

遺伝的アルゴリズムによるタンクモデル定数の同定

田中丸 治哉

PARAMETER IDENTIFICATION OF TANK MODEL WITH THE GENETIC ALGORITHM

By *Haruya TANAKAMARU*

Synopsis

The function optimization methods have been widely used to identify parameter values of rainfall-runoff models, but it is generally difficult to obtain a unique set of optimal parameters by using standard local-type search procedures because of the presence of multiple peaks in the response surfaces. In this study, the genetic algorithm is introduced to cope with such difficulties. The genetic algorithm is a search procedure based on the mechanics of natural selection and natural genetics and is efficient and robust. This procedure is applied to the parameter identification of the Sugawara's Tank model with sixteen parameters and its applicability is examined by numerical experiment. As a result, it is difficult to obtain the true optimum values of all of sixteen parameters, but in the identification of selected eight parameters, the parameter values near the global optimum are obtained successfully.

1. まえがき

菅原¹⁾の直列4段タンクモデルは、簡単な構造であるが物理的に無理がなく、非線形性もよく表現できること、さらに精度もよいことから、長期流出解析に広く利用されている。このモデルの最大の難点は、探索すべき定数が16個と多く、これらの同定が容易でないことがある。このような問題に対処するため、最適化手法あるいは自動化手法によるタンクモデル定数の同定に関する研究が進められてきた。

小林・丸山²⁾は、タンクモデル定数の探索問題を非線形関数の最小化問題に置き換え、Powellの共役方向法を適用して良好な結果を得ている。また、永井・角屋^{3,4)}は、Powell法、DFP法に基準化形式を導入したSP法、SDFP法を提案し、その有効性を示している。さらに渡辺ら⁵⁾は、解析的に求めた感度係数を導入したNewton法を提案し、タンクモデル定数が効率よく同定できることを示している。一方、菅原ら^{6,7)}は、流量をそれぞれの固有半減期を持ついくつかの成分に分解することにより1～3段目の流出孔と浸透孔の7定数を同定する自動化プログラムを開発している。

しかしながら、ほとんどの最適化手法には、局所的な準最適解に陥りやすいという難点がある。一方、タンクモデルなどの雨水流出モデル定数を同定する際の評価関数の応答面の形状は、概して単純ではなく、局所的な準最適解が多数存在していることが少なくない。このため、初期値の設定が適切でなければ、大域的な最適解がうまく求められない。最近、このような多峰性の問題に強い最適化手法として遺伝的アルゴリズム^{8,9)}(Genetic Algorithm, GA)が注目されている。

Holland¹⁰⁾によって提案された遺伝的アルゴリズム(以下GAと略記する)は、選択淘汰、交叉、突然変異といった生物進化の原理に基づくアルゴリズムであり、多点探索を行うという点で他の最適化手法とは大き

く異なっている。GAは、組合せ最適化問題などに対して極めて有効であることが確かめられており、GAの工学的な応用も各方面で進められつつある。水文・水資源学の分野でも、最近 Wang¹⁰⁾は、7定数の概念的雨水流出モデルの同定にGAを適用し、GAが大域的な探索法として有効であることを示した。また Duanら¹¹⁾は、GAそのものではないが、GAに類似した進化の概念を取り入れた新しい大域的最適化手法を提案し、これを6定数の貯留型雨水流出モデルの同定に適用して、その有効性を明らかにしている。

ここでは、GAによって直列4段タンクモデルの16定数ないしは8定数の同定を数値実験的に試み、GAの適応性と問題点について検討した結果を報告する。

2. 遺伝的アルゴリズム

2.1 遺伝子型の設定

関数 $F = F(x_1, x_2, \dots, x_n)$, $x_i^{\min} \leq x_i \leq x_i^{\max}$, $i = 1, 2, \dots, n$ を最小化する決定変数ベクトル $x = (x_1, x_2, \dots, x_n)$ を探索することを考える。GAでは、まず決定変数ベクトル x を N 個の記号を連ねた記号列によって表現する。記号列の内容は、数字、文字などどのようなものでもかまわないが、0と1のビット列が最もよく用いられており、本研究においてもこれを採用する。連続量を扱う場合は、 x_i の探索範囲 $[x_i^{\min}, x_i^{\max}]$ を 2^L 個の点に離散化して、各点に0から $2^L - 1$ までの番号を割り振り、その番号を2進数に変換して得られる L ビットの2進コードを用いる^{9),10)}。このとき n 次元決定変数ベクトル x は、 L ビットの2進コードを n 個連結した $N = n \times L$ ビットのビット列で表現される。この記号を遺伝子、記号列を個体の遺伝子型、これに対応して定まる変数 x の実体を表現型と呼ぶ。

