

[32] GMDH による汚泥処理プロセスの動的モデル同定

近畿大学 工学部 ○馬場 鉄一 山脇 重信
 寝屋川南部広域下水道組合 山村 優 北岸 秀司
 " " 山下 輝幸
 岐阜薬科大学 小瀬 洋喜

1. はじめに

下水処理場の処理プロセスを系統的に総括して見れば、階層大規模システムであって、各サブシステムには、相互干渉をもっている。かかるシステム全体の制御を逐行するためには、システムの入出力定常特性または、動的特性を表現する数学モデルを基盤にする必要がある。そのためには、先ず、各サブシステムの数学モデルの構成が先決問題である。

一般に、物理的・化学的機構だけでなく、生物的・人的操作をも含む系では、その系構造に関する先験的情報は少なく、取得入出力データから数学モデルを構成する必要がある。そのための手法として、時系列分析¹⁾、スプライン関数による出力特性曲線への当て嵌め²⁾、サイバネティクスの方法論を指向するGMDH手法³⁾等がある。これら統計的モデリング法のうち、多層型のサイバネティクスのな入力データの情報処理を行っているという理由で、下水処理システムのモデリングには、GMDH法が最適と考えられる。

本論文は、全下水処理システムの一つのサブシステムである汚泥処理システムを取り上げ、その中でも、特に、濃縮・脱水過程を対象に、一年間の静的数学モデルの構成、すなわち、外部同定を行うために、GMDHに基づく計算アルゴリズムを開発した。また、このアルゴリズムは、動的同定を行うのにも有用であることを示している。この動的同定から、濃縮・脱水システムの制御に直接用い得る、状態方程式、出力方程式を作成し、合せて、操作入力変数の固定化についても考察している。

2. 既報告のGMDH

GMDHアルゴリズムは、多層型構造をしているので、各層において、部分記述推定器の選択が要求される。よって、部分記述推定器、選択器の構成次第で、同定対象モデルによって、その同定能力が左右されるので、対象モデルに対して、現在、種々のGMDHア

ルゴリズムが報告されている。代表的なものとして、

Basic型 Filter型 逐次型、次数自己選択型(Tamura型)がある。³⁾

これら4つの型のGMDHを用いて、寝屋川南部広域下水道組合川俣処理場で実測された濃縮・脱水システムに関して、第1表の入出力変数のデータによって、直接外部同定を行った。ここで、当処理場の、濃縮・脱水機構は、濃縮汚泥量は353(m^3 /日)で、脱水機は、圧力打込ポンプとフィルタープレスで、脱水ケーキ量は40(m^3 /日)である。これらの値は設計値である。

Table. 1. Inputs and Outputs Factors

| 入力変数 | | 出力変数 | |
|------------|-------|------------------------|--------|
| 記号 | 因子名 | 記号 | 因子名 |
| X_1 (°C) | 水温 | Y_1^* (%) | ケーキ含水率 |
| X_2 (%) | 含水率 | Y_2^* (t/cycle) | 乾固形分量 |
| X_3 (%) | 有機分率 | Y_3^* (t/cycle) | ケーキ重量 |
| X_4 (%) | 薬品注入率 | Y_4^* (kg/ m^3 /時) | 濾過速度 |

Y_1^* に関する同定結果は、各GMDHについて、第2表の層数、絶対値平均誤差が得られた。

Table 2 Comparison of Basic, Filter, Sequential,

Tamura's GMDH algorithms

| OUTPUT | Y_1^* | | Y_2^* | | Y_3^* | | Y_4^* | |
|------------|---------|-------|---------|-------|---------|-------|---------|-------|
| | Layer | Error | Layer | Error | Layer | Error | Layer | Error |
| Basic | 11 | 3.9% | 7 | 8.5% | 28 | 13.7% | 51 | 8.0% |
| Filter | 12 | 4.5 | 5 | 11.6 | 6 | 16.9 | 3 | 13.1 |
| Sequential | 20 | 4.0 | 18 | 8.9 | 17 | 12.9 | 46 | 9.2 |
| Tamura's | 5 | 4.3 | 9 | 11.1 | 2 | 17.5 | 5 | 10.6 |

GMDHアルゴリズムは、任意の2個の入力変数から成る関数によって、システムの出力に近似し、次に、前のステップの結果を新しく入力として、システムの出力に近似して、このステップを近似度が最大になるまで繰り返す方策をとっている。この繰り返すステップ数を層数と呼んでいる。よって、層数が多いことは、それだけ、計算効率が悪く、かつ同定されたシステムの入出力数学モデルは複雑になる。この意味で、第2表から分かるように、4型を濃縮・脱水システムの同定に直接

