

〈研究発表〉

ビッグデータ解析を適用した薬品注入ガイダンスシステム

平 林 和 也¹⁾, 占 部 正 敏²⁾, 大 場 正 隆³⁾

¹⁾ (株)安川電機 システムエンジニアリング事業部 社会システム技術部
(〒824-8511 福岡県行橋市西宮市2-13-1 E-mail: Kazuya.Hirabayashi@yaskawa.co.jp)

²⁾ (株)安川電機 システムエンジニアリング事業部 社会システム技術部
(〒824-8511 福岡県行橋市西宮市2-13-1 E-mail: Masatoshi.Urabe@yaskawa.co.jp)

³⁾ (株)安川電機 システムエンジニアリング事業部 社会システム技術部
(〒824-8511 福岡県行橋市西宮市2-13-1 E-mail: Masataka.Oba@yaskawa.co.jp)

概 要

凝集や滅菌処理に使用する薬品（PAC、硫酸バンド、次亜塩素）の注入率をガイダンスするシステムを構築した。このシステムは収集した計測データからビッグデータ解析でモデルを構築し、モデルと直近のデータから現時点の注入率をガイダンスすることができる。今回は、対象項目として、次亜塩素酸を対象として、実プラントで薬注ガイダンスを行った結果を報告する。

キーワード：浄水、薬品注入、ビッグデータ
原稿受付 2018.6.29

EICA: 23(2・3) 16-20

1. はじめに

浄水処理には、急速ろ過、緩速ろ過、膜ろ過など、種々の方法がある。急速ろ過では原水が低濁度であっても、ろ過するだけでは、コロイド懸濁度物質の十分な除去は難しく、前処理として薬品による凝集が不可欠である。また、緩速ろ過方式においても、原水の濁度が30度以上になると、沈殿後の濁度が高くなり、ろ過閉塞するために凝集沈殿によって濁度を下げおく必要がある。

このように原水を浄化するためには、ろ過の前に、凝集、沈殿させる必要があるが、この凝集剤の注入率は、ベテラン運転員によって、これまでの経験を基に設定されていることが多い。特に豪雨などの影響によって原水水質の変動が激しい時期では、薬剤注入量の調整は難しく、ベテラン運用員の経験やジャーテストを行い凝集剤の注入率を決定している。

少子高齢化に伴う技術者不足が予想される中、上水道施設の安定的な運転を継続していくには、ベテラン運転員の運転ノウハウを蓄積し、継承する必要がある。それを達成する方法の一つとして、データ解析技術を活用したガイダンスが有効と考えた。それらをサーバにてモデルを作成し、そのモデルから対象とする項目のガイダンス値を管理用PCに表示することでノウハウの継承と上下水道施設の安定的な維持管理が可能となる。ビッグデータ分析を取り入れた薬注ガイダンスシステムを構築し、実設備で検証を行っている。本論文では、解析手法および実証結果を報告する。

2. 対象浄水場フロー

対象の浄水場のフロー図を **Fig. 1** に示す。施設能力日量 19,700 m³ である。河川から取水し、生物処理池を経て、PAC、硫酸バンドを注入して、凝集沈殿

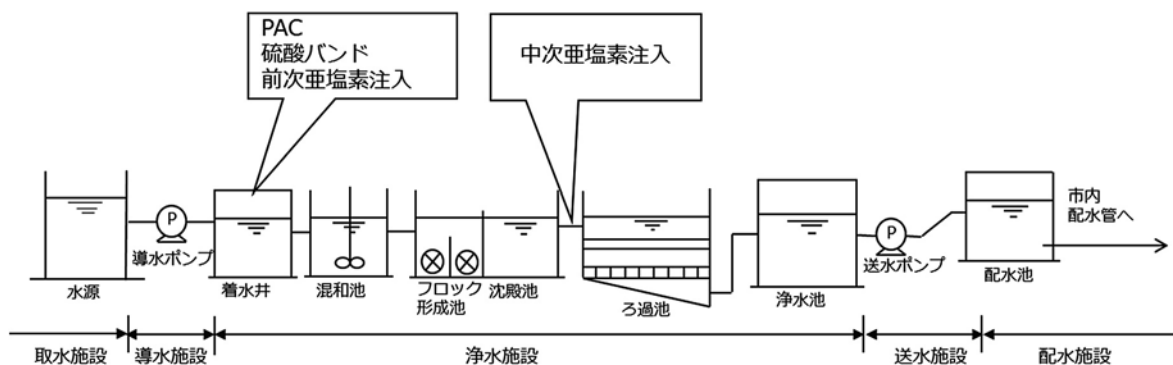


Fig. 1 Purification Plant

している。その後、沈殿池を経て、塩素を注入し、トリハロメタンやカビ臭生成の低減を図っている。急速ろ過を経て、浄水池へと送水し、配水池から各家庭へ水を供給している。