2.2 遺伝的操作

GAでは、 m 個の個体の集合を考え、まず初期の個体集合をランダムに生成する。次いで、この個体集合を選択(selection), 交叉(crossover), 突然変異(mutation)の3種類の遺伝的操作によって新しい世代の個体集合に更新する。このような世代更新を繰り返す度により適応度が大きい個体(より最適解に近い変数)が増えるようにして、集合全体の適応度を大きくしていく。ここでは、適応度の評価および3種類の遺伝的操作について述べる。

(1) 適応度の評価

遺伝的操作の前に各個体の適応度 f を評価しておく必要がある。ここで考えているような関数の最小化問題では、関数値が小さい個体ほど適応度が大きくなるようになるが、これには適応度として関数の最大値からの差をとる方法や関数値の逆数をとる方法などがある⁹⁾。このとき適応度 f の値は非負とする。この他、関数値の順位によって適応度を評価する方法がある。

(2) 選 択

これは、適応度の大きさに応じた選択確率に従って、子孫を残すための親をランダムに選択するという操作である。最も一般的な方法は、選択確率が適応度の評価値に比例するとしたルーレット方式で、個体 j の選択確率 p_j として $f_j / \sum f_i$ を採用する。本研究では、関数値が最も小さい個体を1位として、関数値が小さいものから順に並べ、上位の個体ほど大きな選択確率を与える方法を採用した。この方法はランキング方式と呼ばれるものである。

また本研究では、選択に際してエリート保存戦略を導入した。これは集団の中で最も適応度の大きい個体を無条件で次世代に残す方法である。DeJongは、この戦略を次のように定義している^{9),12)}。 $a^*(t)$ を時間 t までに現れた最良の個体とし、もし個体集合 $A(t+1)$ を通常の方法で生成したときに、 $A(t+1)$ の中に $a^*(t)$ が存在しないならば、 $a^*(t)$ を $A(t+1)$ の $m+1$ 番目の個体として加える。これによると、その時点で最も良い解を交叉や突然変異で破壊せずに保存することができる。

(3) 交 又

交叉は、前述の操作によって2個体A, Bを選び、その特定部位を交換して新しい個体を作る操作である。1点交叉では、交叉位置 k (ただし $1 \leq k \leq N-1$)をランダムに決めて、個体A(B)の $1 \sim k$ 番目の遺伝子と個体B(A)の $k+1 \sim N$ 番目の遺伝子を連結して新しい個体A'(B')を作る。また交叉位置が複数のものを複数交叉という。たとえば2点交叉では、交叉位置 k_1, k_2 (ただし $1 \leq k_1 \leq k_2 \leq N-1$)をランダムに選択し、個体A(B)の $1 \sim k_1$ 番目、個体B(A)の $k_1+1 \sim k_2$ 番目、個体A(B)の $k_2+1 \sim N$ 番目の遺伝子を連結して新しい個体A'(B')を作る。以下に2点交叉の例を示す。本研究ではこの2点交叉を採用した。

$$\begin{array}{c|c|c} A = 10101 & 1110 & 01001 \\ B = 00011 & 0100 & 10101 \end{array} \xrightarrow{\hspace{1cm}} \begin{array}{c|c|c} A' = 10101 & 0100 & 01001 \\ B' = 00011 & 1110 & 10101 \end{array}$$

(4) 突然變異

突然変異は、遺伝子をある小さな確率で変化させる操作である。突然変異が無いときは、探索範囲が初期の遺伝子の組合せによる空間のみに限定されるが、適度な変異を加えることによって、集団中の遺伝子型の多様性を確保し、より広い空間を探索することができる。遺伝子型がビット列で表現されている場合は、任意のビット値をある確率で反転させる。すなわち、突然変異が生じた場合、元のビット値が 1 ならばこれを 0 に反転させ、元のビット値が 0 ならばこれを 1 に反転させる。

