

最近のGMDHの方法論と応用

田村坦之・近藤 正

1. まえがき

最近、システム工学は、環境問題、交通問題、資源問題など現象が複雑で従来の専門科学のみによる取扱いが困難な問題に対して、その問題解決への貢献が期待されている。しかしながら、このような実システムでは、非常に多くの変数やパラメータを含み、しかも、それらが相互に結合しあっているために、専門分野の知識のみに依存してシステムのモデリングを行なうことが一般的にいて困難である。ソ連のキエフ学派に属する A. G. Ivakhnenko らによって提案された GMDH^{1)~4)} (Group Method of Data Handling) は、このような複雑な構造をもった非線形システムに対して、入出力データから発見的自己組織化 (Heuristic self-organization) の原理にもとづいてモデリングを行なう方法論を与えている。GMDH は、

- i) 数少ない入出力データで非線形システムのモデリングを手軽に行なうことができる、
 - ii) システムの構造に関する先験的な情報を必要とすることなく、モデル構造の自己選択が可能である、
- といった利点をもっているが、その反面いくつかの問題点も含んでおり、1968年に基本的 GMDH が提案されて以来、数々の改良がなされてきた。GMDH の改良は、より適切な複雑さをもつモデルの構成の方法に関するものと、モデルの選択規準に関するものがほとんどである。GMDH の基本的概念である自己組織化の原理や基本的 GMDH の構造に関しては、すでにすぐれた解説⁵⁾があるので、ここではこれらについてはふれず、おもに GMDH の実用的な面、つまり、基本的 GMDH のアルゴリズムとそれに加えられた改良、および、これらの GMDH の実際問題への応用などについて解説する。

2. 基本的 GMDH

システムの入力変数 (x_1, x_2, \dots, x_m) と出力変数 ϕ との間に

$$\phi = f(x_1, x_2, \dots, x_m) \quad (1)$$

という従属関係があるものとする。(1)式をシステムの完全表現式という。システムの完全表現式として、たとえば、多項式、Bayes 関数、周期関数、有理関数などを用いた GMDH のアルゴリズム^{2),3),7)} が考えられているが、ここでは、その中で現在もっとも多く用いられており、利用価値の高い Kolmogorov-Gabor の多項式、

$$\phi = a_0 + \sum_i a_i x_i + \sum_j \sum_j a_{ij} x_i x_j + \sum_i \sum_j \sum_k a_{ijk} x_i x_j x_k + \dots \quad (2)$$

を用いたアルゴリズム^{3),4)} について説明する。この多項式は、原理的に 2 変数の組合せによる 2 次多項式 (これを部分表現式という)、

$$y_k = b_0 + b_1 x_i + b_2 x_j + b_3 x_i x_j + b_4 x_i^2 + b_5 x_j^2 \quad (3)$$

$k = 1, 2, \dots, m(m-1)/2$

を仲介させることによって構成することができる。ここで y_k を中間変数という。以下に基本的 GMDH のアルゴリズムを示す。

[ステップ 1]

システムの入出力変数を決定し、必要があればデータの前処理を行なう。

[ステップ 2]

入出力データを、部分表現式((3)式)のパラメータ推定に用いる training データと、中間変数 (y_k) の選択に用いる checking データの 2 組に分割する。分割の方法は training データと checking データを交互にわけける方法、あるいは分散の大きいデータを training データに、分散の小さいデータを checking データにわけける方法などがよく用いられている。

[ステップ 3]

2 変数の組合せをつくり、training データを用いて、部分表現式のパラメータ b_0, b_1, \dots, b_5 を、最小自乗法によって推定する。

[ステップ 4]

得られた部分表現式により、checking データに対する自乗誤差、

$$A_{ch} = \sum_{i=1}^{N_{ch}} (\phi_i - y_{ki})^2 \quad (4)$$

を計算する。そして、この値が小さくなるものから順に L 個の中間変数を選択(閾値的自己選択)し、残りは捨てる。(4)式を regularity criterion とよぶ。 N_{ch} は checking データの個数をあらわす。

[ステップ5]

$x_i=y_i, x_j=y_j$ として再びステップ3と4の操作を繰返す。このようにして、2変数の2次多項式を多層に積重ね、checking データに対する自乗誤差が改善されなくなったところでこの繰返し操作を打切る。

これらの操作により、第 T 層を通過させて計算を終了すると、出力変数 ϕ を、もとの入力変数に関する 2^T 次の多項式によって構成したことになる。基本的 GMDH のブロック図を図1に示す。ここで述べたアルゴリズムでは、部分表現式として(3)式の2次式を用いているが、このほかに部分表現式として提案されているものを以下に列挙する。

