

〈論文〉

異常検知 AI を用いた漏水検知手法の開発

隋 鵬 哲¹⁾，島 村 和 彰²⁾，多 田 啓 太 郎³⁾，角 純 平⁴⁾¹⁾水 ing(株) 研究開発センター
(〒251-8502 神奈川県藤沢市本藤沢4-2-1 E-mail:sui.pengzhe@swing-w.com)²⁾水 ing(株) 研究開発センター
(〒251-8502 神奈川県藤沢市本藤沢4-2-1 E-mail:shimamura.kazuaki@swing-w.com)³⁾水 ing(株) 総合水事業本部
(〒105-0021 東京都港区東新橋1-9-2 E-mail:tada.keitaro@swing-w.com)⁴⁾水 ing(株) 総合水事業本部
(〒105-0021 東京都港区東新橋1-9-2 E-mail:sumi.jumpei@swing-w.com)

概 要

水道管路の漏水事故は、グローバルな課題であり、地域的な減水・断水を招き、道路の陥没などの二次被害を生じる恐れがある。近年人工知能（AI）の急速な進歩があり、本研究では AI を用いた異常検知手法を漏水検知に活用し、配水量の異常変化を検知して迅速に警報を出すことを目標とした。K 市の配水データを用い、One-Class SVM というアルゴリズムを利用し、異なる時刻の配水流量データの関連性や時系列データの周期性を考慮し、時刻毎の One-Class SVM モデルを構築することで、通常の配水流量と比べて高くなった異常流量を検知できることを確認した。また、時刻毎の One-Class SVM モデルの構築より、一日の全体的な流量異常を検知でき、特に夜中の低流量の異常上昇も検知できた。結果より、漏水ではない場合の配水流量上昇を誤検出するという課題があるが、検討した手法は漏水被害の最小化に有効な手法であると考えられる。

キーワード：異常検知，漏水，人工知能

原稿受付 2023.6.6 原稿受理 2023.7.24

EICA: 28(2・3) 3-10

1. はじめに

漏水に関する研究によく使われている言葉である無収水（Non-Revenue Water, NRW）は、給水量と水道料金を徴収した水量の差であり、①実際の漏水（Real Losses, Physical losses；物理的漏水）、②水道メーターの計測誤差等からの見かけ漏水（Apparent losses）、及び③消防等水道料金を徴収しない用水（Unbilled authorized water use）で構成されている。本稿の漏水検知は、その中の①実際の漏水（物理的漏水）を指す。水道管路の漏水事故は、地域的な減水・断水を招き、道路の陥没などの二次被害を生じる恐れがある。水道管路の漏水は、グローバルな課題であり、上水に関する研究の焦点の一つである。ヨーロッパでは、漏水量は給水量の10~50%を占めている¹⁾。日本の漏水率は、全国平均では給水量の5%と海外の都市と比較し低い漏水率を実現している²⁾。しかし、高度経済成長期に整備された水インフラの老朽化及び地震などの自然災害により、断水・漏水事故などの発生リスクの増大が懸念されている。近年では人工知能（Artificial Intelligence, AI）の急速な進歩により、AIを用いた漏水検知技術が注目されている。水道管路に

漏水検知センサーを配置し、AIによる漏水音を検知する手法が報告された³⁾。日本の研究者も、管路に設置する漏水検知センサーの必要数⁴⁾、漏水音の時系列データを用いた AI モデル構築手法の検討^{5,6)}が報告された。また、読売新聞では、愛知県豊田市が人工衛星のデータを活用した水道管の漏水調査を全国で初めて行ったことが掲載された⁷⁾。リモートセンシングの衛星データも、ローカルセンシングの漏水音データも、漏水検知の精度が高いものの、データ収集のコストが高い。一方、漏水が発生する際に、流量や圧力の異常変化により、追加センサーが不要でも水道管路のデータより漏水を検知する可能性がある。フランスの研究者は、大学キャンパスの水道管路を対象とし、ケーススタディで AI より送水流量と圧力データを用いた漏水検知手法が検討された。結果として、教師なし学習した AI モデルより、教師あり学習した AI モデルの予測精度が高いことが示唆された。しかし、実際の活用を考えると、漏水流量などの情報を明確的に確認するのは難しいということは漏水検知 AI 開発の課題になる。教師なし学習した漏水検知 AI の開発は重要である。そこで、本研究では、教師なし AI を用いた異常検知手法を漏水検知に活用し、配水量の異常変化を

検知して迅速に警報を出すことを目標にした。

2. AI を用いた漏水検知手法及びデータ

異常検知については、外れ値検出 (Outlier detection), 変化点検出 (Change point detection), 異常状態検出 (Anomaly state detection) という三つのタイプがある。今回は、外れ値検出に注目し、One-Class SVM⁸⁾ (OCSVM) というアルゴリズムに対して、検知手法の有効性を検討した。OCSVM はサポートベクターマシン (Support Vector Machine, SVM) を 1 クラスの場合に改良したもので、正常データとして 1 つのクラス分を学習させ、識別境界を決定することで、その境界を基準に外れ値を検出する。異常がほとんど発生せず、異常クラスのデータが集まらないようなシステムで異常検知を実現したい場合には有効な検知手法である。