2.3 アルゴリズムの概要

本研究では、個体(n 次元決定変数ベクトル)の遺伝子型を先に述べたように $N = n \times L$ ビットのビット列で表現し、関数值を最小にする個体を次のような手順で探索する。このアルゴリズムは、先に述べたように選択においてランキング方式を、交叉において2点交叉を採用しており、エリート保存戦略を導入している点を除けば、Wang¹⁰⁾が採用しているアルゴリズムと大きな違いはない。

3. GA によるタンクモデル定数の同定

3.1 適用方法

(1) モデル定数の離散化

本研究では、菅原の直列4段タンクモデル (Fig. 1) の同定を試みる。初期水深を含めて16個ある未知定数の探索範囲は、これらの定数が一般的に取り得る値の範囲として、次のように設定する。

$$\left. \begin{array}{l} 0 \leq a_1 \leq 0.635, \quad 0 \leq a_5 \leq 0.0127, \quad 0 \leq Z_1 \leq 127, \quad 0 \leq S_1^o \leq 127 \\ 0 \leq a_2 \leq 0.635, \quad 0 \leq b_1 \leq 0.635, \quad 0 \leq Z_2 \leq 63.5, \quad 0 \leq S_2^o \leq 127 \\ 0 \leq a_3 \leq 0.635, \quad 0 \leq b_2 \leq 0.635, \quad 0 \leq Z_3 \leq 63.5, \quad 0 \leq S_3^o \leq 127 \\ 0 \leq a_4 \leq 0.127, \quad 0 \leq b_3 \leq 0.127, \quad 0 \leq Z_4 \leq 63.5, \quad 0 \leq S_4^o \leq 1270 \end{array} \right\} \quad (2)$$

ここに、 $a_1 \sim a_5$, $b_1 \sim b_3$, $Z_1 \sim Z_4$: 定数 (mm-d), $S_1^0 \sim S_4^0$: 初期水深 (mm)。

これらの定数は、 $L = 7$ として次のように離散化した。

$$\left. \begin{array}{l} a_1 = x_1 \times 0.005, \quad a_5 = x_5 \times 0.0001, \quad Z_1 = x_9 \times 1.0, \quad S_1^0 = x_{13} \times 1.0 \\ a_2 = x_2 \times 0.005, \quad b_1 = x_6 \times 0.005, \quad Z_2 = x_{10} \times 0.5, \quad S_2^0 = x_{14} \times 1.0 \\ a_3 = x_3 \times 0.005, \quad b_2 = x_7 \times 0.005, \quad Z_3 = x_{11} \times 0.5, \quad S_3^0 = x_{15} \times 1.0 \\ a_4 = x_4 \times 0.001, \quad b_3 = x_8 \times 0.001, \quad Z_4 = x_{12} \times 0.5, \quad S_4^0 = x_{16} \times 10.0 \end{array} \right\} \dots \quad (3)$$

ここに, $x_1 \sim x_{16}$: 整数, $0 \leq x_1, x_2, \dots, x_{16} \leq 2^7 - 1 = 127$

GA の適用に際しては、 $x_1 \sim x_{16}$ を 2 進数に変換して得られる 7 ビットの 2 進コードで表現し、この 2 進コードを 16 個連結した 112 ビットのビット列を個体の遺伝子型とする。

(2) 目的閾数

誤差評価関数には、次の最小 χ^2 誤差評価基準を用いる。

ここに, J_{xs} : χ^2 誤差評価基準, Q_o , Q_c : 観測流量, 計算流量, M : データ数。

また再現性の表現には、分かりやすいように次の相対誤差 J_{RE} (%) も併用する。

また次のような制約条件を考慮する。

$$\left. \begin{array}{l} a_1 + a_2 + b_1 \leq 1, \quad a_3 + b_2 \leq 1 \\ a_4 + b_3 \leq 1, \quad a_5 \leq 1, \quad Z_1 \geq Z_2 \end{array} \right\} \dots \quad (6)$$

第1～4の条件は、各タンクの流出孔、浸透孔の和が1を超えないという条件であり、第5の条件は、第1段タンクの流出孔 a_1 と a_2 の底面からの高さが逆転しないという条件である。これらの制約条件の導入に際しては、制約条件が満たされない場合に目的関数が著しく大きくなるようなペナルティ関数を用いる。たとえば、第1の条件に関するペナルティ関数 J_{P1} として次のものを用いる。

$$J_{P1} = 0, \quad a_1 + a_2 + b_1 \leq 1 \\ = (a_1 + a_2 + b_1) \times 10^3, \quad a_1 + a_2 + b_1 > 1 \quad \left. \right\} \dots \quad (7)$$

