

遺伝的アルゴリズムによる水質モデルのパラメータ推定に関する研究

西田 渉*・松下 紘資**
野口 正人*・Dimitri P. Solomatine***Estimation of Model Parameter of Water Quality Simulation
Model using Genetic Algorithm

by

Wataru NISHIDA*, Hiroshi MATSUSHITA**
Masato NOGUCHI* and Dimitri P. SOLOMATINE***

Genetic algorithm is one of the well-known methods to solve global optimization problems and its use is becoming widespread in many branches of engineering. In this study, this optimization method was made use in order to calibrate the water quality simulation model, and its applicability to this kind of problem was discussed. The results obtained from calibrating the water quality model show that the error defined by calculated and observed results gradually decreases as the proceed of alternation of generation, and genetic algorithm succeeded to find the parameter set which makes the error much smaller than by other optimization method. Although several implementations for the practical use of this algorithm, it might be concluded that genetic algorithm has some extent of applicability to the automatic calibration of numerical simulation model.

1. はじめに

実水域の環境管理を適切に行うには、水域内部での物質の循環機構を評価しておくことが必要とされる。現在、この実際的な評価と予測に、数値シミュレーション手法が有効な手段の一つとなっていることは間違いなさそうであるが、その一方で、数値シミュレーションから得られる結果の信頼性を向上させるには、数値モデルそれ自身が実際に水域で生じている現象を矛盾なく表現していることと共に、物質の変化速度等を規定する係数（モデルパラメータ）が適切に評価しておくことが必要とされる。水質モデルについては、物質循環に関する知見の充実に伴って、現在では、水中だけではなく、底泥中の生物生産をも考慮したモデルの構築がされている。ただ、物質の多様な循環過程を表現しようとするモデルには、必然的に相応の数のモデルパラメータが含まれており、モデルの利用に際

してパラメータの評価が不可欠となる。著者らは、数値モデルの有効活用を一層進めるには、パラメータの同定にモデルの精緻性と同レベルの技法が導入されるべきであると考え、従来、手動で試行錯誤的に行われることの多かった最適パラメータ値の推定に最適化手法を導入することで同定手続きの適正化と自動化とを図ってきた。すなわち、パラメータの同定手続きを、観測結果と計算結果との差（エラー）を最小化させる制約条件付の最適化問題として捉え、Solomatineによって提案されている大域的最適化手法：*Adaptive Cluster Covering Method*（ACCO手法）を水質モデルの検定に適用している^{1,2)}。

ところで、大域的最適化には、ACCO手法以外にも幾つかの手法が提案されている。本研究では、その代表的な手法とも言える遺伝的アルゴリズム（*Genetic Algorithm*（GA））を用いて水質モデルのモデル

平成16年4月16日受理

*社会開発工学科（Department of Civil Engineering）

**日建工学株式会社（Nikken Kogaku Co., Ltd）

***UNESCO IHE for water education, Department of hydroinformatics and knowledge management

パラメータの推定を試みることにし、その有効性を ACCO 手法による結果と比較しながら検討することにした。

2. モデルの検定とその手法

数値シミュレーションモデルの適用性の良否は、実測結果の再現性から判断され、モデル検定の手続きの中で、パラメータはエラーを最小化させるようチューニングされることになる。モデルパラメータの数が少ない場合、最適なパラメータセットの推定は、比較的容易に思えるが、数値モデルの構成によっては、エラーの分布が不連続、且つ、多峰性となることもある^{3,4)}。本研究で用いる GA は、探索空間の全域を対象に多点探索する大域探索を得意としており、従来の最適化手法のように、局所解に収束してしまう危険性が少ないことに特徴がある^{5,6)}。GA では、探索点が個体として取り扱われており、探索点の情報は各個体の遺伝子に記録される。この個体に対して、図-1 に示される淘汰、交叉、突然変異といった遺伝的オペレータを作

用させることで、適者生存の進化過程をシミュレートしている。本研究における最適化の手続きは、一般的な SGA に従って構成されており、数値モデルの検定は以下のとおりになっている。

- ①：遺伝子型を決定する。
- ②：各パラメータの探索可能領域からランダムにパラメータ値を抽出し、遺伝子型へ変換する。この処理を繰り返し、個体群を形成する。
- ③：全個体の情報を用いたシミュレーションが実施され、適合度が評価される。
- ④：個体の適合度を基ついて、淘汰、交叉、突然変異の処理がされ、次世代の個体が生成される。

