

# グローバル感度解析に対する カオス力学系やメタモデルの視点

香田 正人

グローバル感度解析について、カオス力学系やメタモデルとの関連を述べる。本稿前半の第1～第3節ではFAST (Fourier Analysis Sensitivity Test) とカオス力学系との関連を論じる。後半の第4～第6節では、メタモデルとの関連からHDMR (High Dimensional Model Representation) を取り上げた後、分散分析 (ANOVA) のHDMRへの適用によるグローバル感度解析 ANOVA-HDMR を解説する。

キーワード：グローバル感度解析, FAST, カオス力学系, メタモデル, ANOVA-HDMR

## 1. FAST におけるサンプル軌道

モデルの入力変数 (パラメータ) を基準値の近傍で局所的に変化させるローカル感度解析と異なり、グローバル感度解析における入力変数はサンプリングの都度、許容されるパラメータ空間内を自由にグローバルに変動できる。すなわち、グローバル感度解析における入力変数は、高次元のパラメータ空間内で一つのサンプル軌道を描くことになる。このとき、サンプル軌道がパラメータ空間全体を一樣に探索するために、軌道のエルゴード性 (ergodicity) が重要となる。サンプル軌道にエルゴード性がなければ、パラメータのサンプリングが部分的で偏ったものとなり入力変数の一樣性が損なわれる。したがって、分散分析 (ANOVA: Analysis of Variance) に基づくグローバル感度解析の有効性を確保するために、サンプル軌道はエルゴード性を備えると同時に、パラメータ空間全体から一樣なサンプリングを行うことが要求される。

FAST (Fourier Amplitude Sensitivity Test) [1] [2] では、高次元のパラメータ空間を正弦関数により1次元空間に変換して得られる Fourier 級数展開に基づき、モデルの出力変数の総分散を関連する入力変数の座標系の各次元における分散に分解する。その特徴は、各パラメータに割り振られた特性周波数 (characteristic frequency) に対応する Fourier 係数のパワースペクトルを使って、感度指標の評価に必要な個

別分散が効率よく計算できることである。FASTの詳細については本特集号の劉氏による第2論文を参照されたい。以下の第3節までの各節では、FASTにおけるサンプル軌道とパラメータの確率分布について考察する。議論にあたって、特定のモデル構造を前提にするものではないことに注意する。

モデルに対する  $k$  次元の入力変数を  $x_i (i=1, 2, \dots, k)$  であらわし、割り当てた特性周波数を  $\omega_i (i=1, 2, \dots, k)$  とする。FASTでは正弦関数を用いて入力変数を次のように定義する [1]。

$$x_i = g_i(\sin(\omega_i s)), i=1, 2, \dots, k \quad (1)$$

ここに  $s$  はスカラーで、 $g_i$  は適当な変換関数である。また、どの特性周波数も他の特性周波数の整数係数線型結合では表されることがない (共約不可能: incommensurate) とする。つまり理想的にはあるが、出力変数の Fourier 級数展開における高周波数成分 (高次ハーモニクス) は、特性周波数に関して干渉がなく独立であるとする。特性周波数の決定法については、文献 [2] を参照されたい。

1次元変数  $s$  を変化させることにより、(1)式からパラメータのサンプル軌道が得られる。変換関数  $g_i$  はサンプル軌道が  $k$  次元のパラメータ空間を余す所なく一樣に探索するよう決定する必要がある。各パラメータが統計的に独立で確率分布が所与の場合に、サンプル軌道のエルゴード性を保証する条件が、パラメータの座標系の各次元における  $g_i$  に関して次の微分方程式により与えられる [2] [3]。

$$\pi\sqrt{1-u^2} p_i(g_i(u)) \frac{d}{du} g_i(u) = 1 \quad (2)$$

ここに  $g_i(0)=0, u=\sin(\omega_i s)$  である。また、(2)式中の

こうだ まさと  
筑波大学 大学院システム情報工学研究科  
〒305-8573 つくば市天王台 1-1-1

$p_i(g_i(u))=p_i(x_i)$  はパラメータの確率密度関数を表す。

したがって、各パラメータが統計的に独立で確率分布が所与の場合、FASTにおける入力変数の変換関数  $g_i$ 、すなわちサンプル軌道は(2)式を満足するように選ぶことが重要となる。逆に、サンプル軌道を(1)式で与えたときに、(2)式から計算されるパラメータの各座標系における確率分布関数が一様分布 ( $p_i(x_i)=1$ ) となる必要がある。

## 2. FASTにおける代表的な変換関数

FASTにおける最初の変換関数としては次式が提案された[1][2]。

$$x_i = g_i(\sin(\omega_i s)) = \bar{x}_i \exp(\bar{v}_i \sin(\omega_i s)) \quad (3)$$