結局、最小化すべき目的関数 J は次式で表される。

$$\left. \begin{aligned} J &= J_{XS} && ; \text{制約条件が満たされた場合} \\ &= J_{P1} + J_{P2} + J_{P3} + J_{P4} + J_{P5} && ; \text{制約条件が満たされない場合} \end{aligned} \right\} \quad (8)$$

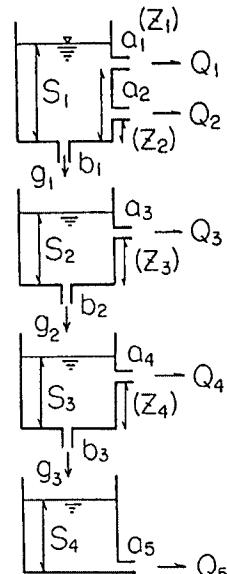


Fig. 1. Tank model.

ここに、 $J_{P1} \sim J_{P5}$: (6) 式の第1～5の制約条件に対応するペナルティ関数。

(3) GA のパラメータの設定

GA の適用に際しては、個体数 m および突然変異の発生確率 P_m を定めておく必要があるが、これらのパラメータは、問題に応じて設定しなければならない。個体数は、広い範囲を探索するためには多い方がよいが、目的関数の評価回数が個体数に比例することから計算時間を短くするためには少ない方がよい。一方、突然変異の発生確率は、集団中の遺伝子型の多様性を確保するためにはある程度の大きさが必要であるが、大きすぎると優れた個体が生まれても突然変異によってことごとく破壊されてしまうことになる。ここでは、これらのパラメータを適当に変えて試行的な計算を数回行い、その結果に基づいて個体数を $m = 200$ 、突然変異の発生確率を $P_m = 0.01$ とした。また世代更新の繰り返し計算を終了させる予定世代数 T は 100 世代とした。 T 世代まで世代更新を行うときの目的関数の累加評価回数は $m(T+1)$ 回であるから、ここでは 100 世代までに目的関数を 20,200 回評価することになる。

3.2 適用結果

(1) 解析資料

解析に用いた資料は、滋賀県愛知川上流に位置する永源寺ダム流域 (131.5 km^2) の流域平均日降水量と月平均日蒸発量で、解析期間は1990年の1年間とした。流域平均降水量は、流域を標高別に4地帯分割し、流域内6雨量観測点の降水量資料を用いて標高別降水量を求め、これを地帯面積で加重平均して算定した¹³⁾。蒸発量は、日射量法の一つである Makkink 式で推定した。この式の適用に際しては、流域に近い彦根気象台の月平均気温（流域平均標高の気温に換算したもの）および月平均全天日射量を用いて、月平均日蒸発量を求めた¹⁴⁾。なお蒸発散量の推定に際しては、無降雨日の蒸発散量は月平均日蒸発量と同じとし、降