これらの処理が計算終了の条件を充足するまで繰り返される。今回の計算では、1つのパラメータを8ビットの2進数で表現することとした。また、モデルの適合度を表す評価関数は、式(1)に示されるとおりであり、エラーの2乗値総和の逆数を用いて評価された。なお、水質モデルの検定では、複数の水質指標のエラーが同時に評価されているため、水質指標ごとの重みも導入されている。

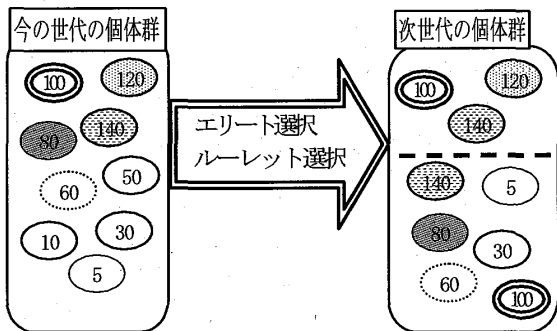
$$SE_{GA} = \frac{1}{\sum_{n=1}^N \sum_{m=1}^M w_{ind,n} w_{obs,m} |OBS_{m,n} - CAL_{m,n}|^2} \quad (1)$$

ここで、 N ：水質指標の総数、 M ：観測結果の総数、 $w_{ind,n}$ ：水質指標 n に対する重み係数、 $w_{obs,m}$ ： m 番目の観測結果に対する重み、 $OBS_{m,n}$ 、 $CAL_{m,n}$ ：観測結果、計算結果、である。

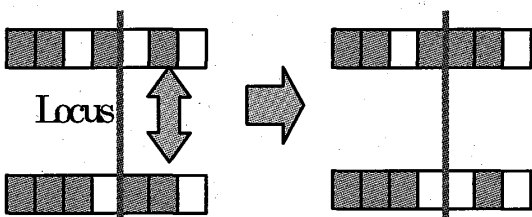
3. 水質シミュレーションモデルの構築

GA によるモデルパラメータの推定がなされる水質モデルは、本研究室で開発されてきたリンの循環型モデルであり⁷⁾、その概要は図-2のとおりである。

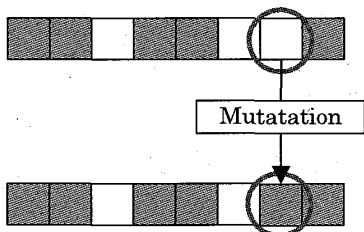
本モデルでは、水中のリンを有機態リン (OP) と無機態リン (IP) とに大別した上で、OP並びにIPをそれぞれ懸濁態物質、溶存態物質として取り扱っている。なお、OPについては、藻類の増殖活動が水温によって異なることに注意して、3種類：珪藻類(OP1)、その他の藻類 (OP2)、無生物体 (OP3) に分類したモデル化がなされている。水域でのリンの循環として、図に示される藻類の枯死、OP の沈降・無機化、IPの底泥からの溶出等の効果が考慮されている。OP1、OP3、IP の収支式は、それぞれ式(2)、(3)、(4)のとおりである。OP2 は OP3 に類似した形となっている。



(a) 選択オペレータ



(b) 交叉オペレータ



(c) 突然変異オペレータ

図-1 遺伝的オペレータ

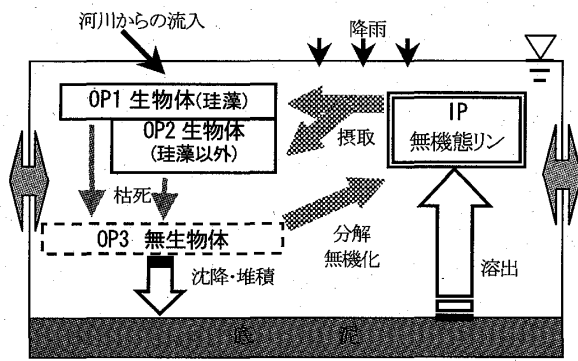


図-2 水質モデルの概念図

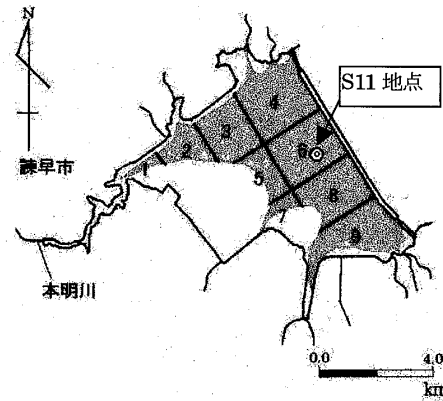


図-3 調整池の概要と水平方向の分割

OP1 の収支

$$Vol \frac{\partial OP1}{\partial t} = F_{OP1} + R \cdot \gamma_{Chl-P} \cdot \alpha_{OP1} \cdot Chl_{OP1} \cdot Vol - \gamma_{Chl-P} \cdot d_{OP1} \cdot Chl_{OP1} \cdot Vol \quad (2)$$

