

## 主成分分析を用いたオンラインデータの異常値検出システム

平林和也\*、藤田逸朗\*、金谷利憲\*  
津村和志\*\*

\* (株)安川電機基礎研究所基盤技術研究室  
北九州市八幡西区黒崎城石2-1

\*\* 京都大学工学部環境地球工学教室  
京都市左京区吉田本町

### 概要

下水処理プロセスにおける水質計測値の信頼性を向上させることを目的として、オンラインデータの異常値検出の研究を進めている。計測値は、処理場、計測器、季節によって特性が異なる。そこで、さまざまな状況に対応できるように、種々の角度から計測値の妥当性をチェックし、その結果を総合的に判定する方法を開発した。本方法では、まず、上下限值からの逸脱、変化量過大、変化量過小、予測誤差、類似計測器、トレンドによる6種の方法によって計測値の妥当性をチェックした。次に、総合判定では、主成分分析手法を用いて6種の異常検出ロジックから出力された値の統合化を行った。本報告では、検出システムの概要と実処理場で収集したデータへの適用例を述べている。

### キーワード

異常値検出、主成分分析、オンラインデータ

#### 1. はじめに

下水処理場では、近年のセンサー技術の進歩により、多くの自動計測器が導入されている。しかし、検出部の汚れ、高湿度雰囲気、腐食性ガスなど、下水プロセス特有の悪環境があるため、長期にわたる計測は難しい状況にある。このことが、下水プロセスの自動制御を困難にしている要因の一つと考えられる。

そこで筆者らは、下水処理場における計測値の信頼性を向上させることを目的として、オンラインデータの異常値検出の研究を行っている。すなわち、種々の角度(上下限值、変化量過大、変化量過小、他の計測器との関係、類似計測器との照合、トレンドなど)から計測値の妥当性をチェックし、その結果を総合的に判定するシステムの開発である。本報告では、この異常値検出システムの概要と実処理場で収集した4種の水質計測値に適用した例について述べる。

#### 2. 異常値検出システム

このシステムの構成は、大きく二つのブロックに分けられる。その一つは、計測値を様々な角度からチェックする異常値検出ロジック、他の一つは、検出ロジックの結果を総合的に判定する総合判定方法である。また異常値検出システムで問題となるのは、処理場、計測器、季節変動などの特性に自動的に対応できる判定基準値の更新方法である。ここでは、判定基準値の自動更新、異常値検出ロジック、総合判定について説明する。

2-1. 判定基準値の自動更新<sup>(1)</sup>

理想的には、その処理場を理解している運転員が判定基準値を逐次設定することが望ましい。しかし、計測器が多数あること、及び季節変動に対応しなければならないことから、運転員が定期的に判定基準値を更新することは困難である。そこで、計算機によって自動的に判定基準値を更新する方法を開発した。通常、計算機による判定基準値の設定は、過去の計測値から標準偏差を求め、この値を基準に判定基準値を設定することが多い。しかし、下水処理プロセスで取り扱う計測値の頻度分布は必ずしも正規分布を示していないことが多い。そこで、正規分布を示さないデータ群の取り扱い方法として提案されている探索的データ解析法を用いた。この方法は、四分位数を基準に判定基準値を設定する。四分位数とは、相対累積度数の25%点、50%点、75%点のことであり、この距離を基準に判定基準値を設定した。この方法により、処理場、計測器、季節によって頻度分布の異なる様々なデータに対応できる判定基準値の設定が可能となる。また、計測値は季節によって変化するため、判定基準値の更新は、直近30日のデータを使用して1日1回行うことにした。

2-2. 異常値検出ロジック

異常値検出システムにおいて、様々な角度から計測値をチェックするため、検出ロジックは、図1に示すように6種の方法を考えた。このシステムは、検出ロジックの追加、変更及びプログラムの変更が頻繁に起こることが予想されるため、追加、変更が容易になるように検出ロジックを独立に構築した。各検出ロジックのチェック方法を記すと次のようになる。

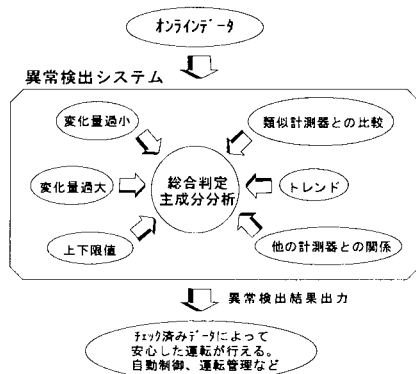


図1 異常値検出システムの構成

a) 上下限值による異常値検出<sup>(1)</sup>

計測値の判定基準値を設定し、計測値がこの値を越えた時異常とする。

b) 変化量過大による異常値検出<sup>(1)</sup>

変化量過大の判定基準値を設定し、変化量がこの値を越えた時に異常とする。

c) 変化量過小による異常値検出

変化量過小の判定基準値を設定し、変化量はその判定基準値内に連続して数点入った時異常とする。

d) 他の計測器の影響を考慮にいった異常値検出(予測誤差)<sup>(1)</sup>

予測誤差の判定基準値を設定し、自己回帰モデルを使用して、検出対象の計測器に影響を及ぼす要因を考慮に入れた予測を行い、その予測誤差が判定基準値を越えた時異常とする。

