

GMDHによる配水量の予測

神戸市 水道局 正員 ○石田 知行
 神戸大学工学部 正員 神吉 和夫
 神戸大学工学部 正員 松梨順三郎

1. はじめに： 神戸市水道局では、テレメータ・テレコントロール方式で送配水施設の集中管理を行っている。集中管理にとって、最適制御は、機器の故障や事故の早期発見とらなる重要な問題であり、配水量の予測は、最適制御を行うにあたり、大きなウェイトを占めている。従来、配水量の予測には、推計式・線形回帰式・時系列による推定法等が用いられて来たが、配水量は、自然的・社会的要因に影響され、その構造は複雑で取り扱いが困難である。本研究では、複雑なシステムの同定予測に優れているGMDHを、配水量の予測問題に適用した。

2. GMDHの配水量予測への適用：^{2),3)}GMDH(Group Method of Data Handling)は、ソ連のA.G. Ivakhnenkoによって開発されたもので、発見的自己組織化(Heuristic Self-Organization)の原理に基づいて、複雑(高次元・非線形等)なシステムを同定予測するための一般的方法である。そのアルゴリズムは、変数のランダムな結合を発生させ、階層構造で、発見的規範(Heuristic Criterion)を用いて閾値的に自己選択するもので、発見的規範の最適化を行うことにより、システムの最適な同定が行われる。わが国におけるGMDHの適用例としては、^{4),5),6)}河川流量の予測や河川水質の予測、大気汚染の予測等がある。

配水量の予測において採用した発見的規範は次のとおりである。

④有用入力変数は、i)予測日の曜日・午前中の天気・午後中の天気・午前中の気温・午後中の気温の5ヶを基本変数とし、これらに、ii)前日と前々日の状態量および7日前の配水量を加えたもの・iii)基本変数の逆数や自乗を加えたもの・iv)基本変数相互の積を加えたもの・v)基本変数相互の和を加えたもの・vi)基本変数の逆数・積・和を加えたものとした。

②データは予測日の前日から逆のぼって、30日・60日・90日間を採用し、トレーニングデータとチェックングデータの分割は、予測日の前日より、交互に、奇数日をトレーニングデータ、偶数日をチェックングデータとした。

③システムの完全記述式としては、Kolmogorov-Gaborの多項式を想定し、基礎関数は、次の3式を用いた。

$$Z_k = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 x_i^2 + a_5 x_j^2, \quad \bar{Z}_k = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j, \quad \bar{Z}_k = a_0 + a_1 x_i + a_2 x_j$$

④各層の部分記述式において、チェックングデータによる予測誤差(2乗平均誤差または相対2乗平均誤差)の小さい変数から順に、5ヶまたは10ヶを中間変数として採用した。

⑤予測誤差の最小値が、前層の最小値を越えたとき、制限層の4または9層に達したとき、計算を中止して最終記述を行った。

以上の発見的規範を与え、その最適化を試みた。なお、基礎関数のパラメータの推定には、最小自乗法を適用しているが、ILL-Condition行列を持つ重共線関係が発生した場合は、その変数の組合わせを除いている。

3. 予測結果および考察： 神戸市の熊内低層配水池のS.50年のデータを用いて、夏期30

日間の配水量を予測した結果の一例を右図に示す。

予測の評価指標EZとして、予測水量 z_k と実績水量 y_k の相対2乗平均誤差 $EZ = \left[\frac{1}{n} \sum_{k=1}^n \left\{ (z_k - y_k) / y_k \right\}^2 \right]^{1/2}$ を用いたとき、 $EZ = 0.055$ であり工学的目的としては十分と思われる。最終記述の平均層は3.2層であり、採用された有用入力変数の平均数は4.5ヶである。

30日間の予測で、配水量の構造（最終記述層・採用された有用入力変数の数および種類）は、日によって異っており、このことは、システムの複雑さとともにGMDFHの柔軟性を示している。種々の数値実験より、配水量の予測に関して、次のようなことが言える。

①有用入力変数は、基本変数とその逆数・積・和を用いるのがよい。

②データ期間は90日がよく、短かすぎるとモデルが不安定になる。

③基礎関数は、 $z_k = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 x_i^2 + a_5 x_j^2$ または、 $z_k = a_0 + a_1 x_i + a_2 x_j$ がよい。

④中間変数の採用数は5の方がよく、有害な変数は、前もって除去しておく方がよい。

⑤最高層の制限は4の方がよい。これはILL条件とも関連するが、層が増すに従って、中間変数間に重共線関係が生じやすくなり、パラメータの信頼性が低くなる。

さらに、トレーニングデータとチェックデータの分割を行わず、構造決定とモデルの再現性のチェックを同一のデータで行い、それ以外のアルゴリズムはGMDFHと同様にして予測した場合、予測精度は悪くなった。

計算時間は、一日の予測に対して、構造の決定およびパラメータの推定を行う場合、約20秒（FACOM230-25/35）であるが、パラメータの推定のみでは約0.5秒であり、オンライン制御としても十分活用できると思われる。

4. おわりに： 以上GMDFHを用いて配水量の予測を行ったが、高次・非線形系の同定予測に適すると言われているGMDFHの特性を発揮し、工学上十分な精度で予測でき、また、処理時間的にもオンライン制御に耐え得ることが明らかとなった。しかしながら、これで十分とは言えず、変数やデータの取り方を始めとする発見的規範をさらに改良し、予測精度の向上を試みると同時に、従来の回帰モデル等との比較検討を行ってみたい。

なお、本研究は、送配水施設の最適制御問題の一環として行ったものであり、本研究の機会を与えていただき、かつ、終止協力していただいた神戸市水道局技術部の皆様に謝意を表する次第である。

参考文献 ①村尾正信：送配水施設の集中管理計画の概要，水道協会雑誌，383，1966。

②AG.Ivakhnenko: Group Handling of Data in Identification of the Static Characteristic of a Multi-External Plant, Soviet Aut. Cont., 14-2, 1969.

③AG.Ivakhnenko: Mathematical Simulation of Complex Ecological Systems, Soviet Aut. Cont., 4-4, 1971.

④池田・榎木：GMDFHと複雑な系の同定・予測，計測と制御，14-2，1975。

⑤市川・池田：発見的自己組織化法による水質予測，土木学会論文報告集，246，1976。

⑥田村・近藤：部族型式の次数を自己選択する改良型GMDFHとその広域気汚染濃度パターンの同定問題への応用，環境汚染の検知と制御 研究報告書，1976。

× 10³ m³/day 図-1 日配水量とその予測値

