

## 〈研究発表〉

### 機械学習を用いた余剰圧力削減のための末端圧力推定技術

山原 裕之<sup>1)</sup>, 松本 隼<sup>1)</sup>, 横川 勝也<sup>1)</sup>, 黒川 太<sup>1)</sup>  
横山 雄<sup>2)</sup>, 平岡 由紀夫<sup>2)</sup>

<sup>1)</sup> (株) 東芝

(〒183-8511 東京都府中市東芝町1 E-mail: hiroyuki.yamahara@toshiba.co.jp)

<sup>2)</sup> 東芝インフラシステムズ (株)

(〒212-0013 神奈川県川崎市幸区堀川町72-34 E-mail: suguru.yokoyama@toshiba.co.jp)

#### 概要

管網に圧力計を設置していない配水プロセスにおいても余剰圧力を削減して配水を省エネルギー化するために、末端圧力を推定する技術を開発している。推定モデルは、入力が必要なパラメータ数が少なく、少量の訓練データで生成できることが望ましい。本論文では、機械学習を用いて、配水施設からの吐出圧力と吐出流量を含む限られた情報を基に末端圧力を推定する手法を提案する。実データを用いたシミュレーションの結果、本手法は30日分の訓練データから生成した推定モデルで、誤差1 mAq以下の推定精度を実現できる可能性を得た。

キーワード：配水管網、余剰圧力、圧力推定、ランダムフォレスト

原稿受付 2018.7.2

EICA: 23(2・3) 143-147

## 1. はじめに

環境負荷削減の観点から、国内電力消費の約1%を占める水道においてもエネルギーの削減が着目されている。そのうち、水の輸送に関わるエネルギーは、送配水過程と取水・導水過程を合わせて約66%を占めており<sup>1)</sup>、配水プロセスにおける省エネルギー化の取り組みが求められている。取り組みのひとつとして、管内の圧力を把握することによる配水コントロール<sup>2)</sup>が挙げられる。自然流下でなく、ポンプ圧送する配水プロセスにおいては、必要最低限の圧力を維持するようにポンプを制御することで、配水を省エネルギー化できる。圧力制御は、漏水削減や運用に関わる費用の削減にも繋がる。

圧力制御方法のひとつとして、センサで取得した末端圧力の値を監視しながら、必要最低限の圧力を維持し、かつ余剰が無いように適切な値に制御する、末端圧力制御が挙げられる。実際には、末端圧力制御を行わず、必要より高い圧力で配水している中小規模の水道事業者も存在する。これらの事業者が末端圧力制御を導入していない理由として、圧力計の設置コストが高いことが挙げられる。センサによる監視を行わず、管網解析<sup>3)</sup>シミュレーションによって圧力を推定する方法も考えられる。管網解析は、管路の接続関係、管長、管径、標高、水需要など、入力が必要な情報が多い。また、スマートメータが無く、水需要をオンラインで取得できない場合、管網解析で圧力を計算するこ

とができない。そこで、管網解析を実施せずに、限られたセンサ情報から簡便に末端圧力を推定する手法を提案する。提案手法は、機械学習を用いて、配水施設からの吐出圧力と吐出流量を含む限られた情報を基に圧力推定モデルを生成する。本手法は、以下の特長を持つ。

- ・管網解析と比べて、少ない情報から推定できる。
- ・短期間の実測データで推定モデルを学習できる。
- ・常設の圧力計を減らすことができ、設備投資コストを抑えて、末端圧力制御を導入できる。