対象の浄水場は、河川より取水しているため、雨天時には、取水濁度が大幅に上昇することがあり、凝集剤の注入量は、取水濁度、温度、これまでの経験などを総合的に判断し、運転員が決定している。

また、次亜塩素酸注入率も残留塩素濃度などを監視しながら、総合的に注入率を決定している。

3. 実証システム構成

実証設備として、クラウドでガイダンスが可能な環境を構築した。その理由として、ガイダンス結果に乖離ある時にリモートでリアルタイムに確認可能なこと、CPU やハードディスクの故障時にバックアップがあり計算用パソコンの管理が必要ないことなどがあげられる。その結果、ガイダンスの実証を安定的に進めることができた。

システムをクラウドで構築した時の流れを Fig. 2 にて説明する。まず、浄水場では PLC からデータ収集を行い、M2M を経由して、クラウド上のサーバへデータを送信する。クラウドでは、M2M より送信された、トレンドなど多数の項目のデータを受信し、過去データを蓄積するデータベースへ保存する。次にそのデータベースから次亜塩素酸、凝集剤などをガイダンス（予測）するために、寄与度の高い項目を抽出してモデルを構築する。

そのモデルと直近のデータから次亜塩素酸や凝集剤などの項目について予測値を計算する。計算した結果は、インターネット経由で浄水場のタブレットで表示する。その予測結果を確認し、オペレータが注入率の設定を行う。

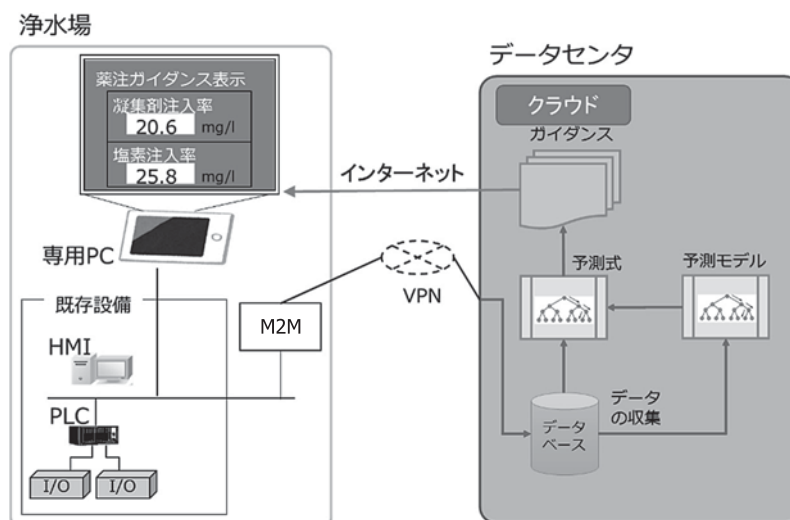


Fig. 2 System Configuration

4. 解析手法

4.1 ランダムフォレスト

この薬注ガイダンスに使用した予測モデルは、ランダムフォレスト¹⁾決定木法を応用した。決定木とは、親から順に条件分岐を辿って結果を得る手法であり、ランダムフォレストはモデル作成用データをランダムに選択し、決定木を複数組み合わせ、各決定木の予測結果を平均する方法である。

ランダムフォレストの特長は3つあり、項目ごとの重要度が確認できること、モデルのパラメータが少ないこと、そして非線形関係も分析できることである。Fig. 3 のようにツリーを構築し、これまでベテラン運転員が判断してきた制御設定値の判断結果をツリーのようなモデルで表現することができるため、ノウハウを可視化することが可能である。

4.2 データ処理方法

今回適用したデータ解析は、数十種類の過去データから解析を行い、目的変数に対する重要度の高い説明変数を抽出し、説明変数から目的変数の状態を予測する。現場からヒアリングした浄水工程における滞留時間や、薬剤効果の時間を考慮した前後の時系列データ