OP3 の収支

$$Vol \frac{\partial OP3}{\partial t} = F_{OP3} + (d_{OP1} \cdot Chl_{OP1} + d_{OP2} \cdot Chl_{OP2}) \gamma_{Chl-P} \cdot Vol - k \cdot OP3 \cdot Vol - w_{OP} \cdot OP3 \cdot Area \quad (3)$$

IP の収支

$$Vol \frac{\partial IP}{\partial t} = F_{IP} - R \cdot \gamma_{Chl-P} \cdot (\alpha_{OP1} \cdot Chl_{OP1} + \alpha_{OP2} \cdot Chl_{OP2}) \cdot Vol + k \cdot OP3 \cdot Vol + \beta \cdot Area \quad (4)$$

ここに、 Vol ：水の体積、 F_{OP1} 、 F_{OP3} 、 F_{IP} ：OP1、OP3、IP の移流・拡散量、 $Area$ ：水底面の面積、 α_{OP1} 、 α_{OP2} ：OP1、OP2 に相当する藻類の増殖速度、 d_{OP1} 、 d_{OP2} ：各藻類の枯死速度、 R ：藻類の IP の摂取速度、 γ_{Chl-P} ：リン-クロロフィル-a (Chl-a) 比、 Chl_{OP1} 、 Chl_{OP2} ：OP1、OP2 における Chl-a 濃度、 k ：分解速度、 w_{OP} ：OP3 の沈降速度、 β ：IP の底泥からの溶出速度、である。

上述のモデルを用いた水質シミュレーションは、諫早湾調整池を対象に実施している。計算領域については、鉛直方向には単層として取り扱い、水平方向には図-3に示される9つの水域に分割している。この調整池へ流入する河川については、本明川を始めとした合計12本が考慮されている。計算時間間隔は1時間である。モデルの検定に用いた水質の測結果は、九州農政局諫早湾干拓事務所によって報告された実測値とした。具体的には、図-3に示される調整池のほぼ中央部 (S11地点) での値であり⁸⁾、1997年4月から2002年4月までの期間に計測された IP と OP について各々253個である。

4. 検定手法の計算条件

ここで、水質モデル検定に最も適した GA のパラメータを定めるために、幾つかの条件で予備計算を実施した。なお、予備計算の対象とした水質モデルは、水中のリンを IP と OP との2種類に大別したものである。

予備計算では、先に、個体の選択・淘汰処理方法であるエリート保存選択の有効性と、優良個体の保存確率とが検定結果に及ぼす影響を調査した。次いで、交叉オペレータに関連して、交叉点の数と交叉確率の計算結果に与える影響が、また、突然変異オペレータに関連して、突然変異確率と逆位方式の有効性が評価された。この他に初期個体数と最大世代数についての評価も行っている。

ここで、交叉確率と突然変異確率の影響について記しておく。予備計算から求められた最大適合度を示すと図-4のとおりである。この図をみると、 $Pc=0.7$ 、 $Pm=0.03$ 付近で適合度が最良となっているのが分かる。ただ、同図から、 Pc と Pm の組み合わせによっては、幾つかの局大値があることも示されており、兩