本論文では、A市の配水データに提案手法を適用し、異なる訓練データサイズで生成した複数の推定モデルによる末端圧力推定精度を比較した結果を報告する。

## 2. 対象とするシステム

図1に本論文で対象とするA市のシステムの概要を示す。A市は複数の配水区を持つが、対象としたシステムは、そのうち配水区Xと配水区Yから成る。配水区Xは、2つの配水場(D1, D2)を持つ。配水区Yは、1つの配水場(D3)を持つ。2つの配水区は、連絡調整弁Vを通じて水を融通する。この際、水は配水区Xから配水区Yの方向に流入する。基本的に1時から6時頃の夜間時間帯のみ弁Vが開かれる。配水区Xは配水区Zとも隣接しており、連絡調整弁Wを通じて水を融通し合う。弁Wに関しては、どちらの方向にも水が流れうる。送水場S1は、弁Wを

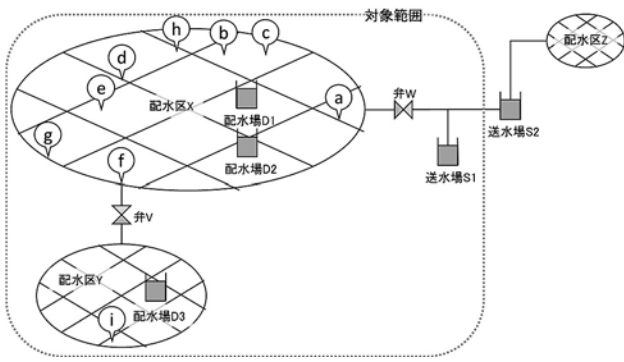


Fig. 1 Target system of the proposed pressure estimation method.

通じて配水区 X に水を供給するとともに、送水場 S2 を介して配水区 Z にも水を供給している。ここで、配水区 Z との間には揚水のための送水場 S2 が設置されていることから、配水区 Z は対象に含まない。本論文では、3つの配水場 (D1, D2, D3)、1つの送水場 S1、2つの弁 (V, W) を含む配水区 X と Y を対象範囲として、提案手法を評価する。

対象範囲内の 9 地点 (地点 a から地点 i) に圧力計が常設され、末端圧力が計測されている。本来は A 市の末端圧力を推定する必要はないが、本論文では、常設の圧力計の値を真値として手法を評価するために A 市のデータを用いた。地点 a, b, c には、主に配水場 D1 から配水される。地点 d, e, f, g, h には、昼間は主に配水場 D2 から配水されるが、1時から6時頃の夜間には D2 の配水は停止され、配水場 D1 から配水される。地点 i には、昼間は配水場 D3 から配水されるが、夜間には D3 の配水が停止されて弁 V が開かれることで配水場 D1 から配水される。各配水場はポンプで配水しており、ポンプを制御することで末端圧力を調整可能である。

### 3. 末端圧力推定手法

提案手法では、機械学習の一種であるランダムフォレストを用いて推定モデルを生成する。ランダムフォレストは、個々に学習した複数の弱学習器を組み合わせる集団学習によって、汎化性能が高い学習器を生成可能なアルゴリズムである<sup>4)</sup>。

提案手法で用いる推定モデルの目的変数 $y_n(t)$ と説明変数を、式 (1) に示す。ここで、 $n=\{a, b, c, d, e, f, g, h, i\}$  の 9 地点をあらわす。

$$y_n(t) = f(t, p, q, j, k, l) \quad (1)$$

式 (1) に示すように、ある時刻 $t(1 \leq t \leq 24)$ における地点 $n$ の圧力 $y_n(t)$ は、時刻 $t$ 、配水場および送水場から配水区に水を送る際の吐出圧力 $p=\{p_{D1}, p_{D2}, p_{D3}, p_{S1}\}$ および吐出流量 $q=\{q_{D1}, q_{D2}, q_{D3}, q_{S1}\}$ 、弁の開度 $j=\{j_V, j_W\}$ 、弁の設置個所における流量 $k=\{k_V, k_W\}$ 、圧力

$l=\{l_V\}$ から推定される。圧力 $l$ に関しては、弁 W では圧力が測定されていなかったため、弁 V の値のみ用いる。推定モデルは地点ごとに生成されるが、説明変数として用いる項目は、推定対象地点によらず共通である。