K 市にある A 配水池及び B 配水池の配水データ (Table 1) を用い、漏水検知への活用を試みた。収集した配水データは、1 時間毎の平均値であり、配水池水位、配水流量、及び流入流量を含んでいるが、今回配水流量に注目し、二つの配水池の配水流量データの異常値を検知した。OCSVM の検討は、Python を使用し、ライブラリー「Scikit-learn」を用いた。OCSVM の計算条件については、RBF カーネル (Radial Basis Function Kernel) を利用した。RBF カーネルのパラメータは、gamma はデフォルトの「Scale」で、nu は 0.5 とした。

Table 1 Water distribution data used for water loss detection

場所	期間	時間間隔	項目	
A 配水池	2014 年 11 月 14 日 2023 年 03 月 16 日	14:00~ 23:00	1 時間間隔	配水池水位
				配水流量
				A1 流入流量
				A2 流入流量
B 配水池	2014 年 11 月 20 日 2023 年 03 月 16 日	16:00~ 23:00	1 時間間隔	配水池水位
				配水流量
				B1 流入流量

3. 結果及び考察

3.1 アルゴリズムの検討

Scikit-learn のワインデータセットを用い、OCSVM の有効性及び課題を確認した。ワインデータセットは、ワインに関する 13 の特徴量がある。その中から、2 つ (color intensity, flavonoids) を選択し、flavonoids を横軸、color intensity を縦軸としてプロットした。このデータの 2 次元分布図から異常データを検知した結果を Fig. 1 に示す。紫色の線はデータの分布より決めた境界線であり、境界線範囲以外の

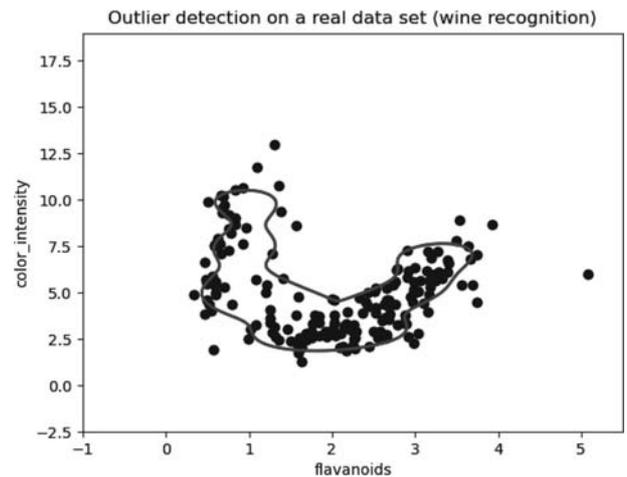


Fig. 1 Test results of One-Class SVM to confirm its property

データは異常値として判断する。Fig. 1 の結果より、境界線の範囲はデータの分布とよく一致していることが示唆された。通常の教師データなしの異常値検知は、データの分布による影響が大きいと言われ、ガウス分布など特定の分布を仮定することが多い。OCSVM は、特定の分布を仮定せず、複雑な分布であるデータにも対応できるという特徴がある。しかし、Fig. 1 に示したように、境界線の範囲が狭く、レギュラーデータを異常値として検知した恐れがある。

配水池の配水流量データは、1 日の変動に対して、配水時刻を横軸、配水流量を縦軸とし、上記検討した手法を用い、配水流量データの異常値を検知できる。しかし、この場合は、配水流量の年間変動 (例えば本年度と昨年度) が少ないという仮説が前提になる。