(i) 1次式の部分表現式

$$y_k = b_0 + b_1 x_i + b_2 x_j \quad (5)$$

部分表現式に1次式を用いると、システムの完全表現式もやはり1次式になる。

(ii) 2次の部分表現式

双線形の部分表現式として、

$$y_k = b_0 + b_1 x_i + b_2 x_j + b_3 x_i x_j, \quad (6)$$

推定すべきパラメータの数を減少させる部分表現式として、

$$X_k = w x_i + (1-w) x_j \quad w: \text{重み} \quad (7a)$$

$$y_k = b_0 + b_1 X_k + b_2 X_k^2 \quad (7b)$$

が提案されている⁹⁾。

(iii) 2次よりも高次の部分表現式^{9),9)}

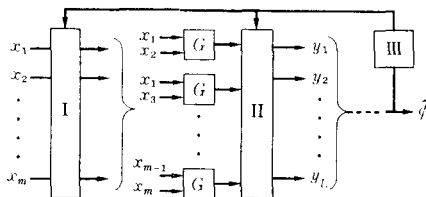
$$X_i = b_0 + b_1 x_i + b_2 x_i^2 \quad (8a)$$

$$X_j = b_0' + b_1' x_j + b_2' x_j^2 \quad (8b)$$

$$y_k = c_0 + c_1 X_i + c_2 X_j + c_3 X_i X_j + c_4 X_i^2 + c_5 X_j^2 \quad (8c)$$

この式を用いると、第 T 層を通過した段階で、もとの入力変数に関して 4^T 次の多項式が構成される。

以上の議論から明らかなように、基本的 GMDH のアル



I: 原データの分割 II: 中間変数の自己選択
III: 閾値の最適化 G: 部分表現式の発生器

図1 基本的 GMDH のブロック図

ゴリズムには、

H1. 入力データの training データと checking データへの分割の仕方を決める、

H2. 各層で選択する中間変数の個数を設定する、

H3. 部分表現式の形を設定する、

といった heuristics を含んでいる。このため、予測精度のよいモデルを得るためには、これらの heuristics の最適化を行なう必要があり、かなりの回数の GMDH の繰返し計算を必要とする。さらに、基本的 GMDH にはつぎのような問題点が存在する、

L1. 同定されるモデルは3つの heuristics H1, H2, H3 に大きく左右される。

L2. 部分表現式として2次以上の式を用いる場合、次数が低くて、数多くの入力変数を含むモデルは得られず、入力変数の数の多いモデルは不必要に複雑になりすぎる。

L3. 同定されたモデルの training データに対するあてはまりはよいが、checking データに対するあてはまりは悪い。

つぎに、これらの問題点を克服するために考えられてきた GMDH の改良点について述べる。

3. GMDH の改良

GMDH の方法的な改良については、問題点 L2 を克服するために部分表現式の構成の方法に改良を加えたものと、問題点 L3 を克服するためにモデルの選択規準に改良を加えたものがほとんどである。まず、部分表現式の構成の方法のみに改良を加えたものを示す。

(i) (3)式に、有意性の検定を用いた変数の逐次選択法を適用して、部分表現式を構成するアルゴリズム¹⁰⁾。

この方法では、heuristics H3 は不必要になり、また、問題点 L1 の一部と L2 が改善される。しかし、部分表現式を構成する際に training データの情報だけを用いるので、問題点 L3 は改善されていない。

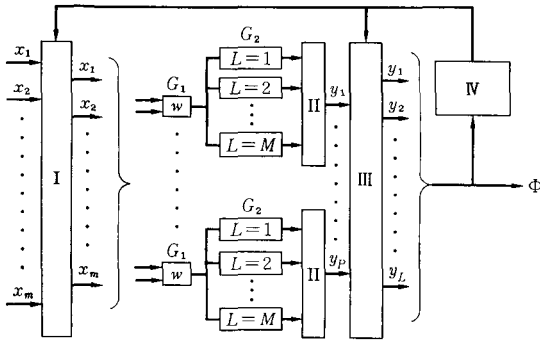
(ii) checking データに対する自乗誤差が最小になるように部分表現式の形を選択するアルゴリズム¹¹⁾。

この方法でも、heuristics H3 は不必要となり、問題点 L1 の一部と L2 が改善される。さらに L3 もかなり改善される。このアルゴリズムのブロック図を図2に示す。

これらの2つのアルゴリズムでは、基本的 GMDH と比べて完全表現式は非常に簡素化され、予測精度もよくなることが報告されている。

つぎに、部分表現式の構成の方法とモデルの選択規準の両方に改良を加えたアルゴリズムを示す。

(iii) unbiasedness criterion を用いる組合せの GM



- I : 原データの分割
- II : 最適な部分表現式の自己選択
- III : 中間変数の自己選択
- IV : 閾値の最適化
- G_1 : フィルター
- G_2 : 単一入力基本関数

図 2 改良形 GMDH(ii) のブロック図¹²⁾

DH のアルゴリズム^{12), 13)}.