適用したのでは、同定誤差が少なければ、層数が増すか、また、その逆の相矛盾した結果となる。よって、濃縮・脱水システム固有の同定アルゴリズムの開発が必要である。

3. 補正型 GMDH

第2表から分るように、層数L、同定誤差Eに対して、

$$L \begin{matrix} \text{Filter} <L & \text{Tamura} <L & \text{Basic} <L & \text{Sequential} \\ \text{Basic} <E & \text{Sequential} <E & \text{Tamura} <E & \text{Filter} \end{matrix} \quad (1)$$

$$E \begin{matrix} \text{Basic} <E & \text{Sequential} <E & \text{Tamura} <E & \text{Filter} \end{matrix} \quad (2)$$

が成立している。ここで、誤差評価として、推定値 Y_j ($j = 1, 2, 3, 4$) を用いて、絶対値平均誤差 ϵ_j を

$$\epsilon_j = \frac{\sum_{i=1}^n |Y_{ji} - Y_{ji}^*|}{\sum_{i=1}^n |Y_{ji}^*|} \quad (8)$$

としている。

(1), (2)より、同定誤差を小さく、層数を少なくするには、結局、Basic・Filter型を基盤とすればよい。

すなわち、Basic型で Y_j^* を近似し、得られた近似値を Y_j 、記述式を Z_{Bj} とすれば

$$e_{ji}^* = Y_{ji}^* - Y_{ji}$$

$$e_{ji}^* = H_j(X_1, X_2, X_3, X_4), j=1, 2, \dots, 4$$

なる関係が得られていることが分り、その時の同定誤差は少ない。

(2) e_{ji}^* を新出力とみなして、Filter型で e_{ji}^* を次に近似すれば、推定値 e_{ji} と記述式 Z_{Fj} が得られる。ここで、ステップ(1)での絶対値平均誤差 ϵ_j とステップ(2)でのそれ、 $\hat{\epsilon}_j$ について、 $\hat{\epsilon}_j < \epsilon_j$ を保証するために、

$$\frac{\sum_{i=1}^n |e_{ji}^* - e_{ji}|}{\sum_{i=1}^n |e_{ji}^*|} < 1 \quad (4)$$

が成立するように、 e_{ji} を決定するものとする。

ステップ(2)の結果、完全同定値 \hat{Y}_{ji} 、完全記述式 Z_j は

$$\hat{Y}_{ji} = Y_{ji} + e_{ji} \quad (5)$$

$$Z_j = Z_{Bj} + Z_{Fj} \quad (6)$$

と得られる。

このアルゴリズムを補正型 GMDH アルゴリズムと

呼び、前章と同データに対する同定結果は、例えば、 Y_1^* については、第1図を得た。その他、 Y_2^* 、 Y_3^* 、 Y_4^* の各同定誤差は Y_2^* : 9.8%, Y_3^* : 14.3%, Y_4^* : 8.9% となった。ここで、添字の後の数字は層数を表わす。

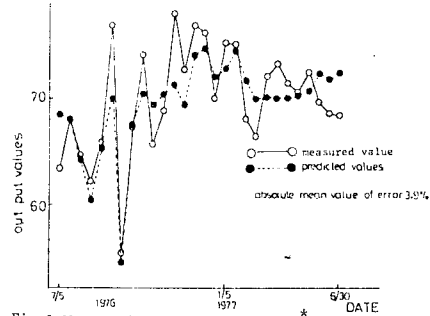


Fig.1 Measured Values of Output Y_1^* via Identified Values at 5th Layer of Revised GMDH Algorithm.

補正型 GMDH は、Basic 型をそのアルゴリズム中に含んでいるので、予測平方和 (PSS) が層数とともに単調減少し、安定な収束性をもったアルゴリズムと云える。

次に、この同定結果が、妥当なものであることを保証する意味で、同一処理場での 1977年7月~1978年8月までのデータに対して予測を行った。その結果、 Y_1^* については、第2図となり Y_2^* : 16.86%, Y_3^* : 22.21%, Y_4^* : 23.65% を得た。

