

ニューラルネットワークによる水道原水の水質予測

齊藤方正、水町昌代、山崎良明

大阪府水道部水質管理センター
大阪府枚方市村野高見台7-2

概要

浄水場における水質計測技術の調査・研究を目的として、ニューラルネットワークモデルを用いた水道原水の水質予測技法を考案した。

当技法は、流域内にある多地点の降雨強度や河川流量などを入力として、原水の過マンガン酸カリウム消費量や濁度を日単位で予測し出力するものである。モデルは、入力層、中間層、出力層の3層構造で構成され、バックプロパゲーション（誤差逆伝播法）により原水水質を出力する。当技法により、高精度の原水水質予測とともに、水源での観測情報の取捨選択等が可能となる。

キーワード

原水水質予測、流域雨量、ニューラルネットワーク、バックプロパゲーション

1 はじめに

浄水場での水処理操作は、主に時々刻々変化する水質計器の情報や操作員の経験をもとに行われている。現在、都市の浄水場では水源汚濁に対処するためにオゾン処理や活性炭処理などの導入を進めているが、このような高度処理の導入や水源汚濁の進行によってプロセス運用の情報量が増えると、水処理操作が複雑・煩雑化し、一層の柔軟性と精緻性が要求されてくる可能性がある。水処理操作を軽減するとともに、安定な水処理を今後とも維持するためには、不確実性や曖昧性の伴う水源情報や原水水質情報を、例えば人工知能（AI）技術等を利用して高度に処理する支援ツールの開発が望ましい。

これまで水処理制御へのAI技術の開発研究は、主に浄水場内の水処理に限定されてきた。水源情報への適用事例については、操作員の得る水源情報が質的にも量的にも限られ、水源情報と水処理を結び付ける経験や知識の蓄積が不十分であることから、あまり見受けられないのが現状である。

筆者らは、浄水場における新しい計測・制御技術としての水源情報処理ツールの開発を念頭に、AI技術の一つであるニューラルネットワークを応用した水道原水の水質予測手法を考案し、過マンガン酸カリウム消費量や濁度について、個別に適用した。

2 原水水質予測のためのニューラルネットワークモデル

AI技術は人間の意識的推論、無意識的判断を計算機に代行させようとするものであり、具体形態としてエキスパートシステム（ES）、ファジィ推論、ニューラルネットワークなどがある。ESの研究・開発事例としては、フレーム・ルール型の知識表現を用いた受変電設備の故障復旧ガイダンスシステム¹⁾や複数送配水網の設計・運用システム²⁾、階層型ニューラルネットワークを応用した凝集剤注入率の予測技法^{3)、4)}やファジィ推論による前塩素注入制御システム⁵⁾などがある。AI技術の開発では、知識の表現より経験・

知識を現場担当者から効率的に獲得することがボトルネックとされているが、ニューラルネットワークでは学習によって自動的に知識が獲得されるため、知識獲得作業を大幅に軽減できることが大きな特徴である。逆にニューラルネットワークからファジィ・プロダクションルールを抽出する手法も研究されている⁶⁾。

本論文のニューラルネットワークでは、学習法として D. Rumelhart らが提案した誤差逆伝播法 (バックプロパゲーション法、以下BP法) を適用した^{7), 8)}。BP法は多階層のニューラルネットに適用される一般化されたデルタルールであり、ニューラルネットの出力信号と教師信号との二乗誤差を最急降下法によって最小化し、中間・出力層ニューロンの結合係数を調整する学習則である。