$p, q$  を説明変数として用いるのは、送り出し元の圧力と流量の変動が、末端圧力の変動に最も影響を与えらるからである。管網に圧力計を設置していない事業者でも、配水場や送水場の運転の監視制御のために吐出流量や吐出圧力の値は計測していることが想定されるため、これらの値を取得するために新たにセンサを設置する必要が無いと期待される。また、 $j, k, l$  を説明変数に用いるのは、弁の情報を用いることで、より精度が高い末端圧力推定が可能になると考えられるからである。時間帯によって開度や水の流れる方向が弁でコントロールされていることがあり、弁の状態は管網全体の状態に大きな影響を与える。

本手法は、圧力計が設置されていない地点の圧力を推定することを目的としているが、推定モデルを生成するにあたり、一定期間は推定対象地点に臨時的圧力計を設置してデータを収集する。生成後は圧力計を撤去する。臨時で設置する圧力計は、必ずしもオンラインでデータ収集する機能を有している必要はないため、比較的安価に実績圧力値を収集可能な利点がある。

### 4. 推定手法の評価

提案手法の推定精度を評価するために、A 市のある年の 1 月 1 日～翌年 6 月 30 日までの 18ヶ月 (547 日) 間の時単位データを用いて、シミュレーションを実施した。推定精度の評価指標として、式 (2) に示す RMSE (根平均二乗誤差) を用いた。

$$\text{RMSE (mAq)} = \sqrt{\frac{1}{U} \sum_u (D_{est_u} - D_{act_u})^2} \quad (2)$$

ここで、 $D_{est_u}$ : 圧力の推定値、 $D_{act_u}$ : 圧力の実測値、 $u$ : 推定対象日時、 $U$ : 統計対象時刻数、である。

図 2 に地点ごとの推定対象時刻別の平均実績圧力を

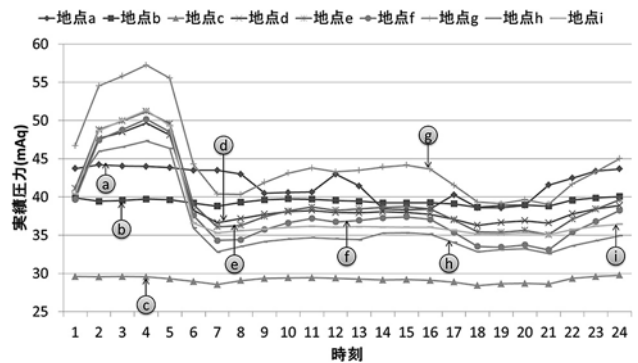


Fig. 2 Average nodal pressure in actual data by time.

示す。昼夜を問わず配水場 D1 から配水される地点 b, c に関しては昼夜の圧力差がほとんどなく 2 mAq 以下である。地点 a ではそれより大きく、1 日のうちで 6 mAq 程度の変動が見られる。これは弁 W を介して水を融通し合った影響だと考えられる。その他の 6 地点は 2 時から 5 時の夜間時間帯に昼間より 15 mAq 程度圧力が高くなっており、他の時間帯に比べて余剰圧力の削減余地がある。

#### 4.1 訓練データ期間長別の推定精度

推定モデルを学習するための訓練データとして使用するデータの期間長を、30 日、60 日、90 日、180 日、360 日に設定した場合の推定精度を比較した。図 3 に、訓練データ期間長が 180 日の例を用いて、547 日分のデータを訓練データと評価データに分割する方法を示す。全データの中から、連続する 180 日のデータを訓練データとし、残りの 367 日分を評価データとする。18ヶ月のうち、各月の 1 日から 180 日間を訓練データとすると、分割パターンは 13 パターン存在する。分割パターンごとに、RMSE を算出し、13 個の RMSE の平均を全体の誤差とする。訓練データ期間長に応じて分割パターンの数は異なるが、180 日の例と同様に分割パターンごとに RMSE を算出して、それらの平均を全体の誤差とする。