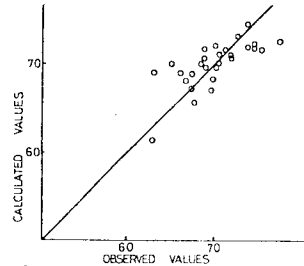


Fig.2 Comparison of Observed Value and Predicted Value (Y_1^*)

予測結果から、 Y_1^* (濾過速度) 以外は、ほぼ、同定誤差と同程度の予測誤差となっているので、補正型 GMDH によって得られる数学モデル式は、当処理場の濃縮・脱水システムの入出力定常特性を表わしていると云える。ただし、 Y_1^* は、フィルターの取り換え周期が不均一であり、予測誤差が大くなったと考えられる。

3. 動的同定手法と制御方程式

濃縮・脱水システムで取得される実データは、物理的操作だけでなく、化学反応、人的操作に基づく情報を含んでいるので、不規則データであり、かつ、定常過程で

もない。よって、動的モデル式を得る一手法である自己回帰モデル法の適用は困難である。

本来、GMDHは、入出力定常モデル式の構成のために開発されたアルゴリズムであって、得られたモデル式には、動的特性を含んでいない。しかし、本論文で用いているデータは時系列データであり、データの適当な取り扱いによって、GMDHの同定結果に動特性が持たせられる。

いま、1976年6月～1977年3月までのデータを第8、4表のように分割し、その同定誤差を求めると8分割の場合が最良であることが分る。ここで用いたデータは、同定結果との比較のため、後程考察する予測に用いる1977年6月～1978年8月のデータと同月同日だけを抽出した。よって、データ数は少なくなり、補正型GMDHアルゴリズムによる同定には、8分割以上の分割ではデータ数が不足するので、ここでは、8分割を最大分割数とした。

Table 3 Identification Error along Division Number.

| Division Number | Subsystem | X ₁ | X ₂ | X ₃ | X ₄ |
|-----------------|-----------|----------------|----------------|----------------|----------------|
| 2 | 1-2 | - | - | - | - |
| | 1-2 | 0.0600 | 0.1000 | 0.1600 | - |
| 4 | 2-3 | 0.1000 | 0.0500 | 0.0020 | - |
| | 3-4 | 0.1300 | 0.1400 | 0.1700 | - |
| | 1-2 | 0.0060 | 0.0230 | 0.0050 | 0.0020 |
| 6 | 2-3 | 0.0000 | 0.0610 | 0.0000 | 0.0003 |
| | 3-4 | 0.0002 | 0.0310 | 0.0030 | 0.0320 |
| | 4-5 | 0.0010 | 0.0120 | 0.0007 | 0.0005 |
| | 5-6 | 0.0020 | 0.0410 | 0.0020 | 0.0050 |
| | 1-2 | 0.0000 | 0.0010 | 0.0000 | 0.0001 |
| 8 | 2-3 | 0.0020 | 0.0000 | 0.0002 | 0.0002 |
| | 3-4 | 0.0001 | 0.0000 | 0.0000 | 0.0000 |
| | 4-5 | 0.0002 | 0.0000 | 0.0001 | 0.0018 |
| | 5-6 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 6-7 | 0.0000 | 0.0002 | 0.0012 | 0.0000 |
| | 7-8 | 0.0000 | 0.0000 | 0.0001 | 0.0001 |

Table 4 Prediction Error along Division Number

| Division Number | Subsystem | X ₁ | X ₂ | X ₃ | X ₄ |
|-----------------|-----------|----------------|----------------|----------------|----------------|
| 2 | 1-2 | - | 2.51 | - | - |
| | 1-2 | - | 1.30 | 10.9 | - |
| 4 | 2-3 | - | - | - | - |
| | 3-4 | 0.32 | 1.43 | 18.5 | - |
| | 1-2 | 10.9 | 5.60 | 27.1 | 13.2 |
| 6 | 2-3 | 11.2 | 12.4 | 8.74 | - |
| | 3-4 | - | 3.81 | 48.8 | 42.9 |
| | 4-5 | 37.9 | 16.2 | 14.4 | 31.3 |
| | 5-6 | 9.66 | 2.30 | - | - |
| | 1-2 | 19.3 | 6.40 | 18.6 | 44.7 |
| 8 | 2-3 | 35.9 | 1.90 | 7.30 | 5.28 |
| | 3-4 | 13.4 | 3.00 | - | - |
| | 4-5 | 47.7 | 7.20 | - | - |
| | 5-6 | 27.2 | 3.30 | - | - |
| | 6-7 | 14.6 | 5.70 | 5.30 | 38.3 |
| | 7-8 | 10.5 | 9.30 | 25.8 | - |