ニューラルネットワークは、複数のユニットからなる入力層 (添字 i)、中間層 (添字 j) および単一ユニットの出力層の計3層で構成し、下位ユニットを上部ユニットすべてに結合させる。そして出力層、中間層ニューロンの内部関数として U_{op} 、 U_{mpj} を定義し、これらを式(1)、(2)のようにニューロン間の結合係数 W_{oj} 、 W_{mij} とそれぞれのニューロンへの入力 M_{pj} 、 I_{pi} の積和で表わす。ここで p は複数の学習パターン (入出力) セットを示す添字である。出力層、中間層からの出力 O_p 、 M_{pj} は式(3)、(4)のように、微分関数が原始関数で表わされるシグモイド関数 ($f(x) : df/dx = \mu \cdot f(x) \cdot [1-f(x)]$) を用いて内部関数 U_{op} 、 U_{mpj} から計算する。なお入力層については入力 I_{pi} をそのまま中間層へ出力するものとする。

$$U_{op} = \sum W_{oj} \cdot M_{pj} \quad \dots \dots \dots (1)$$

$$U_{mpj} = \sum W_{mij} \cdot I_{pi} \quad \dots \dots \dots (2)$$

$$O_p = f(U_{op}) = 1/[1+\exp(-\mu \cdot U_{op})] \quad \dots \dots \dots (3)$$

$$M_{pj} = f(U_{mpj}) = 1/[1+\exp(-\mu \cdot U_{mpj})] \quad \dots \dots \dots (4)$$

学習に際しては、出力層と中間層、中間層と入力層の結合係数 W_{oj} 、 W_{mij} を出力層からの出力 O_p と教師信号 T_p との二乗誤差の各学習パターンの総和 (E_t ; 式(5)) を評価基準としたデルタルールによって、それぞれ式(6)、(7) および式(8)、(9) に従い、学習ステップ (添字 k) 毎に修正していく。ここに η は加速係数、 α は右辺第二項によるオーバーシュートを抑制するスムージング係数である。

$$E_t = \sum (T_p - O_p)^2 / 2 \quad \dots \dots \dots (5)$$

$$\Delta W_{oj} = \eta \mu \sum (T_p - O_p) \cdot O_p \cdot (1 - O_p) \cdot M_{pj} \quad \dots \dots \dots (6)$$

$$W_{kj} = W^{k-1}_{oj} + \Delta W_{oj} + \alpha \cdot \Delta W^{k-1}_{oj} \quad \dots \dots \dots (7)$$

$$\Delta W_{mij} = \eta \mu \sum (T_p - O_p) \cdot O_p \cdot (1 - O_p) \cdot W_{oj} \cdot M_{pj} \cdot (1 - M_{pj}) \cdot I_{pi} \quad \dots \dots \dots (8)$$

$$W_{kij} = W^{k-1}_{mij} + \Delta W_{mij} + \alpha \cdot \Delta W^{k-1}_{mij} \quad \dots \dots \dots (9)$$

実際の学習の流れとしては、まず、乱数発生によって結合係数 W_{oj} 、 W_{mij} の初期値を設定したのち、中間層ユニットの U_{mpj} 、 M_{pj} を、次いで出力層ユニットの U_{op} 、 O_p を式(1)～(4)を用いて各学習パターン毎に計算する。次に各学習パターンの諸量から出力層と中間層ユニットの W_{kj} を、さらに中間層と入力層ユニットの W_{kij} を式(5)～(9)により更新していく。このような学習を予め設定した学習回数、あるいは予測 (想起) 誤差が任意の基準を満足するまで繰り返していく。

3 ニュラルネットワークによる原水水質の予測

3.1 ニュラルネットの構成

1) 出力層ユニット (予測対象) 1993年12月、微量な有機化学物質を中心に飲料水の水質基準が大幅に改訂された。現在、過マンガン酸カリウム消費量 (以下 $KMnO_4$ 消費量) には基準として 10mg/l 、快適目標として 3mg/l の規制値が設定されている。 $KMnO_4$ 消費量で表される河川の有機性汚濁は、ダムなどの貯水施設による滞留時間の延伸等によって増加しており、また、降雨に伴う流量増加時には図1 (大阪府水道部村野浄水