3.2 K 市の配水データの統計解析結果

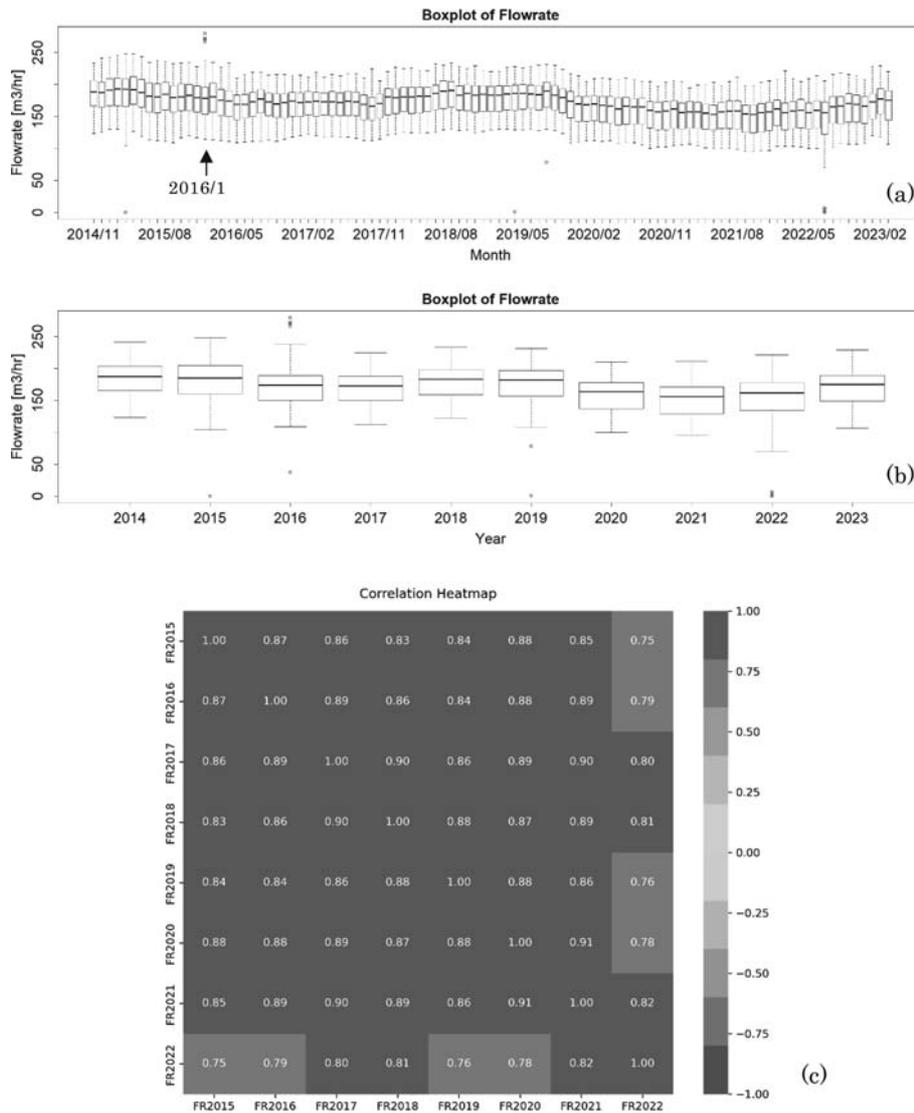
K 市の A 配水池と B 配水池の配水データの統計量解析結果を Table 2 にまとめる。A 配水池は、2014 年 11 月から 2023 年 3 月まで、約 8 年半の 1 時間間隔の配水流量 (平均値) のデータ (73,060 個) があり、配水流量データの平均値 $167.1 \text{ m}^3/\text{h}$ 、最小値 $0.2 \text{ m}^3/\text{h}$ 、最大値 $279.3 \text{ m}^3/\text{h}$ であった。B 配水池は、2014 年 11 月から 2023 年 3 月まで、約 8 年半の 1 時間間隔の配水流量 (平均値) のデータ (72,915 個) があり、平均配水流量 $18.2 \text{ m}^3/\text{h}$ 、最小値 $0.5 \text{ m}^3/\text{h}$ 、最大値 $67.9 \text{ m}^3/\text{h}$ であった。

A 配水池の配水データの解析結果を Fig. 2, B 配水池の解析結果を Fig. 3 に示す。2014~2023 年の 9 年間における、配水流量の月別箱ひげ図を Fig. 2 (a) と Fig. 3 (a) に示す。配水流量の年別箱ひげ図を Fig. 2 (b) と Fig. 3 (b) に示す。配水流量の年別の相関関係のヒートマップを Fig. 2 (c) と Fig. 3 (c) に示す。

A 配水池の解析結果について、Fig. 2 (a) より、2016 年 1 月に配水流量の異常上昇、箱ひげ図で異常

Table 2 Statistical analysis results of water distribution data

配水池	項目	データ数	平均値	標準偏差	最小値	中位値	最大値
A 配水池	配水池水位 [m]	73.060	3.0	0.20	0.57	3.09	3.26
	配水流量 [m ³ /h]	73.060	167.1	29.0	0.2	171.6	279.3
	A1 流入流量 [m ³ /h]	73.060	11.6	8.6	0.3	12.4	36.6
	A2 流入流量 [m ³ /h]	73.060	113.8	11.9	0.2	111	140.6
	A3 流入流量 [m ³ /h]	73.060	60.3	1.7	0	60.7	67.9
B 配水池	配水池水位 [m]	72.915	4.8	0.30	0.11	4.81	5.28
	配水流量 [m ³ /h]	72.915	18.2	5.2	0.5	18.7	67.9
	B1 流入流量 [m ³ /h]	72.915	18.0	17.0	0	16.6	36.5



(a) : Monthly boxplot of distributed water flowrate
 (b) : Yearly boxplot of distributed water flowrate
 (c) : Yearly correlation of distributed water flowrate

