

〈研究発表〉

原水水質のリアルタイム推定

松本 隼¹⁾, 山原 裕之¹⁾, 横川 勝也¹⁾, 毛受 卓¹⁾
 黒川 太¹⁾, 横山 雄²⁾

¹⁾ 東芝インフラシステムズ(株) インフラシステム技術開発センター

(〒183-8511 東京都府中市東芝町1番地 E-mail: jun8.matsumoto@toshiba.co.jp, hiroyuki.yamahara@toshiba.co.jp, katsuya.yokokawa@toshiba.co.jp, takashi.menju@toshiba.co.jp, futoshi.kurokawa@toshiba.co.jp)

²⁾ 東芝インフラシステムズ(株) 社会システム事業部

(〒212-8585 神奈川県川崎市幸区堀川町72番地34 E-mail: suguru.yokoyama@toshiba.co.jp)

概要

浄水場における塩素注入システムの自動化・最適化に向け、浄水プロセスにおける残留塩素の消費・分解を予測するためのモデル構築を進めている。塩素注入率に対する残留塩素濃度を化学反応に基づき予測する場合、化学反応式への入力情報として原水中の各反応物質の濃度が必要となるが、それらの中には浄水場のオンライン測定項目には含まれず手分析でしか取得できないデータがある。筆者らは不足する手分析データ情報を原水 pH や水温など他の浄水場内のオンライン測定項目からリアルタイム推定する手法を提案した。本稿では、既報における課題であった、手分析データ推定モデルのパラメータ最適化手法の提案および本推定値の実サンプルデータとの比較による妥当性検討について検証結果を報告する。

キーワード：残留塩素, 水質推定, AI

原稿受付 2020.7.3

EICA: 25(2・3) 33-36

1. はじめに

浄水場における塩素注入運転の自動化・最適化に向けた要素技術として、浄水プロセスにおける残留塩素の消費・分解を予測するためのモデル構築を進めている。筆者らは、各処理の化学反応式など明示的に定式化したホワイトボックスモデルと実サンプルデータセットを用いた機械学習に基づくブラックボックスモデルのハイブリッド化による残留塩素濃度予測手法を提案、評価した [1]。化学反応式への入力情報として用いる水質項目のうち、オンラインで計測されていない採水分析による水質データ（以下、手分析データと呼ぶ）をリアルタイム推定することによりハイブリッド化を実現し、浄水残留塩素濃度実測値平均 0.440 mg/L に対し絶対誤差平均 0.048 mg/L、正規化値 10.9% となることを確認した。ただし、手分析データ推定モデルのパラメータ最適化により精度改善の余地がある点、また手分析データ推定値の実サンプルとの比較検証を未実施である点が既報における課題であった。本報告では、手分析データ推定モデルのパラメータ最適化手法を提案し、推定結果の実サンプルデータとの比較検証を実施する。検証においては手分析データのトレンド比較による定性的評価および絶対誤差による定量的評価、またモデル全体への影響としてろ過水残留塩素濃度の予測精度を評価する。

2. 手分析データ推定モデル

Fig. 1 に手分析データ推定のモデリング概要図を示す（詳細なモデリング手法については [1] を参照）。浄水場内に設置されたセンサや外部の気象観測機関から取得されるオンラインデータ、年 1 回以上の実施が想定される手作業によるオフライン測定で得られる手分析データを元に、未測定期間にあたる手分析データをリアルタイム推定する。化学反応式に基づく塩素消費モデル（ホワイトボックスモデル）は非線形であるが、筆者らが提案する数理計画問題として定式化した手法を用いる際は、モデルを線形化する必要がある。そこで、手分析データ値を元に塩素消費モデルを線形化し、本線形モデル上でオンライン測定した水質に基づき数理計画問題として定式化、多数の解候補の中から妥当な解を手分析データとしてリアルタイム推定し導出する。[1] では、予備シミュレーションの結果に基づき、求解パラメータとして $r_1=3.0$, $r_2=1.0$, $r_3=9.0$ を採用したが、本値を調整することで、より適切な解探索範囲が設定されると想定される。手分析データより得られる水質項目 x の標準偏差 σ_x 、求解パラメータ $\mathbf{r}=[r_1, r_2, r_3]$ に基づく手分析データ推定モデルにより連続するデータ区間において計算した各水質項目 x の標準偏差 $\hat{\sigma}_x(\mathbf{r})$ を元に、下記の式に基づき最適となるパラメータ \mathbf{r} を計算する。

