

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）
「人口減少社会における情報技術を活用した水質確保を含む管路網管理向上策に関する研究」
分担研究報告書

送配水管における水質等の変化の予測及び実証

研究分担者 氏名：荒井康裕 所属：首都大学東京

研究要旨

小規模水道事業体における効率的な管網管理手法の提案を目的に、個人宅に設置した自動水質測定器のモニタリングデータを用い、管網末端での残留塩素濃度を推定するモデルを構築した。具体的には、機械学習のひとつであるニューラルネットワーク（NN）を用いて残留塩素濃度減少を推定し、NNモデルの有用性に関する評価を試みた。平成29年度に作成した重回帰モデルと比較して最大絶対誤差が改善され、実測値の局所的な微細な変動を再現できた。

A. 研究目的

管網末端で必要な残留塩素濃度を維持するためには、送配水中の塩素の消費量を考慮し、浄水場における塩素の注入量を適切に管理する必要がある。本研究では、平成29年度において、K浄水場残留塩素濃度と個人宅残留塩素濃度の差を「残留塩素濃度消費幅」と定義し、残留塩素濃度の変動に影響を与えると考えられる水温や流量のデータを説明変数、残留塩素濃度消費幅を目的変数とした重回帰モデルを推定した。平成30年度は機械学習のひとつであるニューラルネットワーク（NN）を用いて残留塩素濃度減少を推定し、NNモデルの有用性に関する評価を試みる。

B. 研究方法

分析対象となる地域は図1に示す送配水ネットワークである。分析対象となるデータは、K浄水場計測データ（送水流量・浄水濁度・残留塩素濃度・pH）、各配水流量データ（S系第一配水流量・M系-d配水流量・M系-o配水流量）、個人宅計測データ（濁度・残留塩素濃度・pH・水温・色度・電気伝導率・水圧）の14種類である。なお、各データは

平成28年4月1日から平成29年3月31日までの時間単位データである。

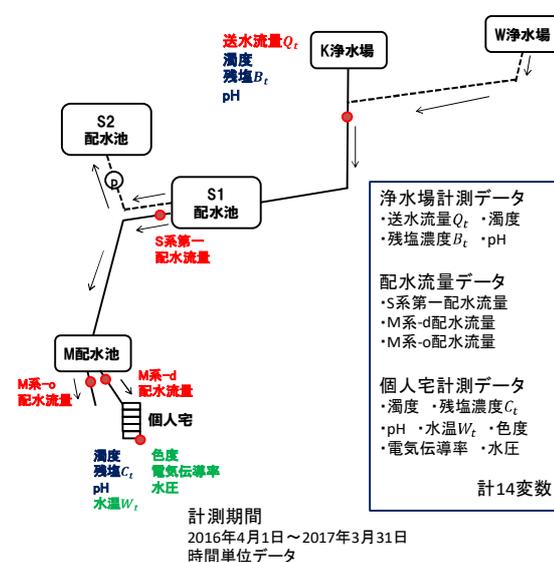


図1 対象とする送配水ネットワーク

以上の時系列データに対して、相関ヒートマップと散布図を作成した。これらを用いて変数間の関係を整理した上で、モデルの説明変数および目的変数を設定する。さらに、得られたモデルの精度を向上させるため、学習

方法を適宜変更することで本研究の最適な NN モデルを決定した。

C. 相関ヒートマップ・散布図による要因関連分析

相関ヒートマップは、各変数間の相関係数を段階ごとに色分けして表示させたものである(図 2 参照)。また、散布図は対応するデータをグラフ上にプロットすることで、相関係数だけでは判断できない変数間の関係を得ることが期待できる。

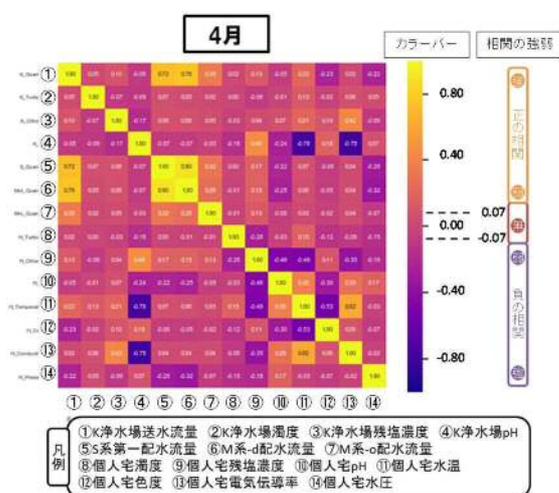


図 2 相関ヒートマップの一例 (2016年4月)