いま、分割されたデータ列を $\{x_i(\sum_{j=1}^{k-1} \tau_i + j)\}_{j=1}^{\tau_k}, \ell = 1, 2, \dots, 4$ なる数列で示すものとする。ここで、 k は分割時点の番号で、 $k = 1, 2, \dots, N$ 、 x_i は入力変数を表わし、 τ_k は k 番目の分割区間中の入力データ数である。一方、この入力列に対応する次の分割区間の入力列は $\{x_i(\sum_{j=1}^k \tau_i + j)\}_{j=1}^{\tau_{k+1}}$ である。取得データ個数が、 $\sum_{i=1}^N \tau_i$ 個であれば、濃縮・脱水全システムは、入力データについては、データの的に、第8図のように分割されたことになる。

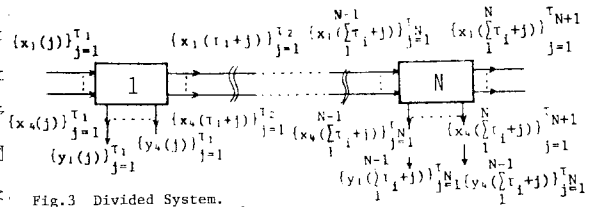


Fig.3 Divided System.

入力データを $\{x_i(\sum_{j=1}^{k-1} \tau_i + j)\}_{j=1}^{\tau_k}$ 、出力データと $\{y_i(\sum_{j=1}^k \tau_i + j)\}_{j=1}^{\tau_{k+1}}$ とみなして、補正型GMDHによる同定によって得られるモデル式は

$$x(\hat{k}+1) = g(\hat{k}, x(\hat{k})), \hat{k} = 1, 2, \dots, N \quad (7)$$

であり、また、同一の入力データ列と別な出力データ列 $\{y_i(\sum_{j=1}^{k-1} \tau_i + j)\}_{j=1}^{\tau_k}$ に対しては、

$$y(\hat{k}) = f(\hat{k}, x(\hat{k})), \hat{k} = 1, 2, \dots, N \quad (8)$$

を得る。ここで、 y_i は実出力変数を表わし、 \hat{k} は $\hat{k} = \{\sum_{i=1}^{k-1} \tau_i + 1, \sum_{i=1}^{k-1} \tau_i + 2, \dots, \sum_{i=1}^{k-1} \tau_i + \tau_k\}$ なる集合で $x \in R^4, y \in R^4$ で g, f はいずれも4次元ベクトル値関数である。

(7)、(8)を得るのに際して、各分割されたデータ列を、トレーニングデータと、チェックングデータに分けて同定し、適用性を付与した。その結果は、例えば、含水率についてみれば、(7)は第4図となり、また、ケーキ含水率についてみれば、(8)は第5図となった。また1977年6月～1978年3月のデータに対する予測結果は、同一因子について、それぞれ、第6図、第7図を得た。他の因子については、ほぼ同様の結果を得たが、ケーキ重量に対する予測結果のみが非常に悪くなった。この原因は同定・予測用のデータ数が少なく、かつ抽出時点での値が、たまたま、両者において、大きな開きがあったものと推察される。例えば、1976年12月24日は3.49

($\frac{t}{c}$)で1977年同日は1.87($\frac{t}{c}$)である。

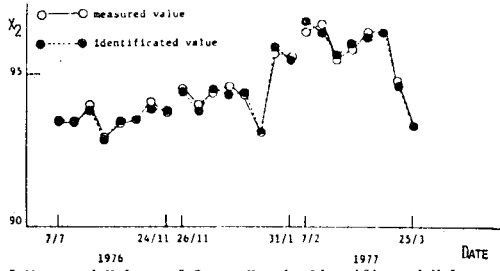


Fig. 5 Measured Values of Input X_2 via Identified Values - Divided Case into Checking and Training Data.

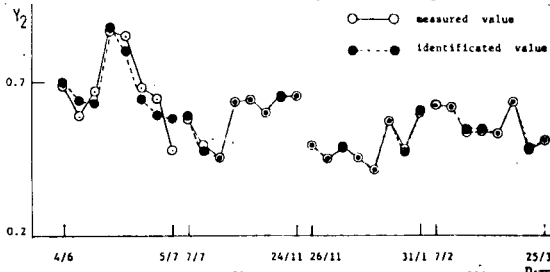


Fig. 6 Measured Values of Output Y_2^* via Identified Values - Divided Case into Checking and Training Data.