図 4 に、9 地点の各データ期間長における誤差を示す。どの地点においても、訓練データ期間長が長いほど、

誤差は小さくなる傾向がある。データ量が多く、年間の様々な時期のデータを基に学習した推定モデルの方が高精度であることは、想定できる結果である。しかし実際には、訓練データ期間長が 30 日と 360 日を比較しても、誤差に大きな差はない。最も差が小さい地点 i は、30 日で 0.80 mAq、360 日で 0.74 mAq であり、その差は 0.06 mAq である。最も差が大きい地点 a でも、30 日で 4.31 mAq、360 日で 3.56 mAq であり、その差は 0.75 mAq である。夜間時間帯の実績圧力の最低値と比較して最大でも 2% 未満の値であり、差が大きくないと考えられる。この結果から、提案手法は 30 日程度の短期間の訓練データで、360 日(約 1 年)の訓練データを用いた場合と性能に大差が無い推定モデルを生成できたと言える。臨時的圧力計を設置すべき期間が短くなることから、短期間の訓練データで推定モデルを生成できることは望ましい結果である。以降は、訓練データ期間長が 30 日のシミュレーション結果に着目する。

#### 4.2 訓練データ期間長 30 日の推定精度

図 4 の訓練データ期間長 30 日の場合の誤差に着目する。各地点の誤差の大小には、配水場からの配水距離、弁からの距離、1 日の圧力変化の大小、といった特徴が関係していると思われる。これらの特徴を、表 1 に示す。

地点 a の誤差は 4.31 mAq であり、他の地点より大きい。送水場 S2 への供給の影響は推定モデルで表現できていないため、推定精度を悪化させたと考えられ

Table 1 Influencing characteristics on each node.

地点	配水場からの配水距離		連絡調整弁からの距離	1 日の実績圧力の変化
	昼間	夜間		
a	短	短	近	中
b	短	短	遠	小
c	中	中	遠	小
d	中	中	遠	大
e	中	中	遠	大
f	中	長	近	大
g	長	長	遠	大
h	長	短	遠	大
i	短	長	遠	大

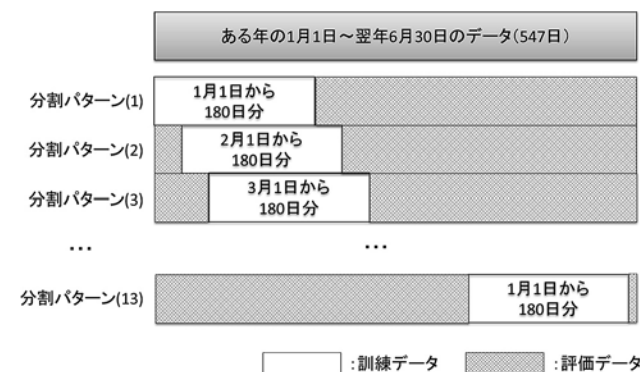


Fig. 3 How to divide sample data set to training data set and evaluation data set.