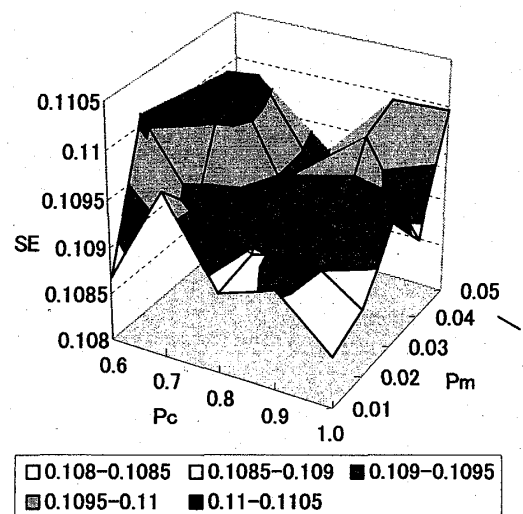


図-4 交叉確率、突然変異確率の違いによる最大適合度の変化

確率の設定条件についても、相応の注意が必要であることが分かる。上記の事項に対して行われた予備計算から求められた最良の設定条件を示すと表-1のとおりである。なお、検定計算の終了判定条件については、設定した値以上の適合度を持つ個体が発生した時や、個体群の平均適合度の値が設定値を超えた時、等が考えられるが、本研究においては、GAの適用性を探ることを目的としたことから、世代更新の回数を計算終了の条件とした。また、GAのモデルパラメータ推定への適用性については、ACCO手法による検定結果との比較に依ることとする。ここにACCO手法の詳細は記さないが、本手法の計算条件として、抽出される試行点の数(Nsmp)と形成されるクラスターの数(Mclr)とを、以前に行われた条件を参考に、Nsmp=200, Nclr=4としている。計算の停止条件については、パラメータの探索領域が、初期に与えた領域の5

表-1 GAの計算条件

| | |
|----------|---------------------|
| 抽出個体数 | 200 |
| 選択方法 | ルーレット選択とエリート保存選択の併用 |
| エリート保存確率 | 0.1 |
| 交叉点 | 1点交叉 |
| 交叉確率 | 0.7 |
| 突然変異方法 | 単一遺伝子座突然変異 |
| 突然変異確率 | 0.03 |
| 最大世代数 | 500世代 |

%以下となった時に最終的な探索を実施することとした^{1,2)}。

5. 水質モデルの検定と ACCO 手法との比較

検定の対象となる水質モデルには、14個のモデルパラメータがあり、各パラメータの初期制約条件は、文献値等を参考に表-2のとおりに与えられた。

ここで、GAとACCO手法で評価された計算回数に対するSEの変化を示すと、図-5、図-6のとおり

表-2 初期制約条件

| min | パラメータ | max |
|------------------------|----------------------------|-----------------------|
| 0.05 | K_I : 日射の半飽和定数 | 0.70 |
| 0.75 | α_0 : 増殖速度 | 2.00 |
| 0.02 | D : 光の減衰係数 | 0.75 |
| 0.001 | K_p : 半飽和定数 | 0.075 |
| 0.005 | w_{OP} : リンの沈降速度 | 0.04 |
| 0.60 | R : 摂取速度 | 1.20 |
| 0.02 | K : 分解速度 | 0.125 |
| 0.0075 | β_0 : 溶出に関わる係数 | 0.03 |
| 1.50×10^{-4} | τ_{CP} : 無次元限界掃流力 | 2.50×10^{-4} |
| 1.00×10^{-4} | λ_0 : SSの巻上げに関わる係数 | 3.0×10^{-4} |
| 1.00×10^{-7} | w_{SS0} : SSの沈降速度 | 1.00×10^{-6} |
| 3.00 | A_I : 沈降に関わる係数 | 10.0 |
| 0.80 | U_C : 沈降に関わる流速 | 1.20 |
| 7.724×10^{-5} | τ_{*C} : SSの無次元掃流力 | 2.00×10^{-4} |