この方法は、Ivakhnenko によって提案されたものである。

組合せの発生による部分表現式の構成法¹²⁾ :

これは(3)式の2次多項式を用いるかわりに、表1に示した変数の組合せを発生し、その中である規準(unbiasedness criterion)をもっとも小さくするものを部分表現式として用いる方法である。

unbiasedness criterion¹³⁾ :

この規準は問題点 L3 を改善するために考えられたものである。

まず、入出力データを2組 A_1, A_2 に分割し、それぞれのデータの個数を R_1, R_2 とする。 A_1 を training データ、 A_2 を checking データとして、部分表現式 $y_k^* = f_1(x_i, x_j)$ を求める。つぎに、データの割り振り

交換して、 A_1 を checking データ、 A_2 を training データとして、部分表現式 $y_k^{**} = f_2(x_i, x_j)$ を求める。第 T 層の k 番目の組合せに対する unbiasedness index は、

$$n_{Tk} = \frac{1}{R_1 + R_2} \sum_{r=1}^{R_1 + R_2} (y_{kr}^* - y_{kr}^{**})^2 \quad (9)$$

である。第 T 層における unbiasedness criterion は、

$$N_T = \frac{1}{F} \sum_{k=1}^F n_{Tk} \quad (10)$$

である。ここで、 F は第 T 層において選択された中間変数の個数をあらわす。

unbiasedness criterion を用いる組合せ的 GMDH のアルゴリズムはつぎのようになる。

まず第1層では、入出力データを2組に分割し、各入力変数の組合せに対して、表1の多項式の unbiasedness index を計算し、各組合せに対する部分表現式を決定する。つぎに、unbiasedness index の値の小さい中間変数を F 個選択し、(10)式の unbiasedness criterion を計算する。第2層以下では、第1層と同じ計算を繰返して行なう。多層構造の層の打切りは N_T が改善されなくなったところで行なう。

このアルゴリズムでは、問題点 L1 の一部と L2, L3 が改善されている。しかしながらデータを2組に分割しているため、依然として heuristics H1 を必要とする。また基本的 GMDH に比べて計算量はかなり増加することになる。

(iv) 構造安定性とパラメータ安定性を用いるアルゴリズム¹⁴⁾.

構造安定性 :

データを2等分割し、最小自乗法でデータ1から求めた部分表現式のデータ2へのあてはまりの程度を自乗平均誤差で評価し、つぎにデータの役割を交換して

表 1 部分表現式の構成¹²⁾

D	1	2	3	4	5	6
個数	$2^0=1$	$2^1=2$	$2^2=4$	$2^3=8$	$2^4=16$	$2^5=32$
部分表現式の右辺	$z_1 = a_0$	$z_2 = a_1 x_2$ $z_2 + z_1$	$z_3 = a_2 x_2^2$ $z_3 + z_1$ $z_3 + z_2$ $z_3 + z_2 + z_1$	$z_4 = a_3 x_1$ $z_4 + z_1$ $z_4 + z_2$ $z_4 + z_2 + z_1$ $z_4 + z_3$ $z_4 + z_3 + z_1$ $z_4 + z_3 + z_2$ $z_4 + z_3 + z_2 + z_1$	$z_5 = a_4 x_1 x_2$ $z_5 + z_1$ $z_5 + z_2$ $z_5 + z_2 + z_1$ $z_5 + z_3$ $z_5 + z_3 + z_1$ $z_5 + z_3 + z_2$ $z_5 + z_3 + z_2 + z_1$ $z_5 + z_4$ $z_5 + z_4 + z_1$	$z_6 = a_5 x_1^2$ $z_6 + z_1$ $z_6 + z_2$ $z_6 + z_1 + z_2$ $z_6 + z_3$ $z_6 + z_3 + z_1$

自乗平均誤差を求め、それらの和の小さいものを構造的に安定な部分表現式であるとする。大きいものほど不安定な部分表現式であるとする。

パラメータ安定性：

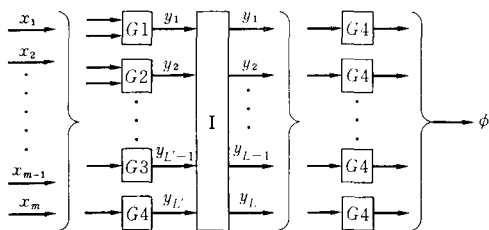
各2次多項式において、データ1, 2で独自に求めた回帰係数が互いに近い値を示す項をパラメータ的に安定な項、大きく異なっている項を不安定な項とよぶ。