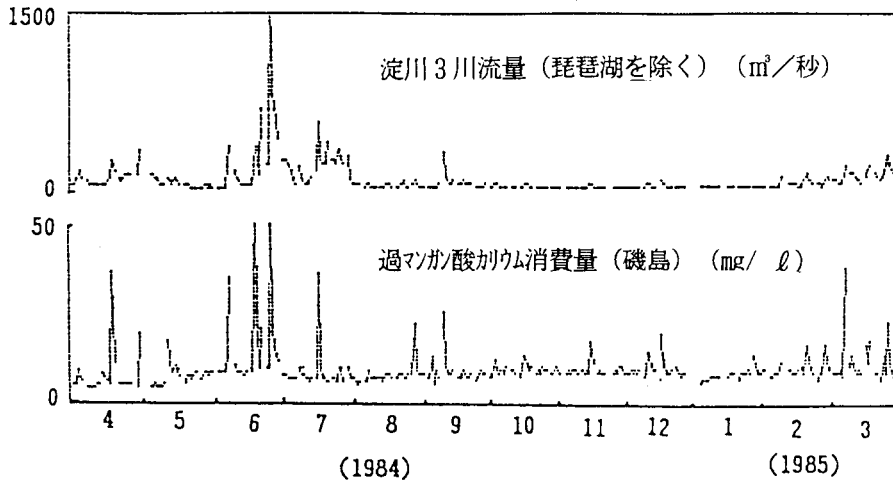


図1 淀川3川流量およびKMnO₄消費量の日単位時系列
(1984年4月1日～1985年3月31日)

場原水)⁹⁾に示すように、河床からの巻き上げや面的な発生源(ノンポイント・ソース)からの流出によって急激に濃度が上昇し、砂ろ過池でのろ抗上昇や水道水での濃度上昇などを引き起こすことがある。さらにKMnO₄消費量はトリハロメタンや全有機ハロゲン化合物の生成と強い濃度相関を有する。

また濁度は、浄水処理において凝集剤注入率を決める重要なファクターであり、最近では水道におけるクリプトスポリジウムの暫定対策指針として、ろ過池出口において常に0.1度以下を維持することが求められている。このような点から本論文では、水道原水のKMnO₄消費量と濁度をニューラルネットワークの予測対象とした。

ニューラルネットの適用期間は1984年4月から翌85年3月までの1年間(図1)とし、期間中の日単位手分析データのうち、KMnO₄消費量については対前日差が+3.0mg/l以上あるケース(学習パターン数21点)を、また、濁度については対前日差が+20度以上あ

- 木津川流域 : (a) 曾爾、(b) 針、(c) 大宇陀
- (d) 田辺
- 琵琶湖流域 : (e) 今津、(f) 彦根、(g) 八幡
- (h) 土山、(i) 大津
- 宇治川流域 : (j) 鷺峰山、(k) 信楽
- 桂川流域 : (l) 京都、(m) 園部

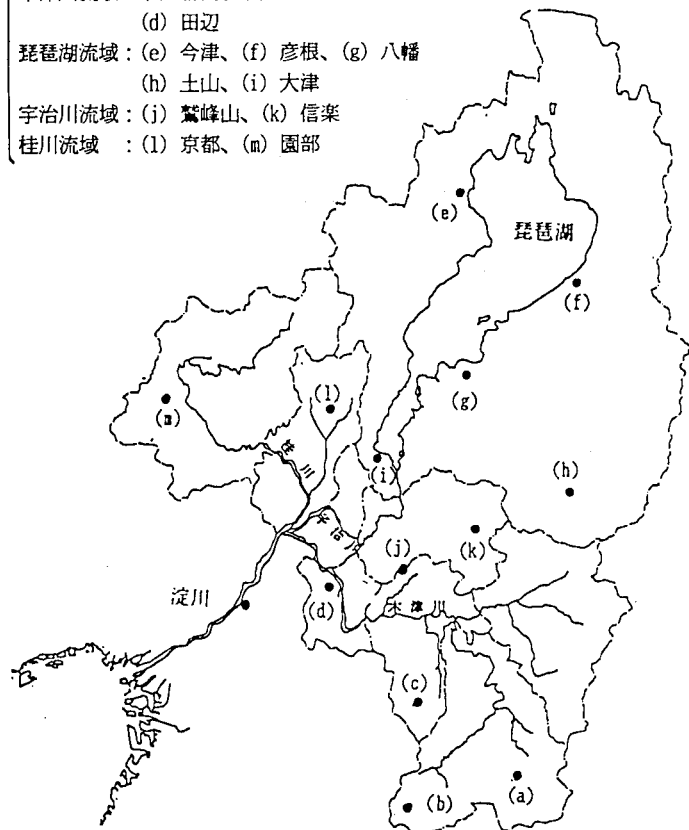


図2 琵琶湖・淀川流域におけるAMeDAS雨量観測地点

るケース (学習パターン数17点) を出力ユニット (予測値) の対象とした。出力データ (教師データ) は0と1の値をとらぬように増分量を期間中の最大値の1.1倍で除して規格化した。