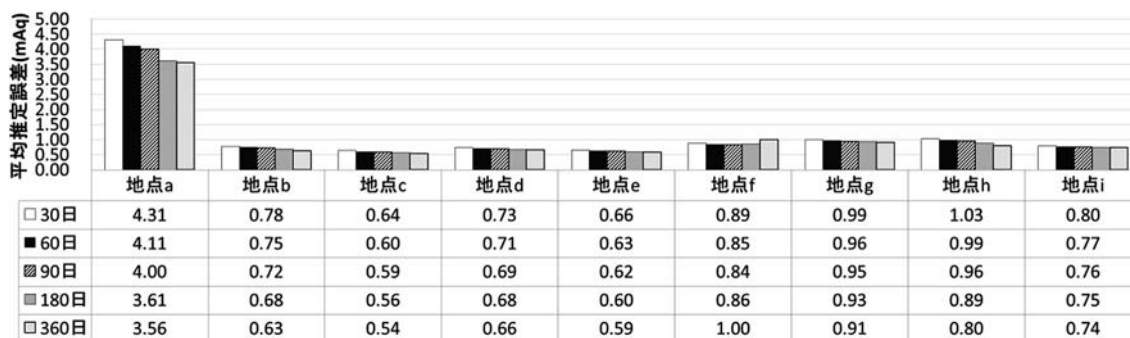


Fig. 4 Estimation accuracy by size of training data set.

る。送水場 S1 が供給する水は地点 a がある配水区 X だけでなく、送水場 S2 にも供給される。また、地点 a は弁 W の近く位置している。しかしながら、前述のとおり弁 W に関しては圧力が計測されていない。仮に弁 W の圧力の値が得られた場合は、地点 a の推定精度は高くなる可能性がある。

地点 a を除く 8 地点は、概ね 1 mAq 以下となった。提案手法は、末端圧力と配水施設からの吐出圧力および吐出流量の関係性をモデル化しており、管路の接続関係や摩擦損失など管網の細かな情報は用いていない。そのため、配水距離が長くなるほど、管網の複雑さの影響を受けて、誤差が大きくなる可能性がある。このことから、8 地点の中でも地点 h と g は昼間の配水距離が長いので、推定精度が下がったと考えられる。また、次に誤差が大きい地点 f は、配水距離の長さに加え、弁 V の近くに位置することから、他の地点に比べて流量や圧力の変化が多様であることが影響して、誤差が大きくなったと考えられる。

#### 4.3 訓練データ月別の推定精度

図 5 に、訓練データ月別の RMSE を示す。地点 c, d, e, g, i に関しては、訓練データが何月のデータであっても、ほとんど誤差に差はない。最大最小差は 0.1 mAq 以下である。地点 b, h は、最大最小差がそれぞれ 0.18 mAq, 0.17 mAq である。2 地点に共通の傾向として 7 月～8 月のデータを訓練データとした場合に、他の月のデータを用いるよりも誤差が小さくなっている。しかし、その差はわずかであり、やはり訓練データ月は推定精度に影響がないと言える。地点 f は、最小最大差が 0.38 mAq である。1 月を訓練データとした場合のみ、誤差が大きい。他の月を訓練データとした場合の誤差は 0.85 mAq 前後でほぼ揃っていることから、1 月のデータには何か特殊な管網の状態が含まれていたと考えられる。地点 a も同様に、6 月だけ突出して誤差が大きい。特別なイベントによ

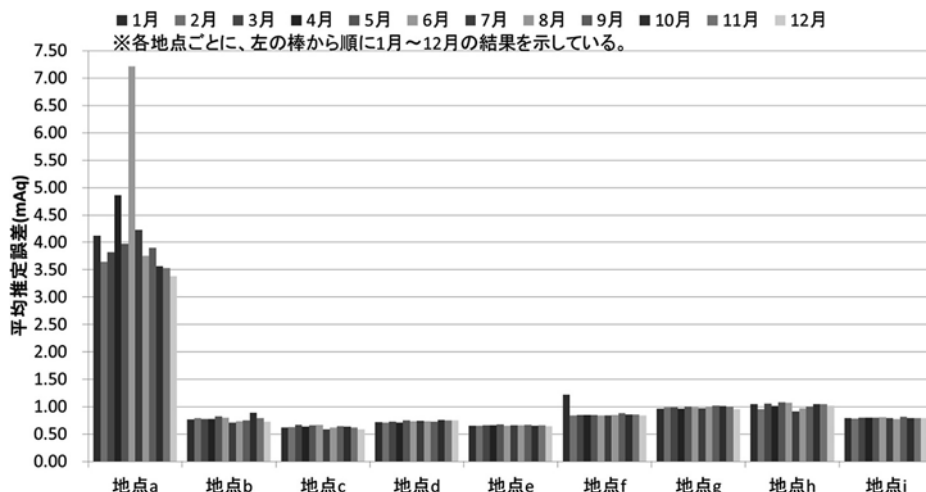


Fig. 5 Estimation accuracy by month of training data.