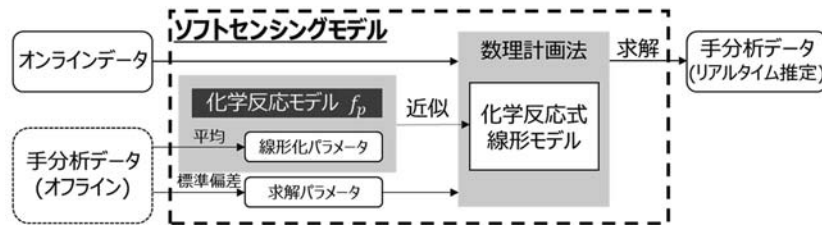


Fig. 1 Manual-sampling data estimation model.

$$\mathbf{r} = \arg \min_{\mathbf{r} \geq 0} \sum_x |\sigma_x - \hat{\sigma}_x(\mathbf{r})|$$

最適化手法には最急降下法や共役勾配法などが考えられるが、 $\hat{\sigma}_x(\mathbf{r})$ が非連続かつ複数の局所解を有することが懸念される点や、本パラメータ導出がオフライン計算であり計算速度について言及しない点を鑑み、モンテカルロ法を採用する。この手分析データに基づく最適パラメータ \mathbf{r} の導出過程を学習として定義し、学習データに応じた手分析データ推定モデルを構築する。

3. シミュレーション評価

3.1 評価手法

評価項目1として、手分析データ推定値の妥当性を実サンプルデータとの比較により評価する。長野市犀川浄水場を評価対象機場とし、オンラインデータとして1分単位の原水流量、前次亜塩注入率、沈澱池入口残留塩素濃度を、オフラインデータとして本浄水場内において約1時間単位（採水時刻による最大20分程度のバラつきあり）で取得した24時間分の手分析データ2セット（Set 1, 2と表記）を用いる。手分析データの平均、標準偏差をTable 1に示す。本データにおける分析項目は溶解性鉄（以降、Feとして表記）、溶解性マンガン（Mn）、溶解性全有機炭素（TOC）、還元性窒素化合物であるが、還元性窒素化合物については検出限界以下であったため本報告では省略する。手分析データ推定モデルは異なる2条件において評価を実施し、条件1については既報と同値となる $r_1=3.0, r_2=1.0, r_3=9.0$ を求解パラメータとして適用、条件2については2章において提案したパラメータチューニングによる最適値として、学習データSet 1に対し $r_1=2.0, r_2=1.0, r_3=4.3$ 、学習データSet 2に対し $r_1=4.5, r_2=1.0, r_3=5.2$ を求解パラメータとして適用する。評価は、クロスバリデーションで行う。すなわ

Table 1 Summary of water analysis data.

		Fe	Mn	TOC
Set 1	平均	0.037 mg/L	0.006 mg/L	0.610 mg/L
	標準偏差	0.007 mg/L	0.001 mg/L	0.024 mg/L
Set 2	平均	0.043 mg/L	0.004 mg/L	0.542 mg/L
	標準偏差	0.008 mg/L	0.001 mg/L	0.024 mg/L

ち、Set 1または2を学習データ（入力用のオフラインデータ）として用いた場合の、Set 2または1のデータ期間の手分析データ推定結果を評価する。

評価項目2として、手分析データ推定モデルの有無によるろ過水残留塩素濃度の予測精度への影響について評価する。本報告では伏流水との合流による影響を排除する目的でろ過水を対象とする。犀川浄水場の2019年6月25日～2019年12月18日を対象にろ過水残留塩素の予測値と実測値との絶対誤差について、手分析データ推定モデルなしの場合と、ありの場合（条件1, 条件2）で評価する。手分析データ推定モデルなしの場合は、化学反応式への入力として先述の手分析データ2セットの平均値を全データ期間一定の値として用いる。手分析データ推定モデルありの場合は先述の条件1および2のもと手分析データ推定を実施する。ただし、条件2については、手分析データ2セット分を学習データとして用いる。学習は連続データに基づくが、2セット間は連続しない。そのため、データセット毎に算出した最適パラメータ値の平均値 $r_1=3.2, r_2=1.0, r_3=4.8$ を求解パラメータとして用いる。