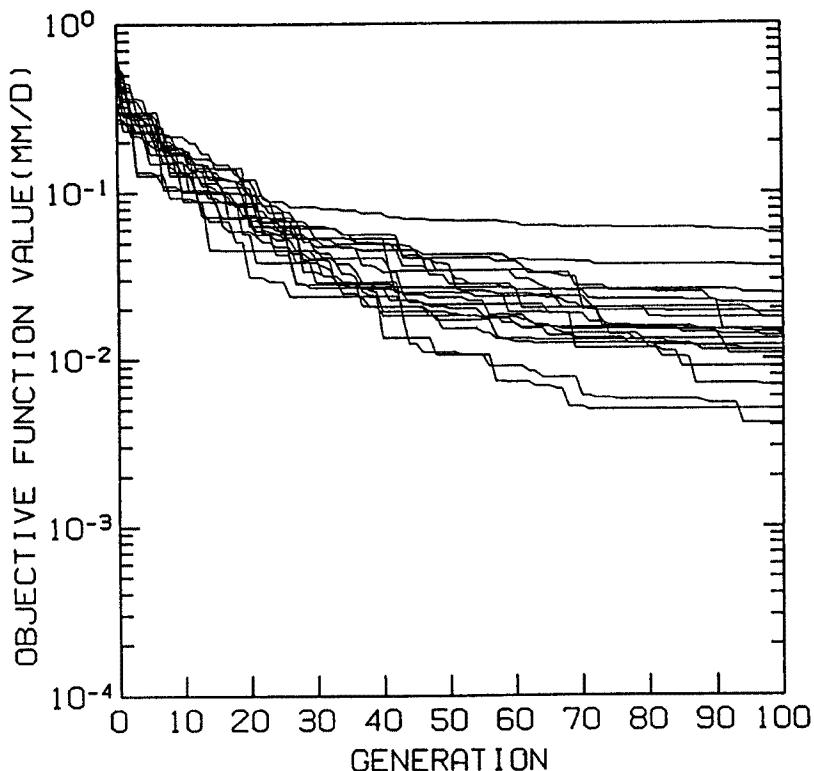


Fig. 2. Objective function values plotted against generations for the twenty runs of identification of sixteen parameters.

雨日の蒸発散量は0とした。

流量資料には、あらかじめ適当な定数を設定したタンクモデルに先の降水量と蒸発量を入力して得た計算日流出高を用いることにした。ここで設定した定数は、探索すべき真値に相当する。この場合、完全な真値が同定されれば、目的関数値は0になる。

ここでは、真値として $(x_1, \dots, x_{16}) = (60, 40, 20, 20, 40, 40, 10, 20, 40, 30, 20, 20, 20, 20, 20, 50, 30)$ を採用した。以下の検討では、タンクモデル定数の表示には、離散化された変量である(3)式の (x_1, \dots, x_{16}) の値を用い、これらの値を定数値と呼ぶ。これらの値は、(3)式によって通常の定数値に変換される。

(2) 16定数の同定結果

まずGAにより16定数の同定を試みた。Fig. 2は、各個体の目的関数値の内、最小のものを世代ごとに示したもので、乱数の種を変えた20回分の試行結果を示している。これによると、世代が進むにつれて目的関数値が減少していくことが分かる。Fig. 3は、100世代目に得られた定数値であり、先の20回分の試行結果を示している。この図では、 \times が同定された定数値、縦線が真値を表しており、 \times が縦線の近傍にあれ果を示している。この図では、 \times が同定された定数値、縦線が真値を表しており、 \times が縦線の近傍にあれば、真値に近い定数が同定されたことになる。これによると、真値に比較的近い値が求められている定数 $(x_1, x_2, x_6, x_9, x_{10}, x_{13})$ は、いずれも第1段タンクの定数である。一方、第2～4段タンクの定数の多くは、試行ごとに値が大きくばらついている。とくに $x_4(a_4)$ と $x_7(b_2)$ は、多くの場合真値とはかなり異なる値が