これらの安定性を用いたアルゴリズムはつぎのようになる。各選択層における部分表現式は、(3)式の2次多項式に含まれている項の中でパラメータ的に安定な項のみによって構成し、中間変数は構造的に安定なものから順に選択する。そして、構造安定性が改善されなくなったところで層の打ち切りを行なう。構造安定性は Ivakhnenko が提案した unbiasedness criterion とほとんど同等なものであり、このアルゴリズムにおいても、問題点 L1 の一部と L2, L3 が改善されるが、データを2組に分割しているため、heuristics H1 はやはり必要である。

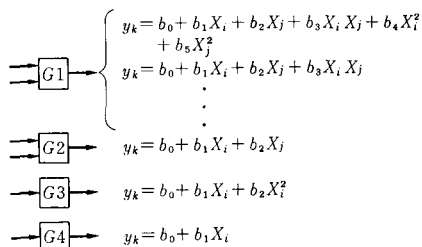
(v) モデルの選択規準に、予測平方和 (Prediction Sum of Squares) あるいは情報量規準 AIC を用いるアルゴリズム^{15), 16)}。

このアルゴリズムのブロック図を図3に示す。図3において、G1~G4は部分表現式の発生器を意味する。どの発生器を選択するかは、N個のすべてのデータを用いて、

$$PSS = \sum_{i=1}^N \left(\frac{\phi_i - \hat{\phi}_i}{1 - \mathbf{x}_i^T (X^T X)^{-1} \mathbf{x}_i} \right)^2 \quad (11)$$



I : 中間変数の自己選択
(a) ブロック図



(b) 最適な部分表現式の発生器

図3 改良形 GMDH(v) のブロック図¹⁵⁾

または、

$$AIC = N \log e S_k^2 + 2(k+1) + C \quad (12)$$

$$S_k^2 = \frac{1}{N} \sum_{i=1}^N (\phi_i - \hat{\phi}_i)^2$$

k: 変数の個数

を計算し、PSS¹⁷⁾ または AIC¹⁸⁾ を評価規準にした変数選択の方法によって自動的に行なう。Lは中間変数の選択個数の最大値を意味するが、この値は heuristic に決めるのではなく、計算機メモリの許される範囲内でできるだけ大きくとる。層を積重ねるにつれて、中間変数の個数は次第に減少すると同時に発生器 G4 の個数が増加する。部分表現式の発生器がすべて G4 になった層において、層の積重ねを停止する。なぜならば、さらに層を重ねても、前層と同一の中間変数を生成し、PSS または AIC の値が改善されないからである。最終的に、システムのモデルは、最終層に残った中間変数の(重みつき)平均の形で構成する。このアルゴリズムでは、入出力データを training データと checking データに分割せずに、すべてのデータを用いてモデルの予測誤差を推定し、この推定値をできるだけ小さくするように部分表現式の変数選択、中間変数の選択、多層構造の層の打ち切りを行なっている。その結果、heuristics H1~H3 を必要としないアルゴリズムを構成しており、問題点 L1~L3 を克服している。さらに他の GMDH と比べて、より少ない計算量で、簡単でしかも予測精度のよいモデルを容易に見つけることが可能である。

以上の GMDH の改良は、問題点 L1~L3 を克服するために、部分表現式の構成の方法とモデルの選択規準に対して行なわれたものである。このほかにも実用的な見地からいくつかの改良が考えられている。以下にこれらのアルゴリズムについて簡単にふれておく。

(vi) 長期予測を目的とした balance-of-variables criterion を用いるアルゴリズム^{19), 20)}。

これは Ivakhnenko によって提案された長期予測を目的としたアルゴリズムであり、balance function を用いるもの、また、direct function と inverse function を用いるものなどいくつかの方法が提案されている。これらのアルゴリズムに関しては紙数の関係上説明を省略するが長期予測に興味のある方は巻末の文献を参照されたい。

(vii) 新しい入出力データに対して逐次的にパラメータを更新させるアルゴリズム⁹⁾。

これは、基本的 GMDH によってすでにモデルの構造が求まっている時に、さらに新しい入出力データが得られた場合、モデルの構造は変化させずに逐次形最小自乗法を適用してパラメータの値のみを更新してい

くアルゴリズムである。このような方法を時系列データに適用すると、安定な予測値が得られることが報告されている。

(viii) システムの中間表現式を構成するアルゴリズム²¹⁾。

いままで述べてきた GMDH のアルゴリズムでは、各選択層において中間変数を入力変数にした部分表現式を構成して、これらを多層に積重ねているために、システムの入出力変数間に存在する物理的意味を明確に表現した多項式の構造を正確に求めることは、一般的に困難であった。しかし、各選択層において、部分表現式のかわりに、もとの入力変数の一部と出力変数の関係をあらわす中間表現式を構成することによって、ある程度の可能性を見出すことができる。中間表現式を用いる GMDH のブロック図を図 4 に示す。ここで、 $f_i^{(1)}$ 、 $f_i^{(2)}$ は各選択層において構成した i 番目の中間表現式をあらわす。このアルゴリズムでは、情報量規準 AIC を用いて、中間表現式に含まれる変数の選択、中間表現式の選択、多層構造の層の打ち切りを行っており、heuristics $H_1 \sim H_3$ を必要とせず、問題点 $L_1 \sim L_3$ を克服している。