ここに  $\bar{x}_i, \bar{v}_i$  はそれぞれ、入力変数の代表値とパラメータ変動の限界値をあらわす。この分布は、パラメータが指数関数で変換されるために非対称でテールを有し、明らかに一様分布ではない。

続いて、Kodaら[4]は以下の変換関数を提案した。

$$x_i = g_i(\sin(\omega_i s)) = \bar{x}_i(1 + \bar{v}_i \sin(\omega_i s)) \quad (4)$$

これは(3)式を Taylor 展開したときの線型項までの近似式であり、その分布は対称な Chebyshev 関数 (図1に示す) となるが、やはり一様分布ではない。

簡単のため、これ以降の議論においてはパラメータ (入力変数) を正規化して  $0 < x_i < 1 (i=1, 2, \dots, k)$  とし、パラメータ空間として  $k$  次元の単位超立方体  $\Omega^k = [0, 1]^k$  を考える。FASTの変換関数(4)式の分布が Chebyshev 関数となることは、(4)式で  $\bar{x}_i=1/2, \bar{v}_i=1$  とおいて(2)式に代入し、 $p_i(x_i)=1/\sqrt{x_i(1-x_i)}$  を得ることにより直接確かめられる。

最終的に、パラメータの一様分布を与える変換関数

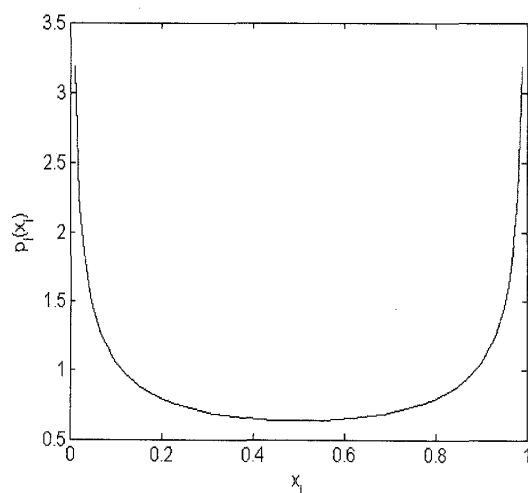


図1 Chebyshev 関数

は Saltelliら[5]により次式で与えられた。

$$x_i = \frac{1}{2} + \frac{1}{\pi} \arcsin(\sin(\omega_i s)) \quad (5)$$

これは拡張 FAST (Extended FAST) と呼ばれる。拡張 FAST の変換関数(5)式で得られるサンプル軌道は図2に示すように、与えられた特性周波数で振動するテント型の区分的線型関数 (三角波) となる。パラメータの一様分布の条件  $p_i(x_i)=1$  を(2)式に代入して得られる微分方程式

$$\pi \sqrt{1-u_i^2} \frac{d}{du_i} g_i(u_i) = 1 \quad (6)$$

を  $g_i$  について解けば、拡張 FAST の変換関数(5)式が得られる[6]。

## 3. FASTとカオス力学系

上で述べた議論を前提に、以下では FAST とカオス力学系との関連を論じる。ここでカオス力学系を取り上げる理由は、確定系における不規則挙動として得られるカオス系列が、FASTのサンプル軌道に必要とされるエルゴード性を備えているからである。

### 3.1 ロジスティック写像

カオス挙動を示す代表的な非線型離散力学系として、以下のロジスティック・システムを考える。

$$y_{n+1} = f(y_n) = 4y_n(1-y_n) \quad (7)$$

ここに  $y_n (0 < y_n < 1)$  はシステムの状態変数でスカラーとし、添字  $n$  は計算ステップを示す。このとき、

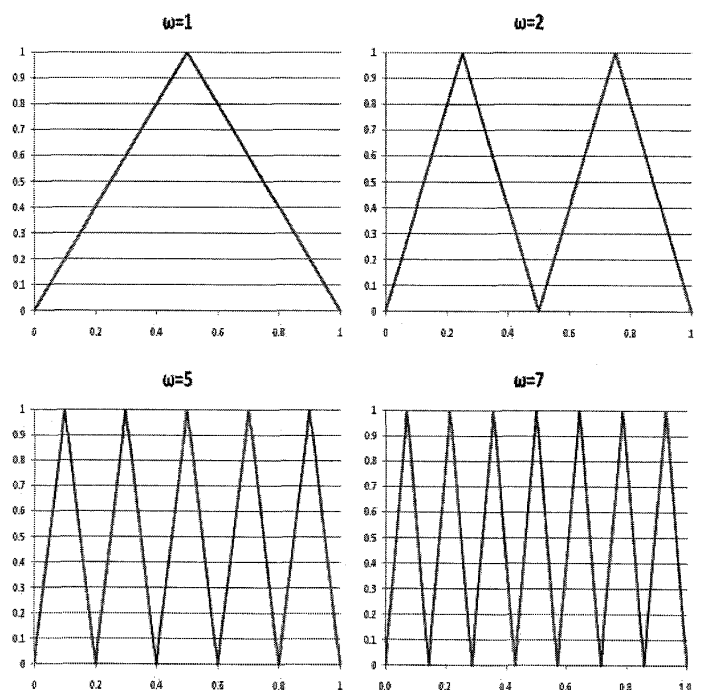


図2 拡張FASTの区分的線型関数

(7)式右辺中のロジスティック写像  $f(y)=4y(1-y)$  は絶対連続で、その不変測度 (invariant measure) は図1で示した Chebyshev 関数  $p(y)=1/\sqrt{y(1-y)}$  となる[7].