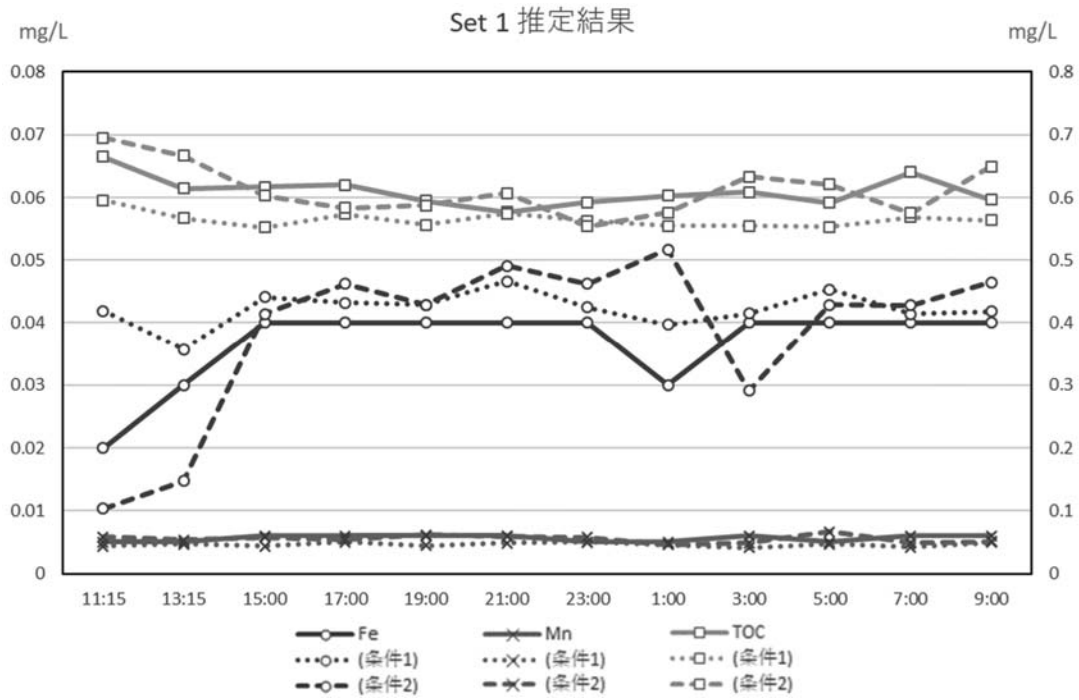
3.2 評価結果

3.2.1 手分析データ推定

Fig. 2にSet 2または1を学習データ（入力用のオフラインデータ）として用いた場合の、Set 1または2のデータ期間の手分析データ推定結果を示す。実線は各水質項目の実測値、点線および破線はそれらの手分析データ推定値（それぞれ条件1および2）を示す。なお、スケールが大きく異なるため、TOCのみ第2軸（右軸）に示す。

定性的評価として、いずれの条件下でも実測値の近傍値を推定している。ただし条件1は学習に用いた平均値を中心として変動が小さいのに対し、条件2は変動が大きく実測値のトレンドに追従している。その一例としてFeに関して、Fig. 2(a)における11:15-15:00にかけての上昇や23:00-1:00における減少、Fig. 2(b)における5:00-10:15にかけての減少について実測値と同傾向の推定結果が見られる。求解パラメータ最適化により、追従性能が向上していると考えられる。その他の水質項目については、サンプリングデータ区間における変動が少なくトレンドとしての比