以上の GMDH の改良のほかにも、Ivakhnenko らのグループにおいて数多くの研究がなされており、これらは文献²²⁾、²³⁾に集大成されている。

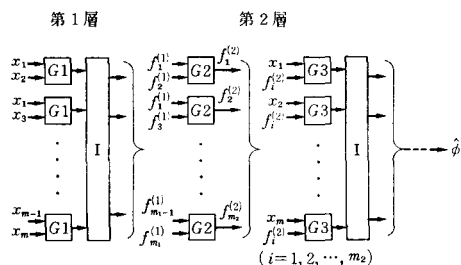
4. 応用例

4.1 基本的 GMDH の応用

ここでは、基本的 GMDH の応用例を紹介する。

イギリス経済のモデリング⁸⁾：

Ivakhnenko は、図 5 に示すイギリス経済のデータに GMDH を適用している。ここで、 v_1 は住民税、 v_2 は法人税、 v_3 は失業者に支払う保険料、 v_4 は資本、 v_5 は政府支出、 φ_1 、 φ_2 は経済指標をあらわし、1954年～1968年のデータを用いてシステムの同定を行ない、1969年度の φ_1 と φ_2 の値を予測している。同定方法としては、は



I：最適な中間表現式の自己選択

G_1, G_2, G_3 ：最適な中間表現式の発生器

図 4 改良形 GMDH(viii) のブロック図²¹⁾

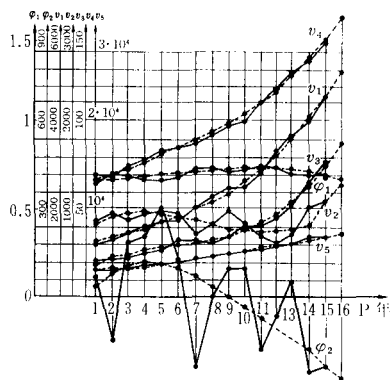


図 5 イギリス経済における入出力変数の動き⁸⁾

じめに、傾向変動を重回帰分析により取り除き、残った変動に対して、基本的 GMDH を適用している。システムの部分表現式としては、(3)式を用いている。

世界人口のモデリング⁸⁾：

GMDH を動的な世界人口のモデリングに応用している。

モデルとして、

$$P_n = f(P_{n-1}, P_{n-2}, P_{n-3}; E_n, E_{n-1}, E_{n-2}, E_{n-3}; F_n, F_{n-1}, F_{n-2}, F_{n-3}) \quad (13)$$

を仮定している。ここで、

P_n ：第 n 時点の世界人口、

E_n ：第 n 時点の 1 人当り世界消費エネルギー、

F_n ：第 n 時点の 1 人当り世界食糧生産指数。

システムの部分表現式としては、(7)式を用いている。

入出力変数の変動が比較的なめらかな問題に対しては、(7)式のような部分表現式を用いることが有効であることが示されている。

河川の流出量予測^{9), 9)}：

測定データとして、日々の烏川流域(利根川上流)における降水量(mm)、水位(m)、流出量(m^3/sec)が昭和40年5月1日～6月30日までの2カ月間にわたって与えられ、このデータに GMDH アルゴリズムを適用して1日先の流出量の予測を行なっている。部分表現式としては、(6)式および(8)式を用いている。同定に用いるデータの使用期間は、予測すべき日の前30日間としている。予測期間は、1965年6月1日～30日である。従来の水文学の立場から開発された流量予測モデル(菅原法および貯留関数法)による予測結果との比較を行なっており、比較結果は図 6 のようになっている。

その他の応用例：

Ivakhnenko らのグループは、1968年以来基本的 GMDH の応用を数多く行なっており、これらの結果は、雑誌「Soviet Automatic Control」に報告されている。

また、わが国におけるその他の応用例としては、佐藤氏

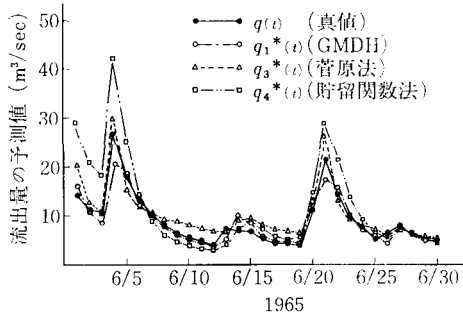


図 6 河川流出量の予測⁹⁾

による空容器回収量予測への応用²⁴⁾, 森田氏らによる製鉄所プロセスのモデリングへの応用²⁵⁾などがある。

4.2 改良形 GMDH の応用

ここでは、改良形 GMDH (i), (ii), (v) の環境汚染問題への応用例を紹介する。他の改良形 GMDH の応用例については、巻末の文献を参照されたい。