Fig. 2 Statistical analysis results of water distribution data in Storage Tank A

値を検出したことが確認された。**Fig. 2 (b)** より、A 配水池の年別配水流量の中央値は、155.8 m³/h から 187.7 m³/h まで変動し、平均値 173.9 m³/h、標準偏差 10.4 m³/h、変動係数 0.0595 であった。また、**Fig. 2 (c)** より、各年で年間配水流量の相関が強く、2022 年以外のデータは、全て各々の相関係数が 0.8 以上であった。これは、A 配水池の配水流量の年変動が小

さいことが示された。

B 配水池の解析結果について、**Fig. 3 (a)** より、2018 年の 7 月と 8 月及び 2022 年の 6 月と 7 月に、普段より高い配水流量となっている箇所が確認された。**Fig. 3 (b)** より、B 配水池の年別配水流量の中央値は、17.7 m³/h から 20.5 m³/h まで変動し、平均値 18.8 m³/h、標準偏差 0.73 m³/h、変動係数 0.039 で

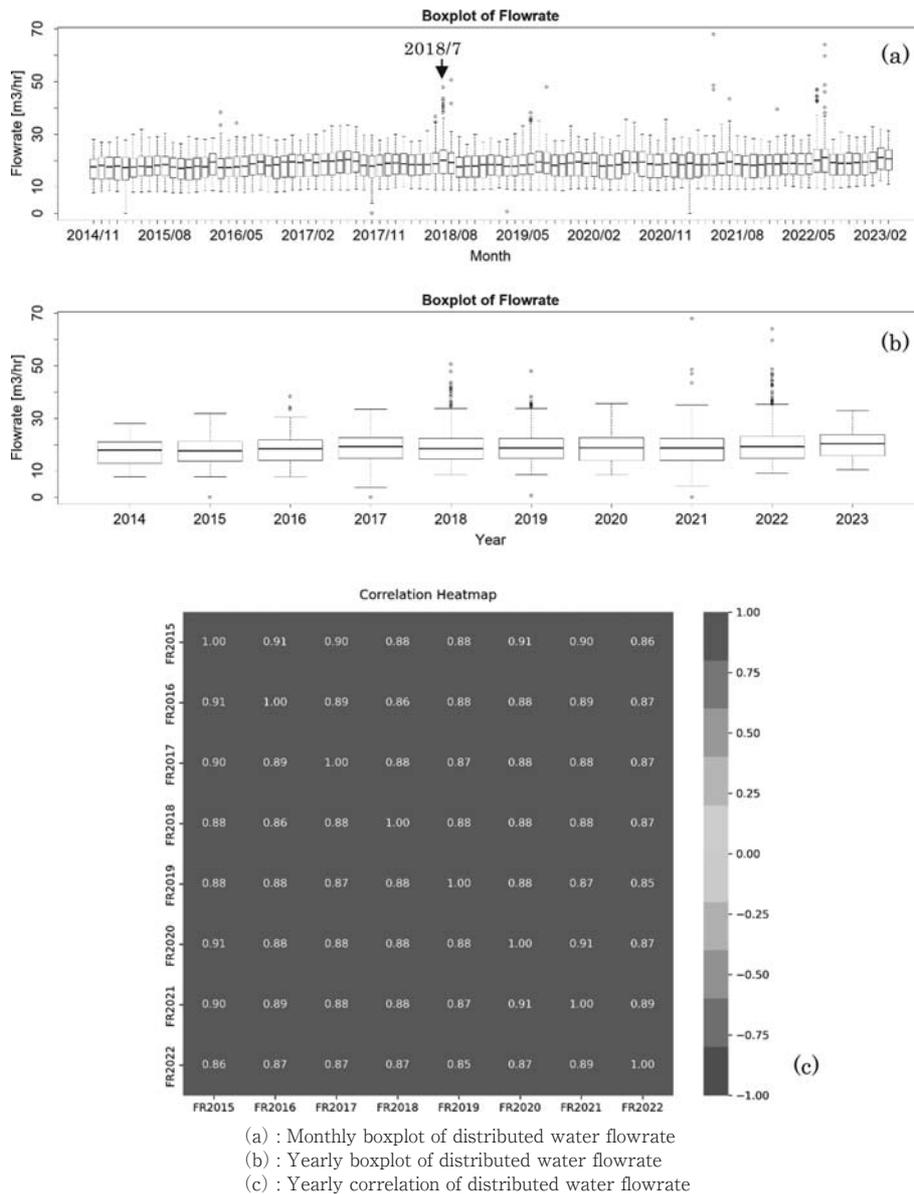


Fig. 3 Statistical analysis results of water distribution data in Storage Tank B

あった。また、Fig. 3 (c) より、各年で年間配水流量の相関が強く、全て各々の相関係数が0.8以上であった。これは、B配水池の配水流量の年変動が小さいことが示された。