(a) training data : Set 2, test data : Set 1



(b) training data : Set 1, test data : Set 2

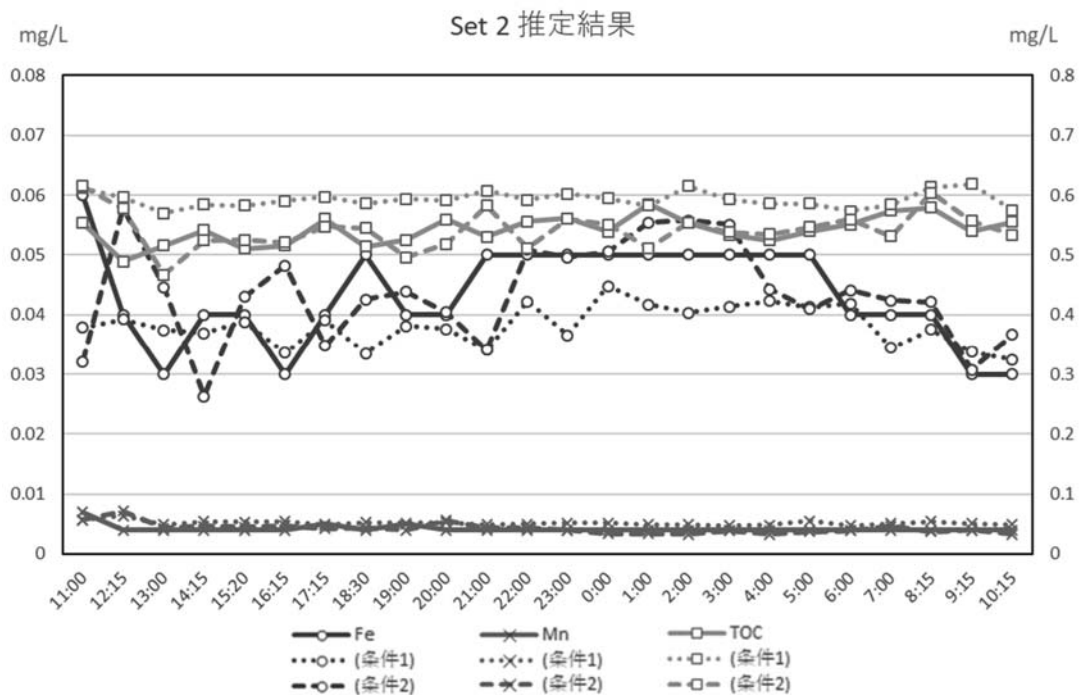


Fig. 2 Comparison between measured data and estimated data. (Primary Y axis : Fe, Mn, Secondary Y axis : TOC)

較は困難である。Fig. 2 (b) の条件 1 において、Fe が実測よりもやや低く、TOC が実測よりも高く推定している理由について、学習データ間で異なる平均値を用いたためと考えられる。一方で Fig. 2 (b) の条件 2 においては上記の乖離が解消されている。パラメータ調整により、学習データおよびテストデータ間で平均値が異なる場合においても性能改善が可能であることを示している。ただし、本報告における手分析データ推定モデルは、滞留時間分過去の前次亜塩素酸注入率に

対する現在の沈澱池入口残留塩素濃度を基準に原水水質を推定し現在値として用いている。そのため、手分析データ推定値と実測値の間には滞留時間分の遅れが発生することが想定され、それらは Fe について、Fig. 2 (a) 上の 23:00-1:00 や、Fig. 2 (b) 上の 11:00-18:30 に顕著に見られる。

定量的評価として、Table 2 に水質項目毎の絶対誤差の平均および標準偏差を示す。最も絶対誤差平均の小さい水質項目 (Set 2 の TOC 手分析データ推定値)

Table 2 Estimation results in normal state. (a) training data : Set 2, test data : Set 1. (b) training data : Set 1, test data : Set 2.

		Fe	Mn	TOC
(a) 条件1	絶対誤差平均 (正規化)	0.00556 mg/L (13.2%)	0.00096 mg/L (20.7%)	0.04507 mg/L (8.0%)
	絶対誤差標準偏差 (正規化)	0.00571 mg/L (13.5%)	0.00064 mg/L (13.9%)	0.01938 mg/L (3.4%)
(a) 条件2	絶対誤差平均 (正規化)	0.00796 mg/L (20.6%)	0.00068 mg/L (12.3%)	0.03416 mg/L (5.6%)
	絶対誤差標準偏差 (正規化)	0.00590 mg/L (15.3%)	0.00047 mg/L (8.5%)	0.01662 mg/L (2.7%)
(b) 条件1	絶対誤差平均 (正規化)	0.00673 mg/L (17.6%)	0.00109 mg/L (21.0%)	0.05096 mg/L (8.6%)
	絶対誤差標準偏差 (正規化)	0.00560 mg/L (14.7%)	0.00044 mg/L (8.5%)	0.02436 mg/L (4.1%)
(b) 条件2	絶対誤差平均 (正規化)	0.00738 mg/L (16.9%)	0.00067 mg/L (15.8%)	0.02792 mg/L (5.1%)
	絶対誤差標準偏差 (正規化)	0.00712 mg/L (16.3%)	0.00065 mg/L (15.2%)	0.02393 mg/L (4.4%)