水質汚染濃度の短期および長期予測¹⁰⁾ :

これはアルゴリズム(i)の応用で、Decatur 湖の Sangamon 河と、Shelbyvill の Kaskaskia 河流域におけるトウモロコシ地帯(Corn Belt)の水質汚染を取扱っている。汚染物質として、硝酸塩をとっており、表 2 に示す入出力変数間の関係を GMDH により求めている。汚染レベルの予測は、離散時間幅が 1 週間の短期予測と、1 年間の長期予測の 2 つについて行なっている。短期予測の結果を図 7 に示す。現象が非常に複雑であるために従来の現象論的方法によるシステムのモデリングはき

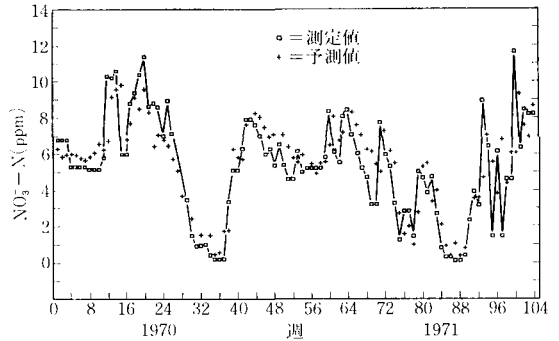


図 7 Sangamon 河における (NO₃-N) レベルの予測¹⁰⁾ (1970~1971年)——週

わめて困難であったが GMDH によって実用的な予測精度をもつモデルを構成できたことが報告されている。

広域大気汚染濃度パターンの同定^{11),15)} :

図 8 に示すような大気汚染濃度パターンの同定を、アルゴリズム(ii)と(v)を用いて行なっている。同定方法のブロック図を図 9 に示す。まず、排出源とモニタリングステーションにおける汚染濃度を関係づける大気汚染伝播マトリックスを重回帰分析により同定し、マトリックスを用いて、空間的な汚染濃度パターンの第 1 近似値を予測する。そして、この予測値とモニタリングステーションにおける観測値との差に、改良形 GMDH を適用し、第 1 近似値を補正することによって広域大気汚染濃度パターンを同定する。観測データのない地点に対する予測精度は、

表 2 Sangamon 河流域における (NO₃-N) の短期予測に用いる測定変数¹⁰⁾

Number	Variable	Location	Source
1.	(NO ₃ ⁻), Sangamon River	Cisco	CBNS*
2.	(NO ₃ ⁻), Sangamon River	Input to Lake Decatur	CBNS
3.	(NO ₃ ⁻), Sangamon River	Monticello	CBNS
4.	δ ¹⁵ N, Sangamon River	Cisco	CBNS
5.	δ ¹⁵ N, Sangamon River	Input to Lake Decatur	CBNS
6.	δ ¹⁵ N, Sangamon River	Monticello	CBNS
7.	Discharge of Sangamon River, cfs	Monticello	U. S. Dept. Int.
8.	Precipitation	Decatur	U. S. Dept. Comm.
9.	Precipitation	Decatur	Dec. Water Dept.
10.	Soil temperature 4" below surface	Urbana	USDC
11.	Soil temperature 8" below surface	Urbana	USDC
12.	Air temperature, maximum	Decatur	USDC
13.	Air temperature, minimum	Decatur	USDC
14.	Pan evaporation	Urbana	USDC
15.	Wind, total/day	Urbana	USDC
16.	NO ₃ ⁻ load (disch. (7) times aver. of (1,2,3))		calculated
17.	Available nitrate	Cerro Gordo	calculated

*Center for the Biology of Natural Systems, Washington University, St. Louis, Mo.