e) 類似計測器比較による異常値検出<sup>(2)</sup>

類似計測器間の判定基準値を設定し、類似計測器の計測値の差がこの値を越えた時異常とする。

f) トレンドによる異常値検出<sup>(2)</sup>

1日のデータの平均値の変化量が、数日間連続して正または負を示した時異常とする。

各検出ロジックの適用例をMLSS濃度計、溶存酸素濃度計(以下DO計と略記する)、pH計、紫外線可視吸光度差計(以下UV計と略記する)を例にあげて図2に示す。これらの計測器を検証に用いたのは、水質計測器の中で、応答の早い計測器、遅い計測器など様々な計測器に、このシステムが対応できるかどうかチェックするためである。サンプリング周期は、すべて60分で行っている。図より、各検出ロジックで検出された点は、それぞれ異なった結果を出力していることが分かる。例えば、図2-(1)のa点では、変化量過大は異常、上下限值、予測誤差、類似計測器の照合などは正常と出力されている。また図2-(2)

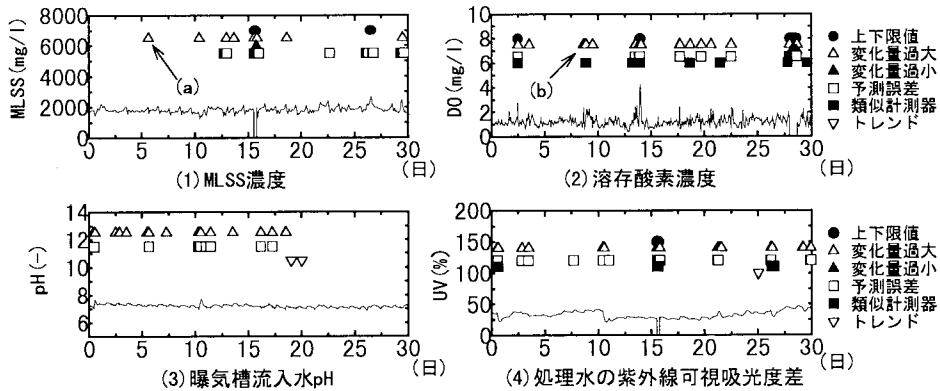


図2 各検出ロジックによる異常値検出結果

のb点では、変化量過大、類似計測器の照合は異常、上下限值、予測誤差などは正常と、異なった結果を出力している。このように6種の検出ロジックはそれぞれ違う角度からデータをチェックしているため、各々、異なった結果が出力される。したがって、これらの結果を計算機が総合的に判定することが要求される。

### 2-3. 総合判定

各検出ロジックから得られた結果を、総合的に評価するために主成分分析<sup>(4)</sup>を適用することを試みた。ただし、6種の検出ロジックの中で、変化量過小による異常検出は計測器の故障などに原因があると考えられ、トレンドによる異常検出は1日1回の異常値検出のため、総合判定を行わず計測値異常として出力することにした。そして、その他の4つの判定結果を対象とした主成分分析を行い総合判定を試みた。<sup>(3)</sup>

主成分分析に使用した変数は、上下限值、変化量過大、予測誤差、類似計測器の照合である。この4つの変数から総合指標である主成分得点を求め、この値を異常値検出の対象データとした。

## 3. 異常値検出システムの検討

### 3-1. 実データへの適用例

総合判定の検証に適用したデータは、図2と同じMLSS濃度計、DO計、pH計、UV計である。図3には、各計測器の主成分得点と判定基準値及びオリジナルデータと異常検出点(●点)を記している。各計測器の異常検出結果を見てみると、応答の早い計測器と遅い計測器いずれについても、少し過敏に計測値異常を検出しているようである。次に、pH計のa点(1日、10日頃)、UV計のb点(1日、10日、22日、27日頃)などは、異常に近い動きを示しているように見える。運転員にとって、日常の運転の中でこのような変動を発見することは困難と思われるが、このシステムではこれらのデータを異常値として検出している。この4種の計測器においては、このように異常と思われる点をほとんど検出している。この結果、主成分分析による総合判定は、各検出ロジックで計算された結果を主成分得点として1つにまとめまた、様々な要因がからみあった計測値の動きをうまく説明しているといえる。

