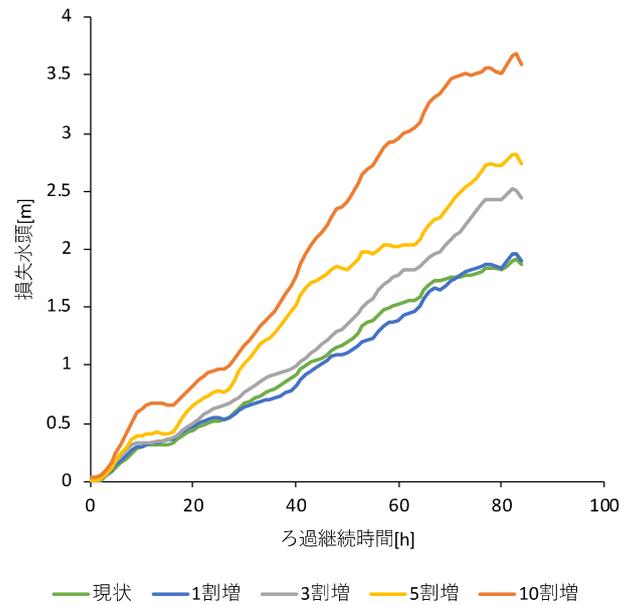


図19. 損失水頭の経時変化予測(2023年1月23日14時～2023年1月27日2時)

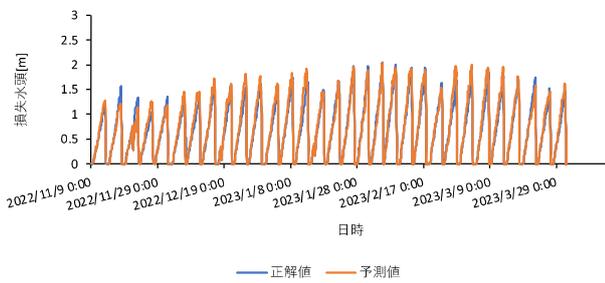


図16. ろ過速度を1割増にした場合の損失水頭の経時変化

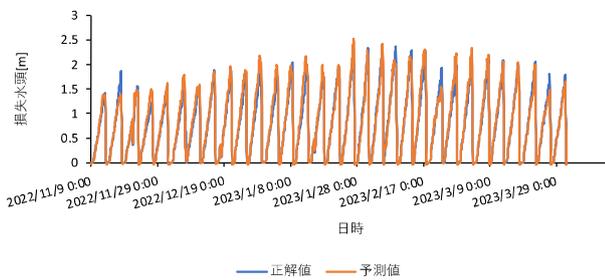


図17. ろ過速度を3割増にした場合の損失水頭の経時変化

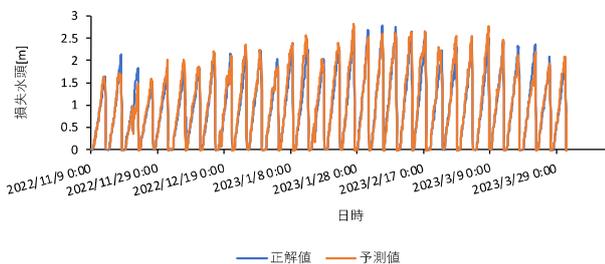


図18. ろ過速度を5割増にした場合の損失水頭の経時変化

表4. シミュレーションモデルの精度

	1割増	3割増	5割増	10割増
R2	0.95	0.96	0.95	0.94
RMSE	0.12	0.13	0.17	0.24
MRMSE	0.14	0.2	0.22	0.32

## D. 考察

### 4.1 損失水頭予測モデルの開発

損失水頭予測モデルの開発として、MLR、RF、DNN、LSTMの4モデルを開発し、最適モデルを検討した。結果より、MLRではRFと同程度の精度で予測されたことがわかった(表3)。これは、ろ過継続時間は損失水頭と正の相関があることに起因していると考えられる(表2)。損失水頭とろ過継続時間に線形関係があることから、統計的手法である重回帰モデルでも高い予測精度を出すことができたと考えられる。