2) 入力層および中間層ユニット 入力データには流域内にあるAMeDAS雨量観測地点 (図2) の日降雨量 (13点)、淀川3川 (木津川、宇治川、桂川) 流量 (午前6時時点) の対前日増分量 (3点) および各3川の前日流量 (3点) を用い、表1に示すこれらの組み合わせ3通りを入力パターンの候補として検討した。なお日単位雨量については、手分析データが午前10時時点の値であることを考慮して当日雨量の半量分と前日雨量の合計量とした。また、入力データは出力データと同様、いずれも最大値で除して規格化した。

ニューラルネットワークによる予測は、まず、KMnO₄消費量を対象に行った。表2は、表1の各入力セットおよび中間層ユニット数3、6、9個の各組み合わせにおける

表1 入力層ユニットの組み合わせ

入力セット	ユニット数	内訳
①	13	日単位雨量 (13)
②	16	日単位雨量 (13) + 流量増分 (3)
③	19	日単位雨量 (13) + 流量増分 (3) + 前日流量 (3)

表2 入力セットおよび中間ユニット数の検討結果

		入力セット			
		①	②	③	
中間 ユ ニ ツ ト 数	3	σ_t	10.6	7.8	9.9
		n	200	200	300
		μ	0.7	0.7	0.7
6	σ_t	8.0	7.4	8.4	
	n	300	300	300	
	μ	0.5	0.5	0.5	
9	σ_t	9.2	8.7	8.9	
	n	300	300	300	
	μ	0.4	0.4	0.4	

*) n: 学習回数、 μ : シグモイドパラメータ

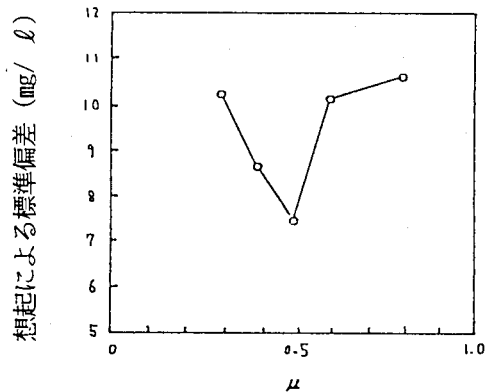


図3 シグモイドパラメータ μ の最適化
(入力セット; ②、中間ユニット数; 6)

学習効率を、二乗想起誤差E から算出される標準偏差 σ_t を評価基準にして検討したものである。学習の出・入力パターンには当該年間の前半部11点 (4~8月分) を降順にソーティングしたものを用いた。検討では加速係数 η とスムージング係数 α をともに1.0に固定した。

各組み合わせにおける標準偏差の比較から、入力層のユニット数として16 (入力パターン)、中間層のユニット数としては3~6が適切であることが判った。また、図3よりシグモイドパラメータは0.5にセットした。なお検討時、学習の効率や結合係数の最適化に関して次の点に留意した。

- ・シグモイド関数のパラメータ μ は学習効率に極めて敏感に影響する。この値が大きいと学習初期に誤差が大きく振幅し、逆に小さいと学習速度 (二乗誤差の降下速度) が遅くなる。シグモイド関数のスロープ部にデータが散布されるようにするのが望ましい。
- ・中間層のユニット数を増やすためには μ を小さくする必要がある。それに伴い学習速度は遅くなる。
- ・データの対数変換は学習効率 (線形分離性) を低下させる。