このデータ解析結果より、いずれの配水池においても、異常流量があったが、年間配水流量の中央値の変動が小さい、かつ配水流量の年別相関が強かったと判明した。上記配水流量の年間変動や季節変動を考慮せずという前提を満足したため、配水時刻を横軸とした1日の変動の異常値を検討できると考えられる。

3.3 K市の配水データを用いた異常流量検知

K市の配水量データを用い、OCSVMを活用した漏水検知手法を検討した。

まず、Fig. 1と同様な手法で、日ごとの1時間間隔配水流量データは、時間を横軸、配水流量を縦軸とし、過去全ての配水データをプロットし、その分布より境

界線を決め、境界線範囲以外のデータを異常値として判断した。

検討結果をFig. 4に示す。Fig. 4では、赤い点線はOCSVMで決めた異常境界線である。白点はレギュラーデータ（通常の供水流量変動範囲内であるデータ）として判断し、赤点は異常値として検知した。検知結果は、青い背景色で示し、青色が濃いほど異常度が高い。Fig. 4 (a) より、11時の最大流量は異常値として検知できたが、12時～14時の最大値は境界線範囲外であっても通常流量として検知した。Fig. 4 (b) より、8時の最大流量は、9時と10時の流量データとの距離が小さいため、レギュラーデータとして判断したが、最大値より小さい40 m³/hの供水流量は、異常値として検出した。Fig. 4の検討結果より、下記2つの課題を洗い出した。

- ① 異なる時刻のデータの関連性：この手法は、データの分布より境界線を決める際に、データ

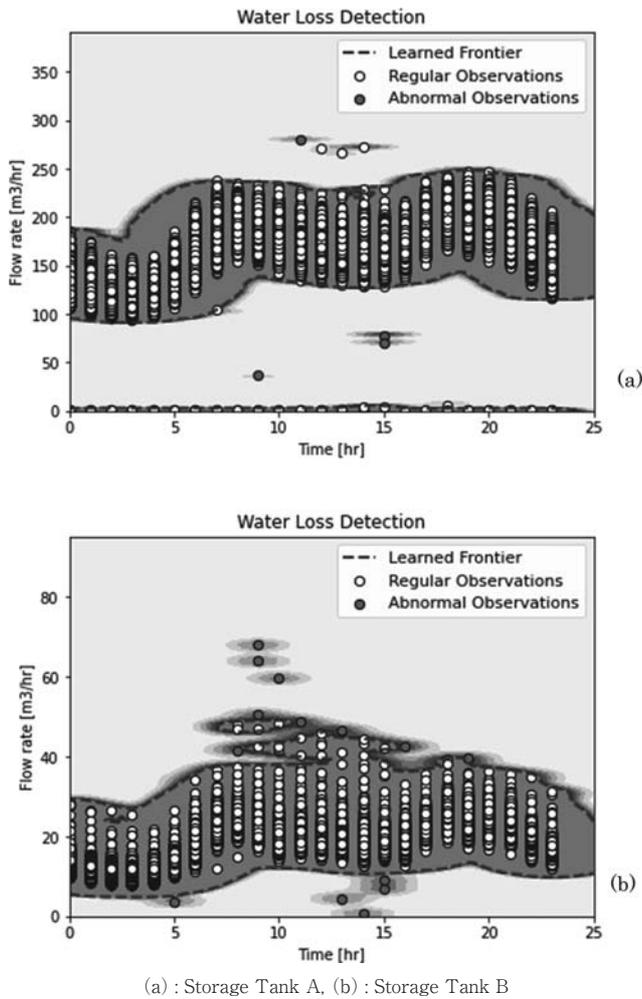


Fig. 4 Outlier detection results for distributed water flowrate without consideration of periodicity