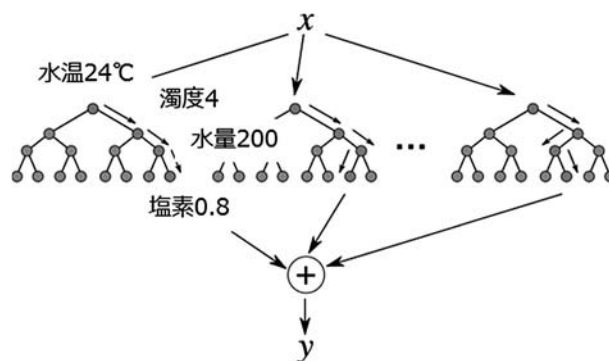


Fig. 3 Schematic of Random Forest

など現場独特のデータを追加して学習させた。各計測器間の相関や目的変数に対する説明変数の寄与度等は、コンピュータが自動学習するので、導入時の設定が容易となる。このように寄与率から抽出した項目の過去データを適用し複数の決定木をモデル化した。

4.3 予測計算

Fig. 3 から予測の方法について説明を行う。モデル作成時には寄与率の高い項目から複数の決定木が作成されている。現時点で計測した水温、濁度、水量などから決定木の条件に一致する分岐を通り、注入する次亜塩素酸注入率の値を求める。モデル形成時に計算した複数の決定木モデルにおいて同様の計算を行い、各決定木の次亜塩素酸注入率を求める。その決定木の予測結果を平均して予測を行う。

5. 説明変数の寄与度評価

浄水場で計測しているデータからランダムフォレストによる計算を行い、目的変数に対する寄与度の高い項目を抽出した。次亜塩素酸注入率に対する寄与度の高い項目中から 10 項目を Table 1 に示す。

残留塩素濃度は次亜塩素酸注入率を決定する主な指標となっているため、対象の次亜注入率以外では最も重要度があると計算された。その他の項目では、寄与率の値に大きな差が無く、配水池水位、中間処理濁度、ろ過池損失水頭などが抽出された。これは、オペレータが次亜塩素酸の注入率を決定するにあたり、直接影響する残留塩素濃度だけでなく、住民に安心・安全な水を供給するために必要な量や水の汚れの状態など注入率の決定に間接的に影響を与える複数の項目も考慮に入れ運転していることが推測される。

Table 1

	項目	寄与度
1	次亜注入率 (設定値)	22.762
2	中間処理残留塩素	8.471
3	次亜注入量	7.994
4	C 配水池水位	7.628
5	ろ過池損失水頭	6.512
6	高区配水量	6.390
7	中間処理濁度	5.198
8	次亜塩素貯留槽液位	5.069
9	A 配水池水位	5.007
10	ろ過池流入量	4.736

6. 次亜塩素酸の予測結果

次に予測結果についての検討を行った。対象の浄水場は河川より取水しているため、原水濁度は晴天時と雨天時では大きく異なることが多い。そこで、予測結

果の評価では、原水濁度変化の少ない晴天時および変化の大きな雨天時を抽出して行った。評価のためのデータは、雨量計測後、3日間濁度に影響が表れると想定して、晴天時と降雨後3日間を抽出した。評価は、原水濁度や中次亜注入率の値から時系列で傾向を確認する定性的評価とオペレータが次亜塩素酸注入率の設定値を変化した点を抽出した定量的評価を行った。

6.1 時系列の定性的評価

(1) 晴天時

ガイダンス結果を Fig. 4 に示す。グラフでは、原水濁度、中次亜注入率の設定値と予測値を表示した。晴天時のため川の濁度の上昇が見られなかったことから注入率も安定した値を設定している。量や濁度以外の水質に影響を受けた設定を行い運転していることが考えられる。予測値は設定値と同様に変化しない予測を行っている。

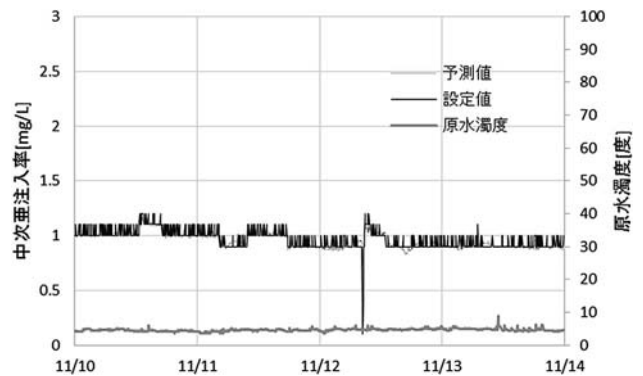


Fig. 4 Prediction result (clear weather)