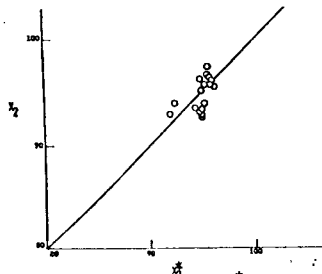


Fig. 7 Comparison of Measured Values X_2^* and Predicted Values X_2^* - Divided Case into Checking and Training Data.

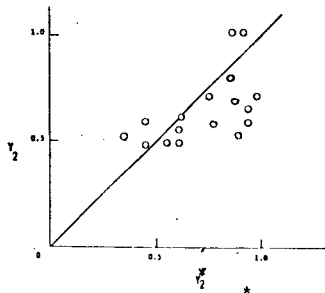


Fig. 8 Comparison of Measured Values Y_2^* and Predicted Values Y_2^* - Divided Case into Checking and Training Data.

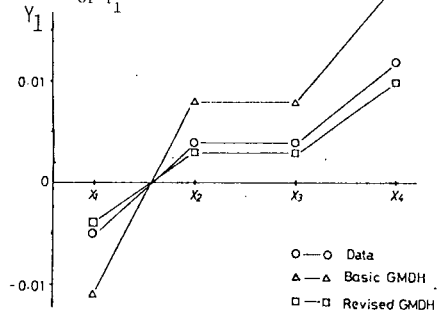
(7)、(8)は、1976年～1978年については、本処理場の濃縮・脱水システムの動的数値モデル式となり得ることは保証されたが、それ以後の予測、すなわち長期予測も可能であることを以下で議論する。

経済システムや工業システムと異なって、下水処理システムには、季節的周期性をもっていると仮定しても、あまり一般性を損わない。よって、濃縮・脱水システムが Seasonal Trend の仮定を満足するとすれば、

これは、同定に用いたデータ長を周期とする周期モデルとなる。よって、同定と同一月数に対する将来の入力変量、出力変量が予測できる。

いま、二種類の入力データ $\{X_i(\sum_{j=1}^k \tau_i + j)\}_{j=1}^{T_{k+1}}$ と $\{\bar{X}(\sum_{j=1}^k \tau_i + j)\}_{j=1}^{T_{k+1}}$ との間、周期 $\sum_{j=1}^{N+1} \tau_j$ の Seasonal Trend の仮定があれば、周期関数の性質から、 $\bar{x}(k) = x(\hat{k})$ となり、 $\{X_i(\sum_{j=1}^k \tau_i + j)\}_{j=1}^{T_{k+1}}$ のモデル式 $\bar{x}(\hat{k}+1) = g(\hat{k}, \bar{x}(\hat{k}))$ に対して、 $g = \bar{g}$ となり、同様に、 $\bar{y}(\hat{k}) = f(\hat{k}, \bar{x}(\hat{k}))$ についても、入力変量に対するモデル式の唯一性より、 $f = \bar{f}$ が成立する。よって、(7)、(8)は将来まで予測できる動的モデル式であると結論できる。次に、(7)、(8)が制御方程式となるためには、操作変量の決定が必要である。ところが、第8図にみられるように、 Y_1^* と X_4 の相関が大きく、 Y_2^*, Y_3^*, Y_4^* についても同様であるので、 X_4 (薬注率) を制御操作変量とできる。

Fig. 8 Correlation Scheme of Y_1



以上のことから制御方程式(状態と出力方程式)は、

$$x(\hat{k}+1) = g(\hat{k}, x(\hat{k}), u(\hat{k})) \quad (9)$$

$$y(\hat{k}) = f(\hat{k}, x(\hat{k}), u(\hat{k})) \quad (10)$$
と得られる。ここで、 $x \in R^3$, $u \in R^1$, $y \in R^4$ である。

4. おわりに

濃縮・脱水システムの静的モデル式を作る同定アルゴリズムを開発し、それに、動的特性を付与できることを明らかにした。また、操作変量の決定を、取得実データから行った。この結果の制御方程式を整形して、既存の制御理論によって、種々の制御問題が解き得る。

参考文献

- 1) G.E.P. Box, G.M. Jenkins; Holden Day Publishing Co., (1970)
- 2) 市田・吉本; 教育出版, 新しい応用の数学 20(199)
- 3) 池田他; システムと制御, 23/12~24/7(1979~1980)