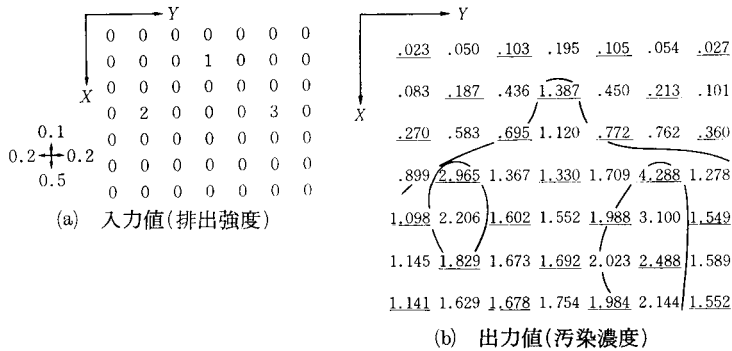


図8 複数排出源に対するシミュレーション結果¹⁵⁾

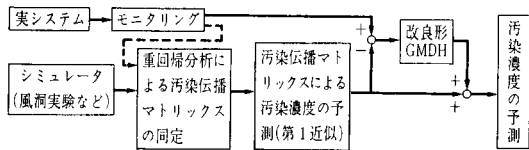


図9 改良形GMDHと汚染伝播マトリックスによる予測システムのブロック図¹⁵⁾

- 基本的 GMDH $J=23.1\%$
- 改良形 GMDH(ii) $J=19.7\%$
- 改良形 GMDH(v) $J=16.2\%$

である。基本的 GMDH に比べて、予測精度が大幅に改善されていることがわかる。

大気汚染濃度の短期予測¹⁶⁾：

アルゴリズム(v)を用いて、汚染質 SO₂ の 1～3 時間先の短期予測を行なっている。出力変数としては、SO₂ の 1 時間先の値をとり、入力変数としては、SO₂ の過去 6 時間分の値をとっている。2, 3 時間先の予測値は、同定したモデルに 1, 2 時間先の予測値を代入することにより求めている。同定に用いたデータの使用期間は 10 日間である。同定は 1 日ごとに繰返し行なっている。図 10 に 1 時間先の予測結果を示す。また、従来から大気汚

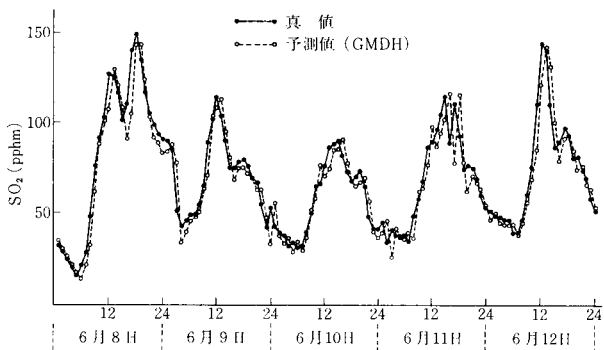


図10 改良形GMDH(v)による SO₂ 濃度の 1 時間先の予測値¹⁶⁾

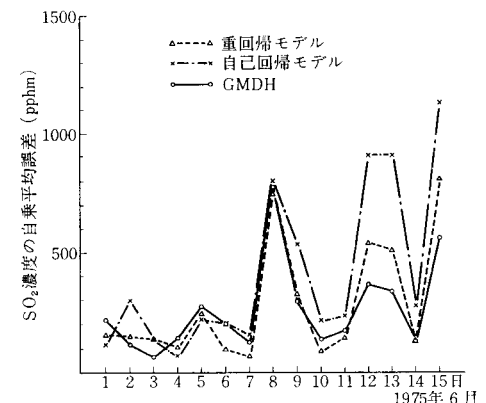


図11 SO₂ 濃度の 3 時間先の予測誤差¹⁶⁾

染の短期予測によく使われている線形モデルとの比較を行なっている。3 時間先の予測精度の比較を図 11 に示す。線形モデルと比べてかなり予測精度が改善されていることがわかる。

5. むすび

GMDH に関する種々のアルゴリズムとその実際問題への応用について述べた。GMDH は原理的には、データさえあれば、どの分野にも応

用が可能であるが、対象分野の理論すなわち現象論とは無関係に、単に入出力データの統計的な解析だけに頼ってブラックボックスモデルを導出すると、対象分野を理論的に研究している人たちからは、「理論なき安易なモデル」としてひんしゅくを買う結果におちいるので注意を要する。現時点における応用に関する結論としては、対象となるシステムの非線形性が顕著で、しかもその分野の個別理論が未発達であり、利用できるデータ数が少なく、モデリングのシステム的方法論としては GMDH 以外には考えられないときに手軽に応用してみる、といった使い方がもっとも適切な使い方ではないかと考える。今後、さらに有用な改良形 GMDH が開発されることを期待したい。

参考文献

- 1) A. G. Ivakhnenko : The Group Method of Data Handling, A Rival of the Method of Stochastic Approximation ; *Soviet Automatic Control*, 1-3, 43/55 (1968).
- 2) A. G. Ivakhnenko : Heuristic Self-Organization in Problems of Engineering Cybernetics ; *Automatica*, 6-2, 207/219 (1970).