(2) 雨天時 (日合計 43 mm, 1 時間最大 8 mm)

24 時間雨量が 43 mm の予測結果を Fig. 5 に示す。グラフでは、原水濁度、中次亜注入率の設定値と予測値を表示した。原水濁度が 40 度程度の上昇が見られ、徐々に減少している。濁度の上昇とともに次亜塩素酸の注入率も徐々に上昇した値を設定し、予測値も追従している。濁度が減少傾向を示した後においても、中次亜注入率は高い値の設定を継続しており、雨量計測後 3 日目から次亜塩素酸中注入率を低下した運転に変わ

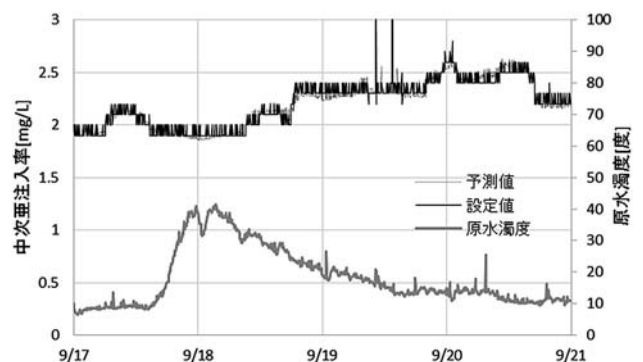


Fig. 5 Prediction result (43 mm for 24 hour rainfall)

りつつある。このように雨天後、濁度が下がった場合においても、注入率の維持や低減するタイミングについても予測値が設定値に追従していることが分かる。

(3) 雨天時(日合計75mm, 1時間最大20.5mm)

24時間雨量が75mmを計測された時の予測結果をFig.6に示す。グラフでは、原水濁度、中次亜注入率の設定値と予測値を表示した。原水濁度が80度程度まで上昇が見られ、徐々に減少している。濁度が40度の時と同様に濁度の上昇と共に中次亜注入率の設定も上昇させ、濁度が減少傾向を示した後も、高い設定値を維持している。雨量75mmでも43mmの時と同様に、雨量計側の3日後から中次亜塩素酸注入率を低下した運転に変わりつつある。24時間雨量75mmというような大量の雨量が計測されていても注入率を維持、低減するタイミングについて予測値が設定値に追従できていることが分かる。

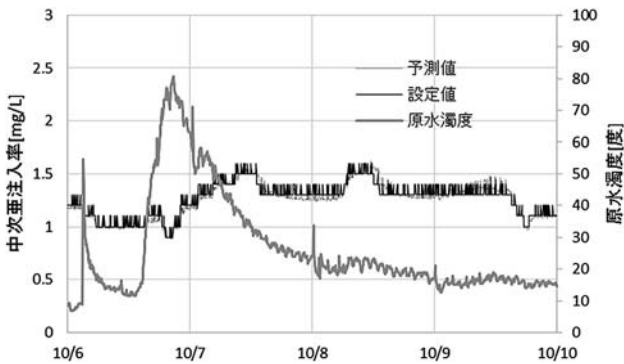


Fig.6 Prediction result (75 mm for 24 hour rainfall)

6.2 設定値変更時評価

時系列での評価はグラフ化して傾向を確認するには向いているが、予測誤差の計算では評価に必要な性が低い安定した期間も取り入れた計算を行うため、本システムの評価は難しいと考えた。つまり、晴天時など設定値の変更が少ないデータおよび雨天時の変化の大きなデータを同じ予測誤差として評価を行い、設定値の変更が少ないデータ数が多くなるため、予測誤差の計算に大きな影響を与える。予測誤差の評価では本来必要な設定値変更時の誤差を適正に評価するために他の手法の検討を行った。

オペレータが設定値を変化させたデータで評価を行うためには、設定値を変更した時の値を抽出し、その設定値と予測値の誤差を評価することにした。全期間、晴天時、雨天時の3パターンで予測誤差について計算を行った。