本システムの検出結果は、全体的に過敏な検出をしているので、主成分得点の値と判定基準値の関係について見てみた。例えば、MLSS濃度計では26日頃、DO計では9日頃に検出された異常は、判定基準値をわずかに越えたものである。これに対してMLSS濃度計の16日頃、DO計の3日、14日、18日頃に検出された異常は、判定基準値を大きく逸脱している。従って実用的には、異常度を大、中、小とレベル分けすることなどが考えられる。異常度のレベル分けは、プロセスによって異なるため、各下水処理プロセスに対応した設定をすることになると考えられる。

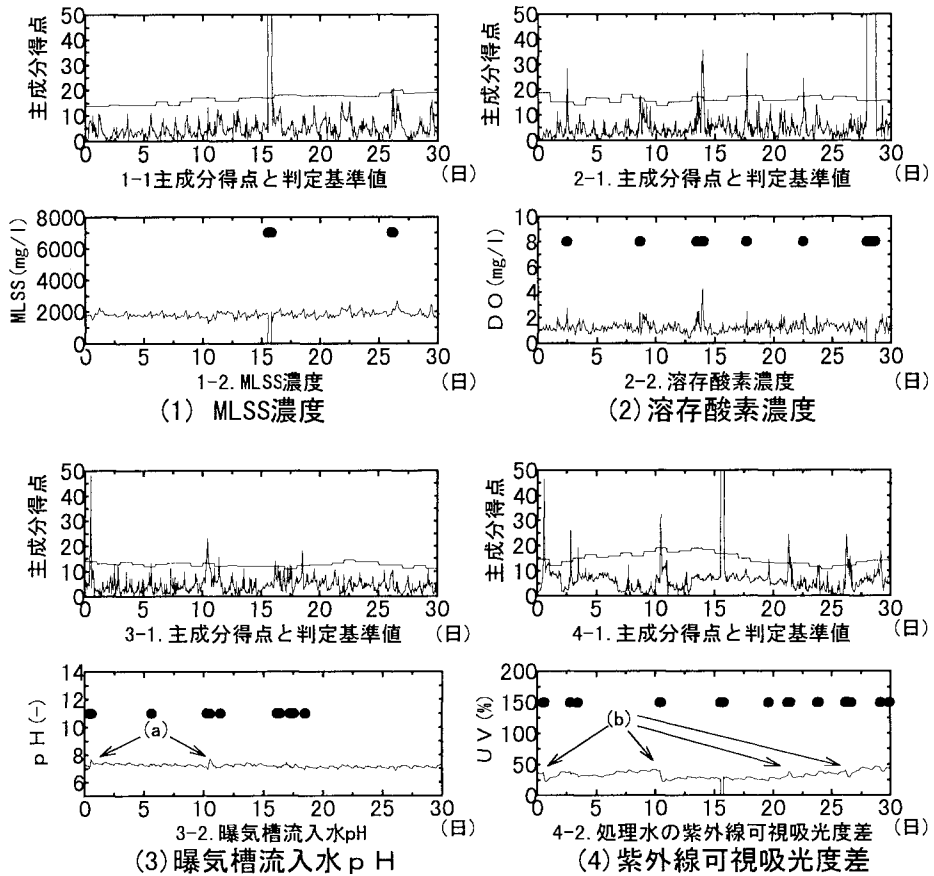


図3 各計測器の異常検出結果

#### 4. 成果と今後の課題

計測値の妥当性をチェックするために、様々な角度からデータをチェックし、各検出ロジックから得られた結果を総合的に判定する手法を構築することができた。今後は、実処理場においてテストを行い、このシステムの適応性、異常度のレベル分けについて検討し異常値検出の精度を向上させていく予定である

また、リン・窒素など今まで計測不可能だった項目が、自動計測されるようになってきている。このように今後は、プロセスをより良く運転することを目指して、新しい計測器の導入、計測ポイントの増大が行われることが予想される。その意味から、今回開発した異常値検出システムは、プロセスの計装において重要な意義を持つと考えている。

#### 【参考文献】

- (1)平林、金谷、津村：水質計測器の異常検出技術の研究、第31回下水道研究発表会(1994)
- (2)平林、藤田、金谷、津村：下水処理場におけるオンラインデータの異常検出、第32回下水道研究発表会(1995)
- (3)平林、藤田、金谷、津村：主成分分析を用いたオンラインデータの異常検出システム、第33回下水道研究発表会(1996)
- (4)T. Kanaya, K. Hirabayashi, I. Fujita, K. Tsumura : Detection of Unusual Data in on-line Monitoring of Wastewater Processing, Water Science and Technology, 1996 vol.33 No.1
- (5)田中、垂水、脇本：パソコン統計解析ハンドブック II 多変量解析編、1984、共立出版