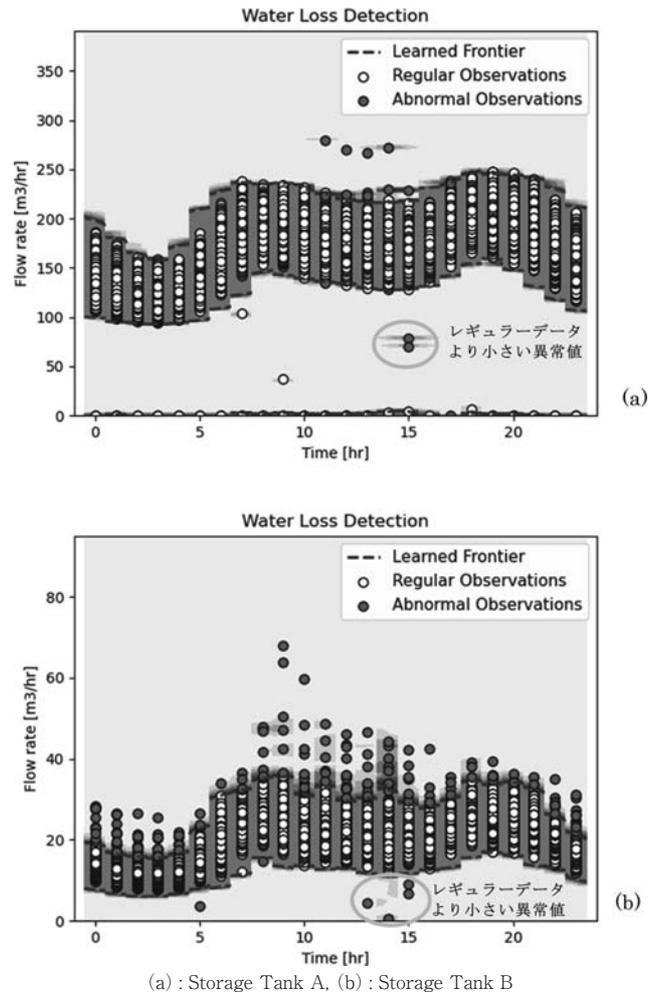


Fig. 5 Outlier detection results for distributed water flowrate considering periodicity

の間の距離を計算する。例えば、15時の流量データの異常値を判断する際に、15時のデータだけではなく、他の時刻の流量データとの距離も計算する。Fig. 4 (a) は、15時の最大流量は、異常流量である可能性があるが、13時や14時の最大流量との距離が小さいため、異常値ではないとして判断した。Fig. 4 (b) は、全ての配水データの分布より決めた境界線は正しくない可能性があり、特に8時から12時までの境界線は異常値のある区間は複数箇所があった。境界線を決める際に、関連性ある時刻のデータのみを利用する方がよい。

- ② 時系列データの周期性：時系列の配水流量データは、1日の周期性がある。例えば、Fig. 4の一番右の23時のデータは、一番左の0時のデータとは、時系列の連続データである。上記検討結果は、この周期性を考慮しなかった。

この2つの課題を解決するために、上記異常値検知手法を改善した。解決策として、対象時刻の異常値を検出する際に、全部データではなく、その前後の時刻

のデータと合わせて3時間のデータ（例えば0時の異常値検出は、23時・0時・1時のデータ）を利用し、OCSVMモデルを構築する。この手法は、各時刻に単独OCSVMを構築し、合わせて24個のOCSVMを構築した。各OCSVMの構築は、RBFカーネルを利用し、パラメータであったgammaはデフォルトの「Scale」で、nuは0.5とした。

改善した手法を用い、A配水池における配水量の異常値検知結果をFig. 5 (a)に、B配水池における配水量の異常値検知結果をFig. 5 (b)に示す。

Fig. 5 (a)とFig. 4 (a)の結果を比較すると、11時～15時の最大値は全部異常値として検出できた。この結果より、OCSVMによって配水流量の異常増大を検知できることが示唆された。また、Fig. 5 (b)より、連続3時間のデータで構築したOCSVMでは、他の時刻から受けた影響が少ないことが示された。例えば10時や12時の境界線上限は11時の上限より低かった。Fig. 4 (b)に示した8時から12時までの境界線が正しくないという課題も解決できた。

各時刻に配水流量の異常値検出数をTable 3にまとめた。A配水池は、73060個の配水流量データから

Table 3 Detected outlier amount for distributed water flowrate considering periodicity