予測誤差は式1にて演算した。

$$Error = \frac{X_p - X}{X} \dots \dots \dots (1)$$

Error: 予測誤差

Xp: 予測値

X: 設定値

(1) 全期間の評価

晴天時と雨天時の全データを対象に、次亜塩素酸注入率設定値を変化した時の設定値と予測値を抽出し、その結果をFig.7に示す。式1で計算した結果、予測誤差が8.4%となった。設定値の範囲が1.0~3.0mg/Lで設定は0.1mg/L単位で変更することが多いため、予測誤差が10%以下は要求仕様を満足していると考えられる。

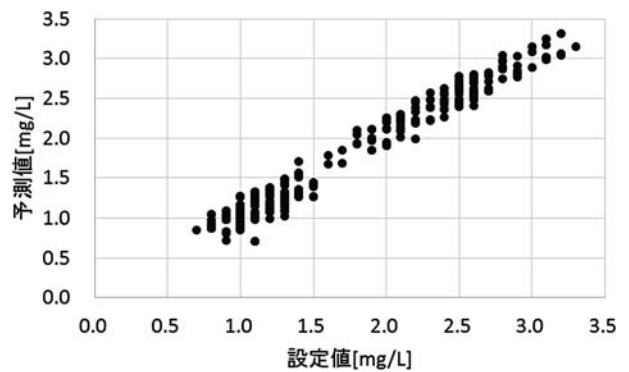


Fig.7 Set value and predicted value when set value is changed (all period)

(2) 雨天時の評価

雨天時の全データを対象に、次亜塩素酸注入率設定値を変化した時の設定値と予測値を抽出し、その結果をFig.8に示す。式1で計算した結果、予測誤差が8.7%となった。全期間の予測後より0.3%低下したが、設定値(1.0~3.0mg/L)や変更値(0.1mg/L単位)から全期間の条件と同様に予測誤差が10%以下は要求仕様を満足していると考えられる。

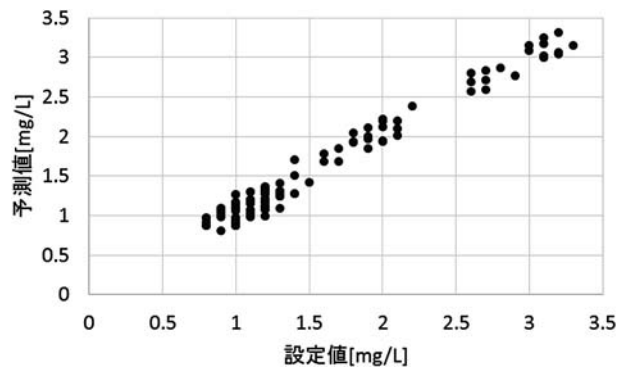


Fig.8 Set value and predicted value when set value is changed (rainy weather)

(3) 晴天時の評価

晴天時の全データを対象に、次亜塩素酸注入率設定値を変化した時の設定値と予測値を抽出し、その結果を **Fig. 9** に示す。予測誤差が 8.2% となった。全期間の条件と同様に誤差が 10% 以下は要求仕様を満足していると考えられる。

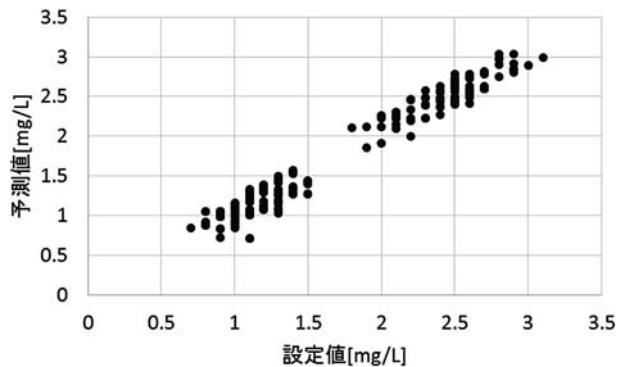


Fig. 9 Set value and predicted value when set value is changed (clear weather)

8. ま と め

薬注ガイダンスを構築し、実プラントで2年程度検証を行った。浄水場にコントローラや通信機器を設置し、オペレータが設定変更する値に注目し厳しい条件で検証を行った結果、晴天時 8.2%、雨天時 8.7% となり、設定値の範囲が 1.0~3.0 mg/L で設定し、0.1 mg/L 単位で変更することが多いため、運用上対応可能なレベルと考えている。

実運用に向けて、データ異常値の処理など運用に強いシステムを構築していきたいと考えている。

参考文献

- 1) Breiman, Leo. "Random forests." Machine learning 45.1 (2001): 5 - 32.