また、DNNは4モデルのうち最も精度が悪く、DNNから算出された予測値は中央変動するような挙動を示した。このような挙動を取る原因として、誤差を小さくするために、モデルが中央値を予測していると考えられる。また、MLRよりも精度が劣った原因として、学習条件やモデル構造の最適化が不十分であったことが考えられる。ノード数や隠れ層等が最適化されることで、RFと同等の予測精度まで向上する可能性が考えられる。

LSTMが最も高精度で損失水頭のピークを記述した。ろ層内の状態は時間と共に変化するため、過去の時系列変化を学習するLSTMが適していると示唆された。したがって、損失水頭予測モデルとして最も適しているモデルは、LSTMであることが示され

た。

### ろ過速度を変動させたシミュレーション

ろ過速度の異なるモデルから損失水頭をシミュレーションした結果、3割、5割、10割の増加で84時間以内に損失水頭上限の2mに達することが示された(図19)。ここから、現状のろ過打ち切り時間を84時間とすると、3割の増加では15時間、5割の増加で27時間、10割の増加で41時間程度、洗浄時間を早める必要があると予測された。また、ろ過継続時間が84時間の時点で、1割の増加で現状との差が見られなかったが、ろ過継続時間を延長することで、差が生じる可能性が考えられる。

また、シミュレーションモデルの精度から、ろ過速度が大きくなるにつれて予測誤差が増加する傾向が明らかになった。その原因として、損失水頭上昇速度の増加が影響を及ぼしていると考えられる。また、ろ過速度の値が大きいため、モデルが学習する際の重みがろ過速度に偏った可能性も考えられる。

### E. 結論

本研究では、損失水頭予測モデルの開発および、ろ過速度を変化させたシミュレーションを行った。

その結果、ろ過池閉塞予測モデルとして、LSTMが最適モデルであることが示された。また、損失水頭を予測するにあたって、モデルに入力したろ過速度、水温、凝集剤注入量、ろ過池流入水濁度の妥当性が示唆された。これらは、浄水場で容易に測定可能なデータであるため、実務へ応用できる簡略さを持つと考えられる。本モデルの活用によって、浄水場の統廃合が行われる前にシミュレーションが可能となったことから、処理水量の変化に迅速に対応可能であると考えられる。

本モデルの課題として、損失水頭に関わる因子である球形度や流路形状を表現できていないことが挙げられる。より高精度なモデルの構築には、球形度および流路形状情報を入力値へ追加する必要があると考える。また、本研究では1箇所の浄水場データでの検証であった。本モデルの汎用性を確かめるため、複数の浄水場のデータを用いて学習する必要がある。また、施設規模が異なる浄水場のデータを用いることで、実データを用いたシミュレーションも可能になると考える。

- 1) 松井佳彦(1992). 急速濾過仮定の数式モデル, 北海道大学大学院工学研究科博士論文.
- 2) S. Song, L. Rong, K. Dong, X. Liu, P. L. Clech. (2020). Particle-scale modeling of fluid velocity distribution near the particles surface in sand filtration. *Water Research*, Vol.117, No.15, p.115758  
<https://doi.org/10.1016/j.waters.2020.115758>
- 3) J. Zhang, Y. Zhu, X. Z, M. Y, J. Y. (2018).

Developing a Long Short-Term Memory(LSTM) based model for predicting water table depth in agriculture areas. *Journal of Hydrology*, Vol.561, pp.918-929  
<https://doi.org/10.1016/j.jhydrol.2018.04.065>

- 4) 生田理弘, 大熊 孝, 大川秀雄, 神立秀明, 中村一郎 (1991). 河川水温と気温等の相関に関する研究 —信濃川水系魚野川を例に—. *水文・水資源学会誌*, Vol.4, No.1,
- 5) 気象庁. “過去の気象データ検索”国土交通省  
<https://www.data.jma.go.jp/stats/etrn/index.php> (参照: 2023-11-30)
- 6) 住谷渚, 安田誠宏, 森信人, 志村智也 (2021). 長短期記憶ニューラルネットワーク LSTM を用いた波高の時系列予測に関する研究. *土木学会論文集 B2(海岸工学)*, Vol.77, No.2, p.I\_151-I\_156

### F. 研究発表

#### 1. 論文発表

なし

#### 2. 学会発表

(発表誌名巻号・頁・発行年等も記入)

なし

### H. 知的財産権の出願・登録状況

#### 1. 特許取得

なし

#### 2. 実用新案登録

なし

#### 3. その他

なし