る水需要の変化、あるいは管路の接続状況の一時的な変化があった場合、その月のデータを訓練データとして用いた推定モデルは、汎化性能が低い可能性がある。

#### 4.4 推定対象時刻別の推定精度

図 6 に、地点ごとの推定対象時刻別の RMSE を示す。多くの地点で、水需要が大きくなる朝の 6 時頃に誤差が大きい。地点によっては、5 時と 6 時の差分が最大 0.62 mAq である。これは、水需要に分単位の急激な変化がある時間帯の吐出圧力・吐出流量と末端圧力の関係性を、シミュレーションで用いた時単位のデータでは表現できていないからだと考えられる。また、地点 d～i に関しては、供給元の配水場が切り替わる時間でもあり、時単位データでは切り替わりの過渡期における関係性を表現できていないことも考えられる。

各地点の誤差の大小は、表 1 に示した特徴と関連付けられる。まず地点 a は、配水距離は短いものの、弁 W に近く、送水場 S1, S2 の影響を受けるため誤差が大きい。最も誤差が小さい 17 時でも、4.37 mAq であ

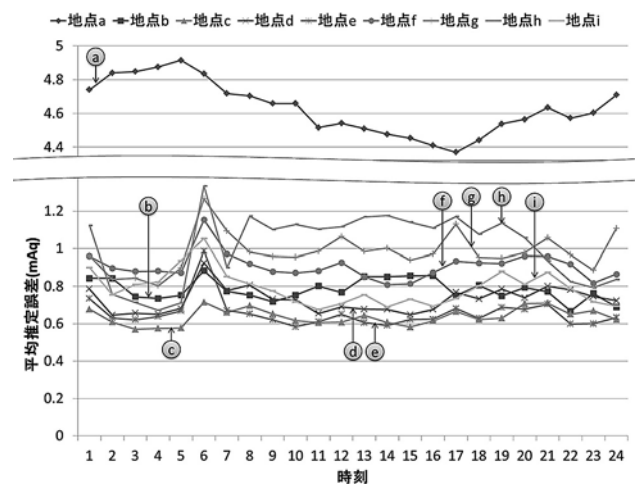


Fig. 6 Average estimation accuracy by time of evaluation data.

る。弁 W は、水の方向が双方向、かつ、弁が開く時間帯も限定されていないことが、推定を難しくしていると考えられる。地点 b, c は、実績圧力の1日の中での変化が小さく、推定しやすいため、誤差はどの時刻においても小さく、0.9 mAq を下回った。地点 d, e は、配水場 D1 から D2 から配水距離が中程度であることから、あまり誤差は大きくなり、朝の6時頃を除き 0.9 mAq を下回った。地点 f, g, h は、配水距離が長いため、地点 b~e より誤差が大きい。昼間は、配送距離が長い g, h の誤差が f よりも大きくなっており、時間帯によっては誤差が 1 mAq を超えている。地点 h に関しては、夜間の誤差は小さくなっている。これは、夜間は配水距離が短くなることの影響したと考えられる。このように、各地点の特徴に応じた時刻ごとの大小の傾向が見られる。しかしながら、地点 a を除くと、2時から5時の夜間時間帯は地点によらず誤差が 1 mAq を下回っている。