で推定誤差 5.1% となった。正規化値で比較した場合に、いずれの結果についても、平均および標準偏差とも TOC が良好な性能を示している。理由として、Fe および Mn は TOC と比較して絶対量が少ないため微小な変動が誤差に大きく影響を与えることが考えられる。特に、Fe は他の水質項目値と比較して短時間での変動が大きく、追従遅れによる推定精度悪化が懸念される。そのため、変動の少ない手分析データ推定モデル (条件 1) の方が Fe については良好な推定精度を示していると考えられる。原水水質の状況により適切にモデルを切り替える、あるいはパラメータ調整することにより更なる性能改善が可能であると考えられる。

3.2.2 ろ過水残留塩素濃度予測

Table 3 にろ過水残留塩素濃度予測の性能評価を示す。条件に関わらず手分析データ推定による予測性能改善が確認され、手分析データ推定あり (条件 2) では手分析データなしと比較し絶対誤差平均が 25% 改善し、正規化値 2.4% となった。対象機場や予測対象水質が異なるため単純比較はできないが、既報における正規化値 10.9% 以下となる良好な結果であるといえる。手分析データ推定の条件 1 と条件 2 を比較すると、条件 2 の方が平均、標準偏差いずれについても改

Table 3 Prediction results of the residual chlorine concentration in the filter basin.

	手分析データ 推定なし	手分析データ 推定あり (条件 1)	手分析データ 推定あり (条件 2)
絶対誤差平均 (正規化)	0.023 mg/L (3.2%)	0.019 mg/L (2.6%)	0.017 mg/L (2.4%)
絶対誤差標準偏差 (正規化)	0.027 mg/L (3.9%)	0.022 mg/L (3.1%)	0.021 mg/L (2.9%)

善されており、パラメータ調整による効果がみられる。48 時間分の限られた学習データであってもパラメータ調整により全テストデータ区間に対し一定の予測精度改善効果があると考えられる。

また、原水水質が実サンプルと大きく異なる場合の影響の検証を目的として、原水濁度 20 mg/L 以上を除外区間とし手分析データ推定あり (条件 2) において再評価した。その結果、絶対誤差平均 0.017 mg/L、絶対誤差標準偏差 0.017 mg/L となった。平均の改善効果はみられなかったものの、標準偏差すなわちバラつきが約 19% 改善した。ろ過水残留塩素濃度予測性能についても、原水水質の状況に応じた適切なモデル切り替えあるいはパラメータ調整により改善が可能であると考えられる。

4. ま と め

本報告では、手分析データ推定モデルのパラメータ最適化手法の提案、および本推定値の実サンプルデータとの比較による妥当性検討を実施した。前者について、実サンプルデータの平均および標準偏差に基づくパラメータ最適化手法を提案した。後者について、クロスバリデーションによる評価を実施し、最も絶対誤差平均の小さい水質項目 (TOC) で平均推定誤差が 5.1% となった。手分析データ推定モデルの有無によるろ過水残留塩素濃度予測精度への影響について評価した結果、手分析データ推定により絶対誤差平均が 25% 改善した。またいずれについてもパラメータ最適化による性能改善を確認した。

謝 辞

本研究は、JWRC 研究課題公表による実証研究 (A-IDEA) のテーマ③：情報技術の活用による水道事業の基盤強化に資する研究かつ実フィールドでの実証を必要とする研究として採択された、「AI を活用した最適化技術 (塩素注入の最適化) の開発」プロジェクトの成果です。本研究の実証先として長野市犀川浄水場様にご協力いただき、浄水場内の計測データならびに検証用水質データの採取にもご協力いただきました。長野市犀川浄水場様に、この場をお借りして御礼申し上げます。

参 考 文 献

- 1) 松本 隼, 山原裕之, 横川勝也, 毛受 卓, 黒川 太, 横山 雄: 機械学習と化学反応モデルのハイブリッドによる残留塩素濃度予測, 環境システム計測制御学会誌, Vol. 24, No. 2/3, pp. 40-47 (2019)