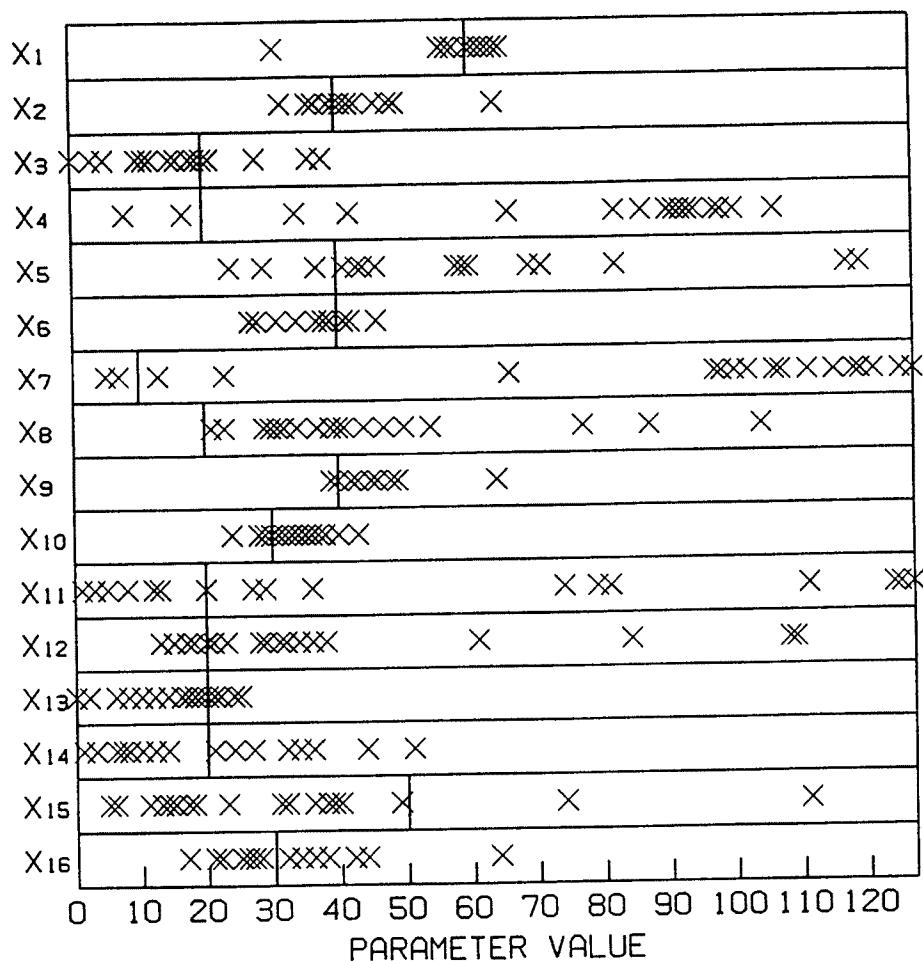


Fig. 3. Parameter values on integer interval for the twenty runs of identification of sixteen parameters.

同定されている。

20回の試行の内、100世代目の目的関数値が最小となった個体の定数値は $(x_1, \dots, x_{16}) = (58, 43, 3, 98, 71, 38, 119, 47, 40, 30, 81, 15, 21, 8, 14, 26)$ であった。この個体でも、第1段タンクの定数は、ほぼ真値に近いものが同定されているが、他の多くの定数は、真値とはかなり異なっている。しかしながら、この個体の日流出高の相対誤差は2.5%であって、再現性は悪くない。20回の試行によって得られた個体の日流出高の平均相対誤差は4.3%であった。

この相対誤差をみると、同定された定数値と真値がかなり異なる場合でも、これらの与えるハイドログラフには、それほど大きな差はないようである。これは、タンクモデル定数がかなり鈍感である上、定数間に相互関連性があるためで、この性質のためにタンクモデルの解の一意性は得にくくなっている⁶⁾。

(3) 8定数の同定結果

次いで、流出孔と浸透孔の8定数のみを他の定数は既知として同定した。Fig. 4 は、各個体の目的関数値の内、最小のものを世代ごとに示したもので、Fig. 5 は、100世代目に得られた定数値であり、いずれも20回分の試行結果を示している。Fig. 5 によると、 $x_8(b_3)$ がやや大きめになっているが、他の7定数は、多くの場合ほぼ真値に近いものが同定されている。なお x_4 と x_7 については、20回の試行の内の2~3回は真値とかなり異なる値が同定されているが、16定数の同定結果に比べれば、真値に近い値を得る確率はかなり高いといえる。

20回の試行の内、100世代目の目的関数値が最小となった個体の定数値は $(x_1, \dots, x_8) = (59, 40, 20, 19, 45, 40, 10, 28)$ で、ほぼ真値に近い定数値が求められている。この個体の日流出高の相対誤差は0.6%であった。また20回の試行によって得られた個体の日流出高の平均相対誤差は2.0%であった。