前述の 14 種類の各変数について、4 月から 10 月までの範囲で 1 ヶ月単位での相関ヒートマップと散布図を作成し、各変数間の関係を分析したところ、次の結果が得られた。まず、浄水場送水流量と他の配水流量には強い正の相関がみられた。したがって、モデルの説明変数には流量に関する変数のいずれか一つを代表して採用する。また、浄水場で計測された変数間の相関関係は無相関であるものが多いが、個人宅で計測されたデータ間の相関関係は水圧を除いて月ごとに变化するものが多くみられた。これは浄水場では残留塩素濃度や濁度の値が一定にコントロールされていることを意味し、モデルの説明変数には浄水場にて管理可能なデータを優先して選択する。

さらに、浄水場残留塩素濃度と個人宅残留塩素濃度間には、月によって相関係数の正負が逆転する現象がみられ、両者の間には複雑な関係が存在することが明らかとなった。

D. 残留塩素濃度減少モデルの構築と評価

(1) NN モデルの概要

本研究では、全結合ニューラルネットワークを用いて残留塩素濃度減少モデルを構築した。ニューラルネットワークは脳の中に存在する神経細胞(ニューロン)のつながりをコンピュータで再現したものであり、入力層から中間層(隠れ層)を経て、出力層へと信号を伝えることで、入力データ-出力データ間の関係を確立させるものである。図 3 に本研究で用いるニューラルネットワークを示す。

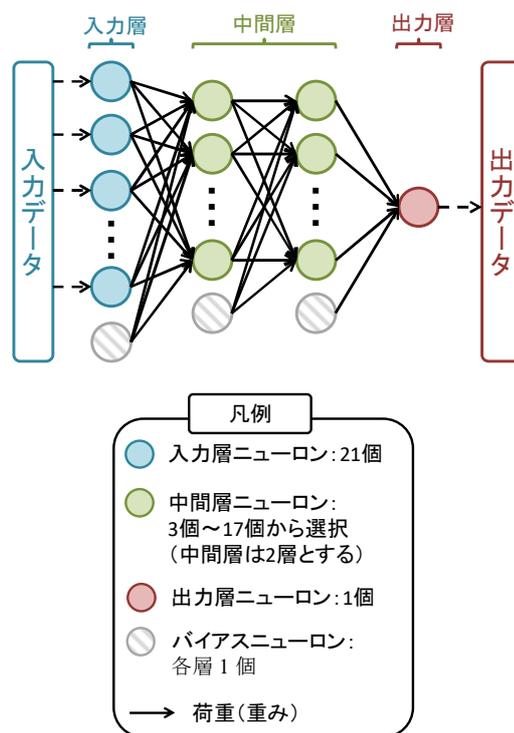


図 3 本研究のニューラルネットワーク

(2) 訓練データとテストデータ (training と test)

モデルの訓練 (training) には 6 月のデータを、テスト (test) には 7 月のデータを各々

用いることとし、説明変数・目的変数は昨年度の研究成果を踏まえて決定した。すなわち、説明変数は 11 時間までの時間遅れを考慮した浄水場残留塩素濃度 $B_t, B_{t-1}, B_{t-2}, \dots, B_{t-11}$ 、2 時間までの時間遅れを考慮した個人宅水温 W_t, W_{t-1}, W_{t-2} 、5 時間までの時間遅れを考慮した浄水場流量 $Q_t, Q_{t-1}, Q_{t-2}, \dots, Q_{t-5}$ の計 21 変数とし、目的変数は残留塩素濃度消費幅 D_t ($D_t = B_{t-11} - C_t$: ただし C_t は個人宅残留塩素濃度) の 1 変数となる。エポック数 (学習回数) の上限を 2000 とし、中間層ユニット数及びバッチサイズについて比較検討したところ、中間層は 12 ユニットでバッチサイズは 72 以下の場合が最適なモデルの候補として選択された (図 4 参照)。さらに、説明変数のうち浄水場残留塩素濃度について、11 時間までの遡行平均を計算したものに置き換えたところ、推定値グラフの時系列としての適合性が向上した。