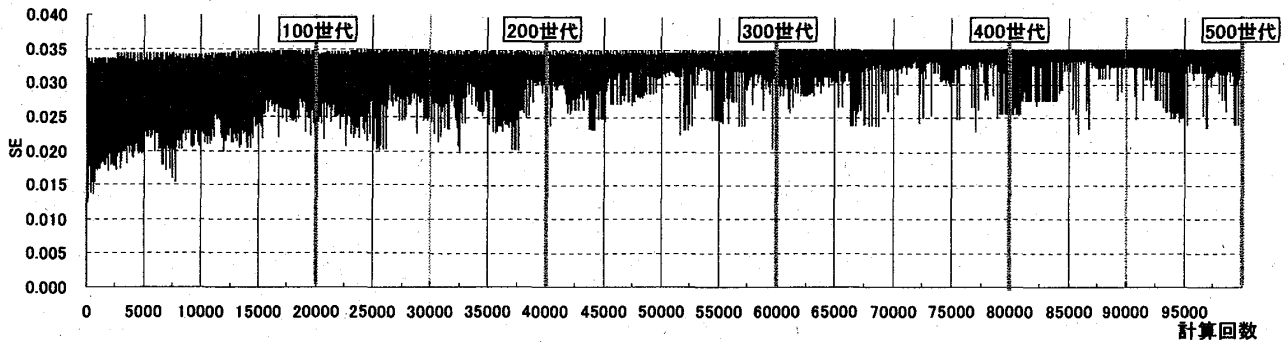


図-5 GAでの計算回数に対するSEの変化

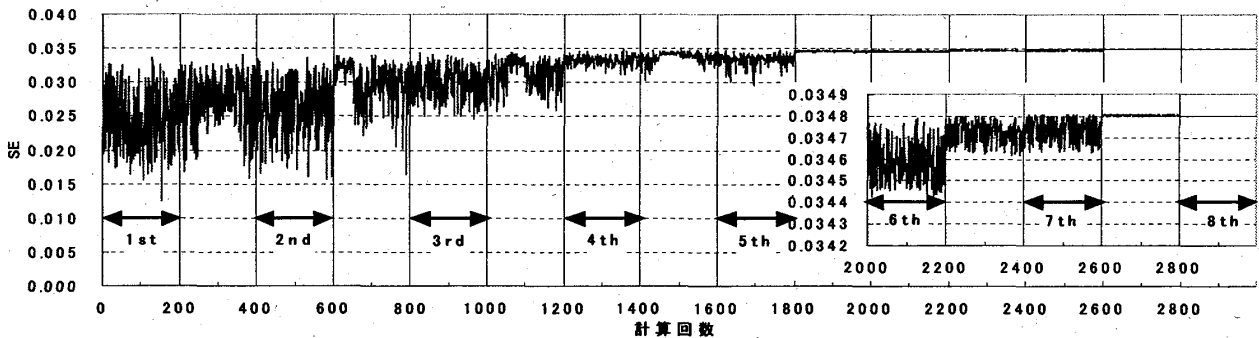


図-6 ACCO手法での計算回数に対するSEの変化

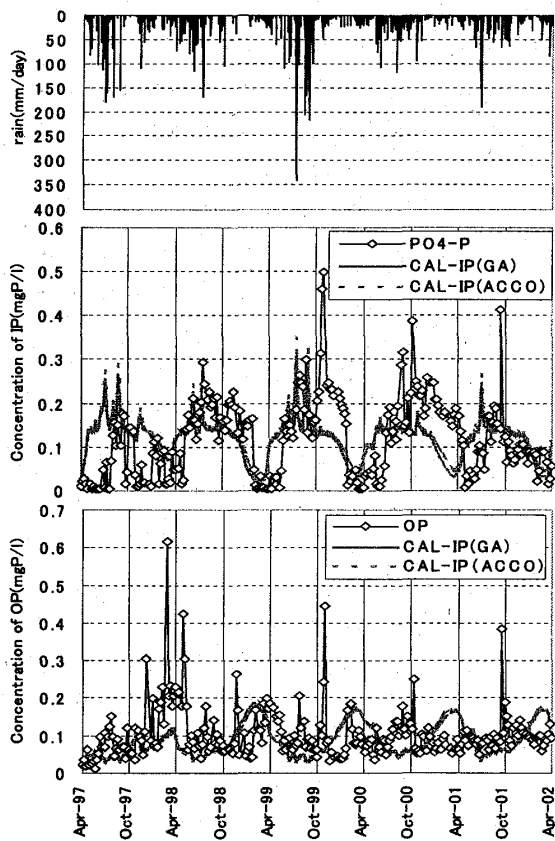


図-7 計算結果と観測結果
(上：降雨，中：IP，下：OP)

りである。また、図-7には、GAによって求められた最大適合度を示すパラメータセットと、ACCO手法によって最終的に探索されたパラメータセットを使って予測された、IPとOPの計算結果が観測結果と共に記されている。

今回の水質モデルのように、モデルパラメータが多数存在する場合、手動によって最適なパラメータの値を探し出し、感度解析を行おうとすると、それに要する時間と労力は莫大なものになることが推察されるが、今回行ったGAによる検定計算に要した時間は、3日間程度であった。このことから、GAを用いて水質モデルの検定を行うことで、試行錯誤的な手法に比べて、短時間の内に最適なパラメータセットの探索が可能になるものと思われる。