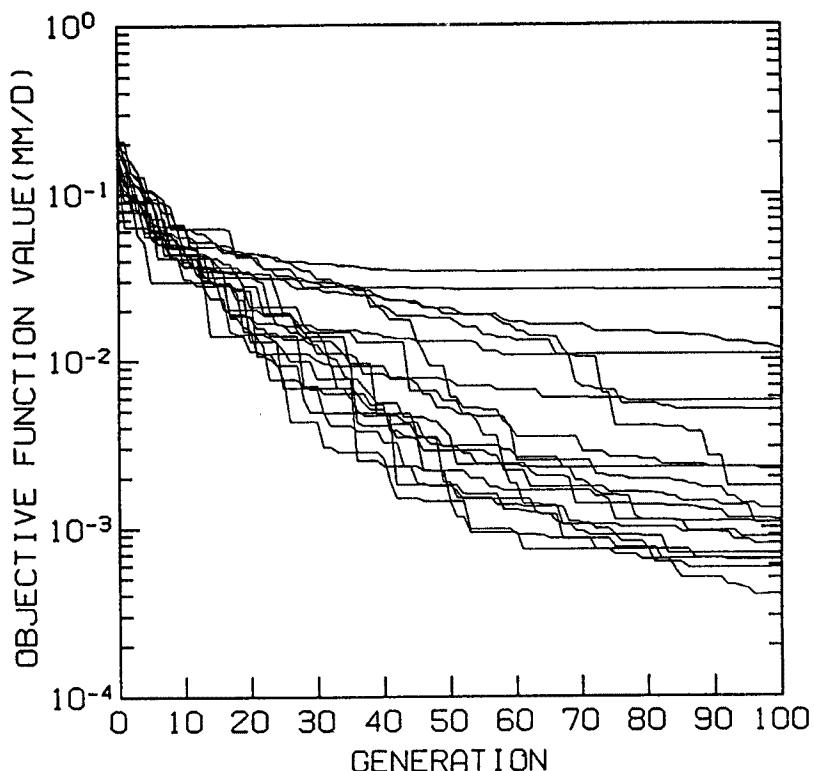


Fig. 4. Objective function values plotted against generations for the twenty runs of identification of eight parameters.

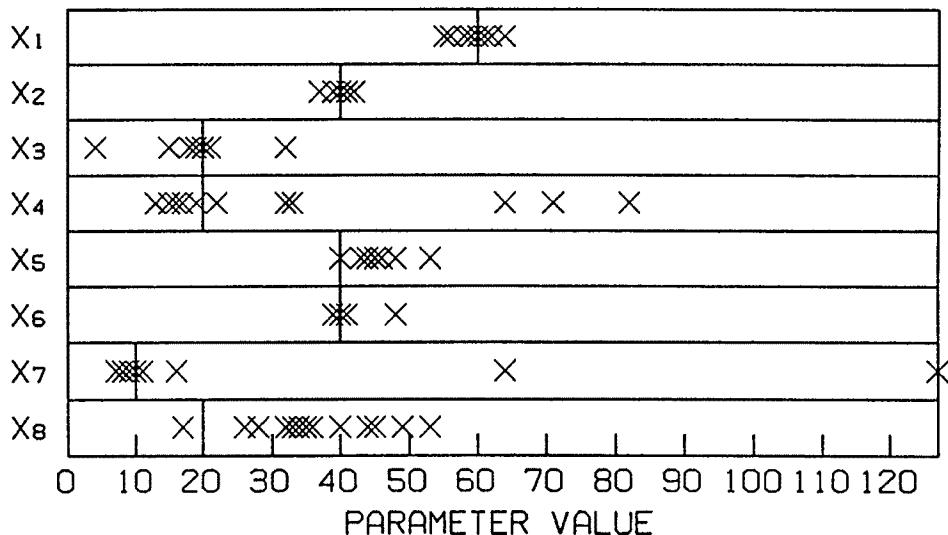


Fig. 5. Parameter values on integer interval for the twenty runs of identification of eight parameters.

3.3 GA の問題点についての考察

前節では、GA を直列4段タンクモデルの同定に適用した。この結果、16定数を全てうまく同定するには到らなかったが、探索する定数が少ないときは、ほぼ真値に近い定数を得ることができた。これより GA がタンクモデルなどの雨水流出モデル定数の同定に十分活用しうることが分かったが、いくつかの問題点も明らかになった。ここでは、GA をタンクモデル定数の同定に適用する際の問題点について考察する。

第一の問題点は、探索すべき定数が多くなると、多峰性の問題に強いとされている GA といえども局所的な準最適解に陥りやすくなることである。たとえば、先の適用例では x_4 と x_7 の真値はそれぞれ 20, 10 であったが、16定数を全て同定した場合、 x_4 の同定値の多くは 90~100 の付近に、 x_7 の同定値の多くは 100~120 の付近に集まっており、明らかに局所解に陥っている。局所解への陥りやすさは、個体の選択確率の与え方と密接に関連している。一般に適応度の大きい優れた個体と適応度の小さい劣った個体との選択確率の差が大きいほど局所解に陥る可能性が高くなり、この差が小さいほど局所解に陥りにくい。しかしながら、選択確率の差を小さくすると、進化の速度が遅くなり探索効率が悪くなるから、一概にこの差を小さくするほどよいとはいえない。すなわち、局所解に陥らないようにしながら、ある程度の探索効率を維持するように選択確率を調節しなければならない。さらに GA の適応性は、すでに述べたように個体数や突然変異の発生確率の与え方によっても変化する。これらのパラメータの与え方は、これから検討を要する課題である。