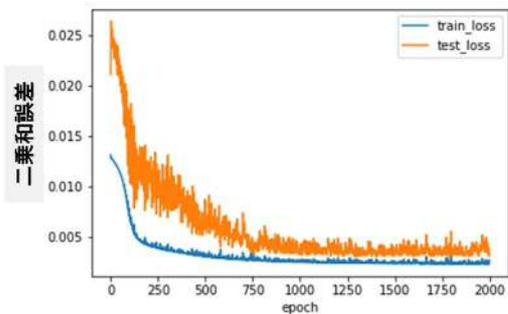


図 4 エポック数と二乗和誤差の関係 (中間層 12 ユニットの case)

したがって、最終的な NN モデルの構造は「中間層=2 層 12 ユニット・バッチサイズ

=24・エポック数=1840」とし、入力データに関しては浄水場残留塩素濃度のみ遡行平均をとった教師データを用いて学習したモデルを選択した。

(3) モデルの精度

得られたモデルの最大絶対誤差は 0.144 [mg/L] 及び 0.157 [mg/L] (訓練期間及びテスト期間) となり、既往の研究成果 (重回帰モデル) の最大絶対誤差である 0.192 [mg/L] 及び 0.158 [mg/L] (訓練期間及びテスト期間) と比較して、訓練期間の最大誤差を改善することができた。また、7 月前半における推定値グラフを示した図 5 より、NN モデルの推定値は実測値の局所的な微細な変動を再現することが可能であり、モデルの汎化能力が確認された。

E. 結論

本研究では、ニューラルネットワークを用いて残留塩素濃度減少を推定するモデルを作成した。その結果、重回帰分析を用いたモデルと比較して最大絶対誤差が改善され、実測値の局所的な微細な変動を再現できた。したがって、ニューラルネットワークを残留塩素濃度減少推定モデルに適用することは十分に可能であると言える。今後は RNN (ニューラルネットワークの出力を別のネットワークの入力として利用するような再帰的構造を持ったニューラルネットワーク) や CNN (深層学習の一種で「Convolutional」(畳み込み) という操作を加えたものであり、人間の脳が視覚情報を処理する際の動きを簡易的に再現

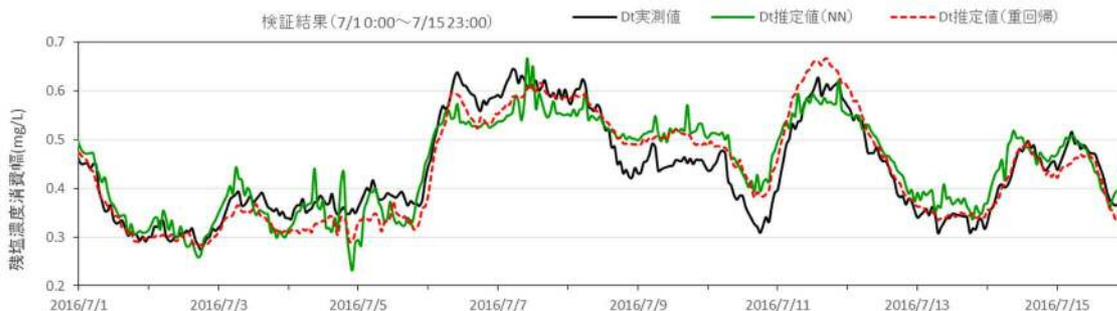


図 5 残留塩素濃度消費幅の推定グラフ (7 月前半)

し、画像認識に特化したニューラルネットワーク)等の適用可能性、原水濁度等の情報を収集し、より多くのビッグデータを活用したモデルの構築が課題である。

F. 研究発表

1. 論文発表

該当なし

2. 学会発表

荒井康裕・稲員とよの・堀口 幸菜・小泉明・佐々木史朗：配水管網の水質監視データを活用した残留塩素濃度シミュレーション、土木学会第73回年次学術講演会、pp.265-266、2018年9月

G. 知的財産権の出願・登録状況

1. 特許取得

該当なし

2. 実用新案登録

該当なし

3. その他

該当なし