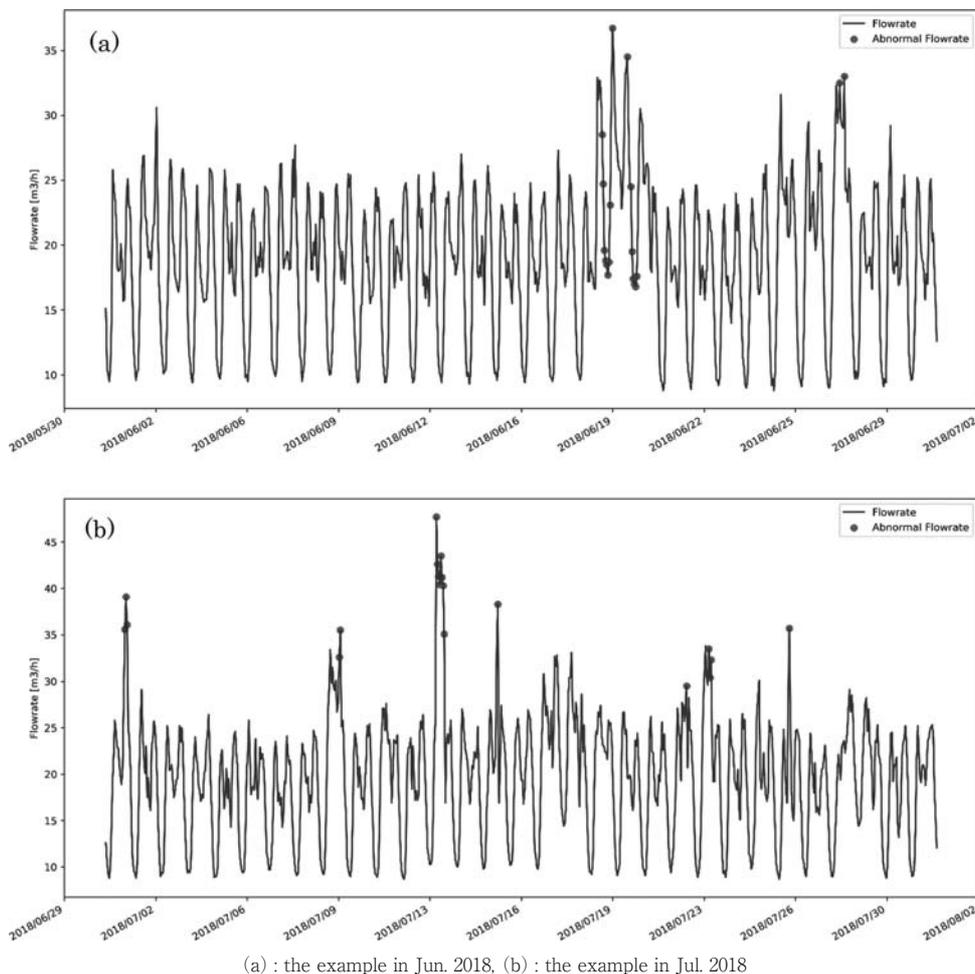
時刻	データ数	A 配水池 異常値数	割合	データ数	B 配水池 異常値数	割合
0	3,044	0	0%	3,038	12	0.395%
1	3,044	0	0%	3,038	13	0.428%
2	3,044	0	0%	3,038	10	0.329%
3	3,044	0	0%	3,038	10	0.329%
4	3,044	0	0%	3,038	8	0.263%
5	3,044	1	0.033%	3,038	5	0.165%
6	3,043	0	0%	3,036	1	0.033%
7	3,044	0	0%	3,037	5	0.165%
8	3,044	0	0%	3,037	8	0.263%
9	3,044	0	0%	3,038	6	0.197%
10	3,043	0	0%	3,038	10	0.329%
11	3,043	3	0.099%	3,038	6	0.197%
12	3,043	2	0.066%	3,038	13	0.428%
13	3,043	4	0.131%	3,038	9	0.296%
14	3,044	2	0.066%	3,037	8	0.263%
15	3,045	3	0.099%	3,038	9	0.296%
16	3,045	0	0%	3,039	7	0.230%
17	3,045	2	0.066%	3,039	4	0.132%
18	3,045	0	0%	3,039	3	0.099%
19	3,045	0	0%	3,039	5	0.165%
20	3,045	0	0%	3,039	3	0.099%
21	3,045	0	0%	3,039	2	0.066%
22	3,045	0	0%	3,039	7	0.230%
23	3,045	0	0%	3,039	10	0.329%

17 個の異常流量を検出した。B 配水池は、72915 個の配水流量データから 174 個の異常流量を検出した。A 配水池より、B 配水池に検出した異常流量数が多かった。

Fig. 5 より、同じ時刻のデータの数は多いため、異常値であるかを判断しにくい。配水流量の経時変化をプロットすると判断しやすい。例として、B 配水池 2018 年 6 月と 7 月の配水流量の経時変化（青線）と検出した異常値（赤丸）を Fig. 6 に示す。この結果より、通常の配水流量と比べて高くなっている点を異常値として検知できていることを確認した。

Fig. 2 (a) と Fig. 3 (a) に示した配水流量の月別の箱ひげ図より、月別データの異常値も判別できる。例えば、Fig. 2 (a) より 2016 年 1 月 A 配水池の配水流量の異常上昇は分かった。さらに、配水流量の日別の箱ひげ図より、一日の流量異常変動も検知できる。しかし、これらの検知は一日の流量変動を考慮せず、一日の最大流量の異常しか判別できない。本検討は、Fig. 6 (a) に示した 2018 年 6 月 19 日の一日の全体的な流量異常を検知でき、特に夜中の低流量の異常上昇も検知できた。

なお、この配水流量の異常値が必ずしも漏水であるとは言えないが、異常値検知時に漏水の可能性が高い

**Fig. 6** Variation of distributed water flowrate with time in Storage Tank B and outlier detection results

という警報を出し、配水先の他の流量データと合わせて確認したり、漏水の可能性が高い場所を現場確認したりすることで、漏水被害の最小化に有効な手法になると考えられる。

ただし、開発した手法では、いくつかの課題や改善点が挙げられる。

- ① A配水池より、B配水池で検出した異常流量数が多かったが、誤検出の可能性があり、異常値の判断基準を更に検討する必要がある。Fig. 5 (b) に示したB配水池の配水流量の異常値は、通常流量と近いデータが多かった。
- ② レギュラーデータより小さい異常値を誤検知していた。Fig. 5 (a) と Fig. 5 (b) のいずれも、OCSVMの特性より、通常流量より小さく、漏水ではないような点も異常値として検出されていた。これらの異常値は、漏水データではない。また、境界線下限より小さい異常値はレギュラーデータとして判断した結果もある。例えば、A配水池 (Fig. 5) 7時と9時の25~100 m³/hの流量データがこれにあたる。