3.2 予測 (想起) シミュレーション

1) 予測 (想起) 結果

入・出力セットをニューラルネットワークが学習する過程を記録過程、またニューラルネットワークに入力データが与え

られ、なんらかのデータが出力される過程を想起過程という。

ここでは全入・出力パターン(21点)のうち、前半の11点を学習パターン数の初期値として、以後2点ずつパターン数を増やして予測(想起)特性の推移をみた。シミュレーションパラメーターは表3のとおりである。図4に予測(想起)結果を示す。図から判るように学習パターン数が増えるほど、誤差に対応する標準偏差は減少し、予測(想起)能力が上昇している。同図には未学習データ2点に対する想起結果も併示したが、学習済みデータと同様に標準偏差は減少しており、ニューラルネットワークへの学習パターンの取り込みに伴い、未学習データに対する予測能力が向上していることが判る。

図5は、全入・出力パターンを学習させたあとの予測(想起)値と実測値を图示したものであるが、相関係数から判断されるように極めて良好な想起結果が得られた。

図6は、濁度について予測(想起)値と実測値を图示したものである。濁度に関するシミュレーションパラメーターは学習回数が800回であるほかは、 $KMnO_4$ 消費量と同一である。濁度についても良好な予測(想起)結果が得られた。

表3 シミュレーションパラメーター

入力セット	表1中の②
中間総ユニット数	5
加速係数 η	1.0
スムージング係数 α	1.0
シグモイドパラメーター μ	0.5
学習回数	250

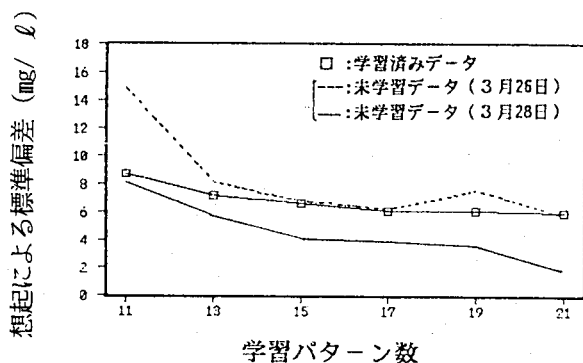


図4 予測(想起)による誤差偏差の推移

2) 感度解析による
学習知識の明示

学習済みニューラルネットワークには、原水水質予測に関する知識がユニット間の結合係数に分散的に内在されていると考えられる。モデルの充実や拡張性、浄水操作員との対話性を考えると、このような内在知識が、例えばファジィ推論のルール

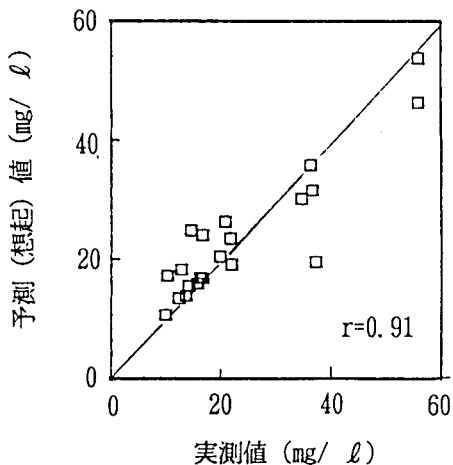


図5 予測(想起)値と実測値の比較 (KMnO₄消費量)

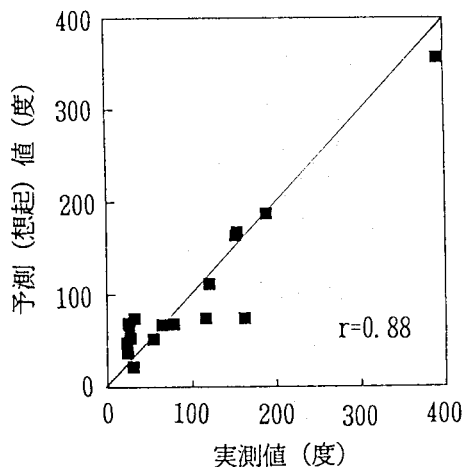


図6 予測(想起)値と実測値の比較 (濁度)