ここで図-5と図-6を見ると、いずれの手法においても、適合性を表すSEは、計算開始の直後に上下に大きく変動しており、その後、検定計算が進展するにつれて、次第に高い値へと収束していくことが分かる。このように一群体群のSEの分布が、検定計算の初期と後期とで大きく異なるのは、計算開始段階には、初期条件として設定されたパラメータの領域から、ラ

ンダムに抽出された値を使ってパラメータセットが生成されるために、再現性のあまり良くないパラメータの組み合わせも生成されるためと考えられる。その後、検定計算が進行するにつれて、SEが高い値に収束していき、GAにおいては、淘汰や交叉等の各進化プロセスによる最適解の探索が、的確に進められたものと推察される。同様に、ACCO手法においても、最適解付近へのクラスター形成、探索領域の絞込み、等の探索手続きが、計画どおりに進められたものと考えられる。

図-7に示される計算結果と観測結果との比較を見ると、IPの計算結果は、観測結果とほぼ同様の変化を示しており、冬季にかけて減少し、春季に再び増加する傾向などが再現されていることが分かる。一方、OPに関しては、観測結果と同様の変化を示すところもあるが、IPの結果に比べると、差が大きいところが多いようである。このようにIPとOPで予測の精度に違いが現れたのは、適合性の評価関数に物質毎の重みを考慮してはいるものの、OPのエラーがIPのそれよりも総合的な適合性の評価に反映され難かったことが原因の一つと考えられる。また、シミュレーションモデルに関連しては、計算期間に調整池で生じた地形変化が計算条件に考慮されていないことも、影響したものと推察される。ここで、ACCO手法による計算結果をGAによる結果と比較すると、両者に然程の違いはなく、終始同様の変化を示している。ただし、最大適合度を比較すると、GAでは0.03536であったのに対して、ACCO手法では0.03480であり、GAによる適合度の方が良好な結果となっている。

6. おわりに

本研究では、GAの水質モデルのパラメータ推定への適用性について検討した。その結果、GAでの計算結果は観測結果とほぼ同様の変化を示したことや、最大適合度が他の手法よりも良好な結果となったこと等から、GAの水質モデル検定への適用性がある程度示されたものと考えている。今後は、GAに採用される進化のプロセスについて他の手法の有効性を調査したい。また、モデルの検定が自動化されたとはいえ、図-5に示されるように、世代交代の回数だけを計算の終了条件とすると結果的に不要な計算を行うことになるため、実用的な適用を目指して、更に検討を進めたいと考えている。

謝辞 本研究で用いた実測値は、諫早湾締切干拓地域でモニタリングされた値であり、九州農政局、長崎県生活環境部環境政策課、諫早湾環境監視機構の関係各位に記して深謝申し上げます。研究を遂行するにあたり、データの取り纏め等にご協力頂いた松田良寛氏を始めとする本学河川工学研究室の皆様にご感謝の意を表します。

参 考 文 献

- 1) Noguchi, M., Solomatine, D.P. and Nishida, W. : "Calibration of water quality model by global optimization techniques", Proc. of 5 th Int'l Conf. on Hydroinformatics, Vol. 1, pp.464-469, 2002.
- 2) 西田 渉, 野口正人, D.P.Solomatine : "大域的最適化手法を用いた水質モデルの自動検定に関する研究", 水工学論文集, 第47巻, pp.1267-1272, 2003.
- 3) Wang, Q. J. : "The genetic algorithm and its application to calibrating conceptual rainfall-runoff models", J. Water Resources Research, Vol. 27, No.9, pp.2467-2471, 1991.
- 4) Solomatine, D. P. and Dibike, Y. B. : "Automatic calibration of groundwater models using global optimization techniques", J. Hydrological Sciences, 44 (6), pp.879-894, 1999.
- 5) 坂和正敏, 田中雅博 : "遺伝的アルゴリズム", 朝倉書店, p.194, 2000.
- 6) 北野宏明編 : "遺伝的アルゴリズム", 産業図書, p.447, 2000.
- 7) 仁木将人, 西田 渉, 野口正人, 橋本篤史 : "諫早調整池における水質変化の予測とその評価に関する研究", 水工学論文集, 第43巻, pp.1007-1012, 1999.
- 8) 九州農政局諫早湾干拓事務所 : "環境モニタリング結果", 資料, pp.(I)1-(I)38, 2002.
- 9) 西田 渉, 野口正人, D.P.Solomatine : "遺伝的アルゴリズムによる水質モデルの検定とその適用性に関する研究", 水工学論文集, 第48巻, pp.1321-1326, 2004.