- 3) A. G. Ivakhnenko : Polynomial Theory of Complex Systems; *IEEE Trans. Syst., Man, Cybern.*, **SMC-1-4**, 364/378 (1971).
 - 4) A. G. Ivakhnenko, et al. : Group Handling of Data in Identification of the Static Characteristic of a Multi-Extremal Plant; *Soviet Automatic Control*, **2-2**, 30/37 (1969).
 - 5) 池田三郎, 榎木義一 : GMDH(発見的自己組織化法)と複雑な系の同定・予測; 計測と制御, **14-2**, 185/195 (1975).
 - 6) A. G. Ivakhnenko and M. M. Todua : Prediction of Random Processes Using Self-Organization of the Prediction Equations, Part 1 Problems of Simple Medium-Term Prediction; *Soviet Automatic Control*, **5-3**, 35/52 (1972).
 - 7) D. L. Dylbokova and I. S. Dylbokov : Prediction of Trends of Development of Digital Computers by GMDH Using Linear-Fraction Partial Descriptions and Balance-of-Variables Criterion; *Soviet Automatic Control*, **8-2**, 24/30 (1975).
 - 8) 井原二郎 : 改良形 GMDH 一動的世界人口モデルの場合一; システムと制御, **19-4**, 201/210(1975).
 - 9) S. Ikeda, M. Ochiai and Y. Sawaragi : Sequential GMDH Algorithm and Its Application to River Flow Prediction; *IEEE Trans. Syst., Man, Cybern.*, **SMC-6-7**, 473/479 (1976).
 - 10) J. J. Duffy and M. A. Franklin : A Learning Identification Algorithm and Its Application to Environmental System; *IEEE Trans. Syst., Man, Cybern.*, **SMC-5-2**, 226/240 (1975).
 - 11) 田村坦之, 近藤正 : 部分表現式の次数を自己選択する改良形 GMDH とその広域大気汚染濃度パターンの同定への応用; 計測自動制御学会論文集, **13-4**, 39/45 (1977).
 - 12) N. A. Ivakhnenko and M. Z. Kvasko : Combination-Generating GMDH Algorithms in which the Regularity of Both Symmetrical and Nonsymmetrical Polynomials is Checked; *Soviet Automatic Control*, **5-5**, 33/38 (1972).
 - 13) A. G. Ivakhnenko, et al. : Discovery of Physical Laws by GMDH Method with the Absence-of-Bias Criterion; *Soviet Automatic Control*, **6-6**, 32/45 (1973).
 - 14) 遠藤明男, 畑中浩 : GMDH による非線形制御系の同定; 制御理論シンポジウム (計測自動制御学会主催), 93/98 (1974).
 - 15) H. Tamura and T. Kondo : Large-Spatial Pattern Identification of Air Pollution by A Combined Model of Source-Receptor Matrix and Revised GMDH; *Proc. IFAC Sympo. on Environmental Systems Planning, Design and Control*, 373/380, Kyoto (Aug. 1977).
 - 16) 田村坦之, 近藤正 : 改良形 GMDH による大気汚染濃度の短期予測; 第9回確率システムシンポジウム講演論文集 (日本自動制御協会主催), **L2**, 127/130 (1977).
 - 17) 奥野忠一ほか : 続多変量解析法; 日科技連(1976).
 - 18) 赤池弘次 : 情報量規準 AIC とは何か一その意味と将来への展望一; 数理科学, no. 153 (1976).
 - 19) A. G. Ivakhnenko, et al. : Long-Term Prediction of Random Processes by GMDH Algorithms Using the Unbiasedness Criterion and Balance-of-Variables Criterion; Part 1, *Soviet Automatic Control*, **7-4**, 40/45 (1974); Part 2, *S. A. C.*, **8-4**, 24/38 (1975); Part 3, *S. A. C.*, **9-2**, 28/42 (1976); Part 4, *S. A. C.*, **9-4**, 16/27 (1976).
 - 20) A. G. Ivakhnenko and B. K. Svetal'skiy : Self-Organization of World Dynamics Model according to Forrester's Data and Control Synthesis by Selecting the Vertices of the Hypercube of Feasible Controls; *Soviet Automatic Control*, **8-1**, 25/40 (1975).
 - 21) 近藤正, 田村坦之 : 情報量規準 AIC を用いた改良形 GMDH; 第3回システムシンポジウム講演論文集 (計測自動制御学会主催), 163/168(1977).
 - 22) A. G. Ivakhnenko : Present State of the Theory of Computer-Aided Self-Organization of Mathematical Models (Survey); *Soviet Automatic Control*, **8-5**, 18/26 (1975).
 - 23) A. G. Ivakhnenko : The Group Method of Data Handling in Prediction Problems; *Soviet Automatic Control*, **9-6**, 21/30 (1976).
 - 24) 佐藤敬 : 空容器回収量予測への GMDH の応用; システムと制御, **20-7**, 365/372 (1976).
 - 25) 森田徹ほか : GMDH による製鉄所プロセスモデルの同定; 鉄鋼オートメーションシンポジウム講演論文集 (日本自動制御協会主催), 47/50 (1977).
- (たむら ひろゆき・こんどう ただし 大阪大学工学部精密工学科)