図7には、地点ごとの推定対象時刻別の最大推定誤差を示す。地点 b, c は実績圧力の変化が小さいことから、推定しやすいため、最大誤差も小さい傾向にある。多くの時刻で 5 mAq を下回っている。地点 d~i の昼間は、配水距離が長いほど、最大誤差も大きくなる傾向が見られる。配水距離が中程度の地点 d, e, i は 2~6 mAq の時刻が多いのに対し、より長い地点 f, g, h は 4~9 mAq の時刻が多い。一方、夜間はどの地点も 4~8 mAq 程度であり、配水距離の違いによらず推定できている。

平均推定誤差、最大推定誤差の両方の傾向から、夜間は昼間に比べて安定した推定が可能であり、提案手法による推定結果を用いて夜間の余剰圧力を削減することができると考えられる。たとえば、夜間の実績圧力が 40 mAq、最低限必要な圧力が 15 mAq とすると、25 mAq の余剰圧力が存在する。これに対し、推定誤差を 8 mAq 見込んだとしても、17 mAq は削減可能

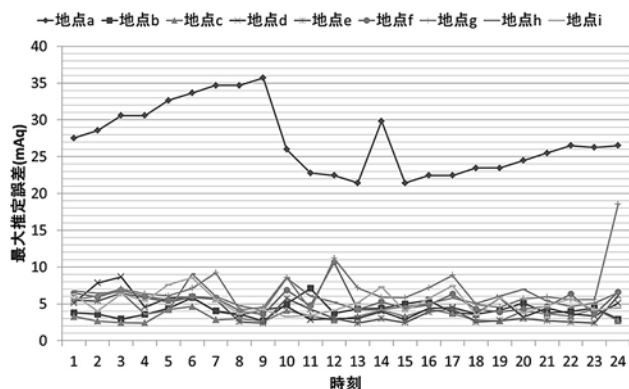


Fig. 7 Maximum estimation error by time of evaluation data.

である。

#### 4.5 フィールド適用に向けた課題

地点 a の推定精度を改善するためには、弁 W の圧力値を説明変数に含めるなどして、送水場 S2 側との影響を考慮した推定モデルとする必要がある。提案手法を実フィールドに適用する場合、弁、揚水施設、異なる配水区との関係性を考慮して、重要な箇所には常設の圧力計を設置し、その値を活用することも考えられる。

臨時の圧力計を用いて訓練データを取得する時期を決定する際、何らかのイベントで特定の日や期間のみ管網の状態が通常と異なる時期は基本的には避ける必要がある。しかしながら、そのようなイベントが年に1回など特定の頻度で発生することが分かっているのであれば、むしろその時期のデータも訓練データに取り込むことで、より多くの状況で高い推定精度を実現しうる。

前述のとおり、時単位のデータでは朝方の急激な水需要の変化の影響を捉えきれず、推定精度が低下している可能性も考えられる。また、末端圧力制御を実現するには時単位の圧力を推定するだけでは時間粒度が大きすぎる。今後、より細かな時間粒度のデータで、提案手法の推定精度を評価する必要がある。

## 5. おわりに

本論文では、管網解析を実施せずとも、限られたセンサ情報から末端圧力を推定する手法を提案した。A市の実データを用いたシミュレーションの結果、概ね 1 mAq 以下の誤差で推定可能であった。また、訓練データが 30 日分であっても、約 1 年分の訓練データを用いた場合と差がない推定モデルを生成できることを示した。今後、推定結果をもとに余剰圧力削減効果を試算する。

#### 参考文献

- 1) 東京都水道局：東京水道エネルギー効率化 10 年プラン (2015)
- 2) 横川勝也、稲貝とよの、小泉明、難波諒、杉野寿治：末端圧力推定に基づく夜間バルブ操作による漏水削減、土木学会論文集 G (環境), Vol. 72, No. 7, pp. III\_449-III\_455 (2016)
- 3) 高桑哲男：配水管網解析の基礎と応用、水道管路技術センター技術レポート, Vol. 12 (1992)
- 4) Leo Breiman: Random Forests, Machine Learning, Vol. 45, Issue 1, pp. 5-32 (2001)