第二の問題点は、GA によると近似解を求めることはできても、厳密解がなかなか求められることである。たとえば、先の適用例では、8定数を同定した場合には、ほぼ真値に近い定数が得られたが、完全な真値を得るには到っていない。これは、GA が局所探索の機能を持っていないためであり、これが探索法としての GA の弱点とされている。このため、GA の大域的な探索の機能を生かしながら、近似解から厳密解へと同定精度を向上させるためには、GA に局所探索の機能を組み込むかあるいは GA とこれまでの局所探索型の最適化手法を組み合わせるなどの工夫が必要と思われる。たとえば Wang¹⁰⁾ は、雨水流出モデル定数の同定において GA による初期探索の後、局所探索型の最適化手法を適用して解の精度を高める方法を提示している。

4. あとがき

本研究では、遺伝的アルゴリズム（GA）によって直列4段タンクモデルの16定数ないしは8定数の同定を数値実験的に試み、その適応性について検討した。この結果、まだ16定数を全てうまく同定するには到っていないが、探索する定数を流出孔と浸透孔の8定数に絞ったときは、ほぼ真値に近い定数を得ることができた。一方、GAをタンクモデル定数の同定に適用する際のいくつかの問題点も指摘した。今後、これらの問題点を解決していくことによって、GAの適応性をより向上させることができると考えている。

GAには、アルゴリズムが簡単なことや目的関数の性質を全く気にせずに最適化できることなどの利点があり、GAは、雨水流出モデル定数の同定だけでなく、水文・水資源学の分野における様々な問題に応用できるものと思われる。

参考文献

- 1) 菅原正巳：流出解析法，共立出版，1972.
- 2) 小林慎太郎・丸山利輔：Powellの共役方向法によるタンクモデル定数の探索，農業土木学会論文集，第65号，1976, pp. 42-47.
- 3) 永井明博・角屋 瞳：流出モデル定数の最適化手法，京都大学防災研究所年報，第22号B-2, 1979, pp. 209-224.
- 4) 永井明博・角屋 瞳：タンクモデルの最適同定法に関する基礎的検討，京都大学防災研究所年報，第23号B-2, 1980, pp. 239-247.
- 5) 渡辺和好・館谷 清・松木賢治・星 清：タンクモデル定数の最適化手法の改良，第33回水理講演会論文集，1989, pp. 55-60.
- 6) 菅原正巳・尾崎睿子・渡辺一郎・勝山ヨシ子：タンク・モデルの構造を自動的に定める計算機プログラムの開発（第1報），国立防災科学技術センター研究報告，第17号，1977, pp. 43-89.
- 7) 菅原正巳・渡辺一郎・尾崎睿子・勝山ヨシ子：タンク・モデルの構造を自動的に定める計算機プログラムの開発（第2報），国立防災科学技術センター研究報告，第20号，1978, pp. 157-216.
- 8) Holland, J. H.: Adaptation in Natural and Artificial Systems, University of Michigan Press, 1975.
- 9) Goldberg, D. E.: Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley, 1989.
- 10) Wang, Q. J.: The Genetic Algorithm and Its Application to Calibrating Conceptual Rainfall-Runoff Models, Water Resources Research, Vol. 27, No. 9, 1991, pp. 2467-2471.
- 11) Duan, Q., Sorooshian, S. and Gupta, V.: Effective and Efficient Global Optimization for Conceptual Rainfall-Runoff Models, Water Resources Research, Vol. 28, No. 4, 1992, pp. 1015-1031.
- 12) 北野宏明：遺伝的アルゴリズム，人工知能学会誌，Vol. 7, No. 1, 1992, pp. 26-37.
- 13) 角屋 瞳・田中丸治哉・永井明博・金口正幸：愛知川上流域への長短期流出両用モデルの適用と実時間洪水予測，水資源研究センター研究報告，第9号，1989, pp. 45-60.
- 14) 角屋 瞳・永井明博：長短期流出両用モデルの永源寺ダム流域への適用と考察，農業土木学会論文集，第137号，1988, pp. 71-78.