セットように明示的に表現されることが望ましい。図7は、ニューラルネットワークの入力ユニットの個々に、規格化した3種類の入力信号である0.25(低強度)、0.50(中強度)、0.75(高強度)をパルス的に与えたときの出力特性(規格化された $KMnO_4$ 消費量の増分)をみたものである。その結果、13の降雨観測点のうちでは特に3地点(大津、田辺および針観測点)にパルス強度に応じた強い応答が見られた。また取水点から距離的に近い理由から淀川3川のうち木津川の応答が最も強かった。試行錯誤的ではあるが、このような感度解析を通して明示的知識やルールの獲得が可能になると考えられる。

4 まとめ

筆者らは、浄水場における新しい計測・制御技術の一つとしての水源情報処理ツール(システム)の開発を念頭に、ニューラルネットワークを応用した原水水質予測のための手法を考案した。

予測モデルは、 KMnO_4 消費量あるいは濁度の増分を出力層として、中間層に1層、入力層には流域の日単位雨量河川流量の対前日増分量を採用した3層型の階層ニューラルネットワークで

あり、学習則には誤差逆伝播法(バックプロパゲーション法)を用いた。モデルパラメーターを trials and errors で最適化し、一年間の学習パターンで予測(想起)シミュレーションを行ったところ、学習済みデータ、未学習データ両者に対して良好な予測(想起)能力を確認することができた。さらに感度解析を通して、モデル内に内在する学習知識をファジールール等の形態で明示化しうる可能性を示した。

当手法による原水水質の予測は浄水場の効率的なプロセス運用に寄与する考えられ、また流域降雨や流量などの外部と件情報の取捨選択は水道からみた水源保全のあり方に新たな視点を提供するもの考えられる。

(参考文献)

- 1) 船橋誠壽、鈴木正義、笠井武郎、福原雅之：上下水道における知識工学応用、日立評論、Vol. 70、No. 6 pp. 5-10(1988).
- 2) 向出正章、竹中道夫、守本修司：工業用水へのAI適用
- 3) 圓佛伊智郎、原直樹、馬場研二、植木茂：ニューロ応用薬注支援システムの実機検証、第43回全国水道研究発表会、pp. 575-577(1992).
- 4) 馬場研二、圓佛伊智郎、松崎晴美、野北舜介：履歴学習型・上下水プラント運運転支援システム、第3回水システム自動計測制御国内ワークショップ論文集、pp. 165-168(1989).
- 5) 吉田俊雄、塚本孝：ファジィ推論を用いた前塩素注入の検討、第43回全国水道研究発表会、pp. 569-571(1992).
- 6) 圓佛伊智郎、馬場研二、松崎晴美、依田幹夫：ニューラルネットを用いたプラント運転ルールの抽出に関する研究：電気学会論文集D、Vol. 111、No. 1、pp. 20-28(1991).
- 7) 中野肇監修：ニューロコンピュータ、技術評論社、pp. 318(1990).
- 8) 中野肇著：ニューロコンピュータの基礎、コロナ社、pp. 248(1991).
- 9) 大阪府水質試験所水質試験成績並びに調査報告(昭和59年度).

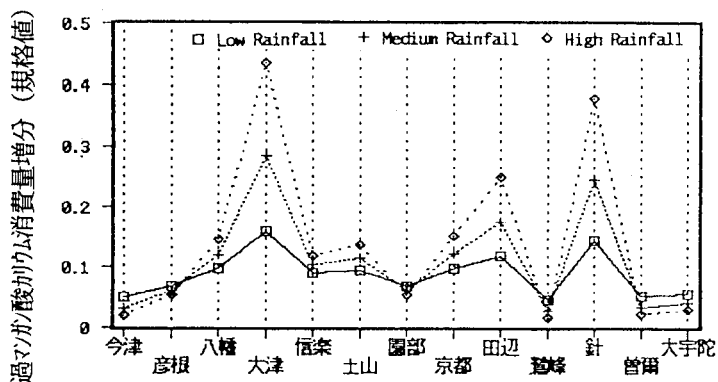


図7 ニューラルネットによる入力ユニットの感度解析