また、本検討では配水流量のデータしか用いていない。配水池の水位や流入流量などのデータは漏水との相関性が低いと考えられるので、本検討では利用しなかった。一方、配水流量の異常値を検知できていない可能性もあるので、流量以外の情報と合わせて検出する手法を改善する方がよいと考えられる。文献考察より配水管内の圧力変化は漏水と強い相関があり、今後配水圧力データを収集し、流量と合わせて検討していく。

4. ま と め

One-Class SVMという外れ値検出用アルゴリズムを検討し、K市の配水流量データに対して、漏水検知への活用を試した。

異なる時刻の配水流量データの関連性や時系列データの周期性を考慮し、時刻毎のOne-Class SVMモデルを構築することで、2つの配水池とも、通常流量より異常に高い流量を異常値として検知でき、今回検討

したOne-Class SVMは、配水流量異常の検知に有効な手法であることが示唆された。時刻毎のOne-Class SVMモデルの構築より、一日の全体的な流量異常を検知でき、特に夜中の低流量の異常上昇も検知できた。なお、今回の配水流量の異常増大は、必ずしも漏水であるとは言えないが、漏水の可能性が高いという警報を出すことで漏水被害の最小化につながる有効な手法であると考えられる。

課題や改善点としては、①漏水ではない場合の配水流量上昇を誤検出すること、②誤検知・レギュラーデータより小さい異常値が検知されたこと、が挙げられる。この改善に向けて引き続き検討していく。また、実際に把握した漏水情報より、検出した配水流量の異常増大と実際の漏水との関係を検討し、開発した漏水検知手法の有効性を確認していく。

参 考 文 献

- 1) Water Loss, <https://www.interreg-central.eu/Content.Node/Digital-Learning-Resources/03-Water-Loss.pdf>
- 2) Water Ranking データで見る水の世界, <http://www.waterworks.jp/vol25/page4.html>
- 3) Vanijirattikhan R, Khomsay, S, et al: AI-based acoustic leak detection in water distribution systems, Results in Engineering, v15, p. 100557, 2022, <https://doi.org/10.1016/j.rinen.g.2022.100557>
- 4) 島田孟親, 荒井康裕, 南泳旭, 國實誉治, 小泉明: 水道管路に設置する漏水検知センサーの必要数に関する一考察, 令和2年度水道研究発表会講演集, pp. 464-465
- 5) 島田孟親, 伊藤開登, 荒井康裕, 國實誉治, 小泉明: ディープラーニングを用いた漏水検知モデル (I) —— モデルに使用する画像サイズに関する分析, 令和3年度水道研究発表会講演集, pp. 470-471
- 6) Muh. Anshari Caronge, 伊藤開登, 荒井康裕, 國實誉治, 小泉明, B. Bakri: 実漏水音を用いた漏水検知モデルの検討, 令和4年度水道研究発表会講演集, pp. 792-793
- 7) 水道管の漏水, 人工衛星から発見, 期間も費用も大幅減, 読売新聞, 2021, <https://www.yomiuri.co.jp/science/20210917-OYT1T50106/>
- 8) Schököpf B., Platt J.C., Shawe-Taylor J., Smola A.J., and, Williamson R.C.: Estimating the Support of a High-Dimensional Distribution, NC, 2001

Detection of Water Loss by AI with Anomaly Detection Algorithm

Pengzhe Sui^{1)†}, Kazuaki Shimamura¹⁾, Keitaro Tada²⁾ and Jumpei Sumi²⁾

¹⁾ Research and Development Center, Swing Corporation

²⁾ Integrated Water Business Div., Swing Corporation

† Correspondence should be addressed to Pengzhe Sui:

Research and Development Center, Swing Corporation

E-mail: sui.pengzhe@swing-w.com

Abstract

Water loss in water supply systems is a global issue, causing regional water outages directly as well as the indirect damage such as road subsidence. With the rapid development of artificial intelligence (AI) in recent years, an anomaly detection method with AI was developed in this study to detect water loss in water supply system in order to respond quickly according to the alarm of AI when water leakage happened. Using the flowrate data in water distribution system of K City and considering the connectivity and daily periodicity of time series flowrate data, an algorithm called One-Class SVM was adopted to detect the abnormal flowrates in every hour. It was confirmed that One-Class SVM was useful for the detection of abnormally high flowrates, including the abnormal flowrates in one whole day, especially the relatively low abnormal flowrates in the night. According to the results, although there were still some issues, such as wrong detection result of high flowrates when water loss did not occur, the developed approach could be effective to minimize the damage caused by the water leakage.

Key words: anomaly detection, water loss, artificial intelligence