不変測度はカオス系列の定常分布として解釈できるものである。ロジスティック写像の不変測度はFASTの変換関数(4)式から得られるパラメータの確率分布と同じく Chebyshev 関数となる。したがって、FASTの変換関数(4)式のサンプル軌道において、1次元変数  $s$  の値を離散的に与えて得られるパラメータ (入力変数) の分布特性は、(7)式から得られるロジスティック写像に基づくカオス系列  $\{y_n\}$  と確率論的に区別できない。

カオス理論の詳細については省略するが、Chebyshev 関数を用いて計算した Lyapunov 指数が正となることから、(7)式の系列  $\{y_n\}$  が初期値への鋭敏な依存性を有することが判定できる。したがって、ロジスティック・システム(7)式の系列  $\{y_n\}$  はカオス系列としてのエルゴード性を備えている。しかし、その定常分布は上述のように一様ではなく、Chebyshev 関数で与えられる。FASTの変換関数(4)式のサンプル軌道から得られるパラメータについても同様である。

図3にコントロール変数  $\mu$  (すなわち(7)式の右辺にかかる係数、(7)式では  $\mu=4$ ) を横軸にとり、ロジスティック・システムから得られる状態変数の系列を示す。コントロール変数  $\mu$  の増大とともに状態変数が倍周期の鞍点分岐 (saddle-node bifurcation) を重ね、Ulam ポイント、すなわち  $\mu=4$  では(7)式から得られるカオス系列  $\{y_n\}$  が全空間  $[0, 1]$  を完全に被覆する様子 (エルゴード性) が示されている。図3に示されたカオス系列の不規則挙動やエルゴード性などの特徴を利用して、カオス力学系が乱数の生成や、グローバル

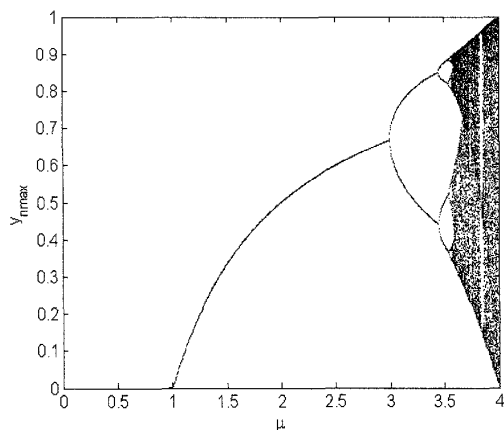


図3 ロジスティック・システムのカオス系列

最適化における解の探索などに広く利用されている。

### 3.2 テント写像

次の微分同相写像  $h(y)$

$$x=h(y)=\frac{2}{\pi}\arcsin(\sqrt{y}) \quad (8)$$

によって、3.1節で導入されたロジスティック写像  $f$  は、以下の区分的線型関数に同型変換 (conjugate transformation) される。

$$\tau(x)=h \circ f \circ h^{-1}(x)=\begin{cases} 2x & (0 < x \leq 1/2) \\ 2(1-x) & (1/2 < x < 1) \end{cases} \quad (9)$$

これはテント写像 (tent map) として知られるが、その不変測度は  $p(x)=1$  となり、拡張FASTの変換関数(5)式が与える確率分布関数と同じく一様分布となる。

テント写像(9)式は、拡張FASTの変換関数(5)式で表されるような、与えられた特性周波数で振動する区分的線型関数 (図2を参照) に容易に拡張できる。図2において、拡張FASTの特性周波数を  $\omega=1$  としたものがテント写像  $\tau(x)$  に他ならない。すなわち、拡張FASTの変換関数(5)式において、スカラー変数  $s$  の値を離散的にサンプリングして得られるパラメータの分布特性は、テント写像  $\tau(x)$  から得られるカオス系列と確率論的に区別できない。したがって、拡張FASTのサンプル軌道はエルゴード性を備えていると同時に、そのパラメータが一様分布するという望ましい特性を有する。そのため、拡張FASTはグローバル感度解析における最も有力な手法の一つとなっている。

本節で論じたFASTとカオス力学系との関連は、筆者[8]により初めて示されたものである。

## 4. メタモデルとHDMR

グローバル感度解析は“モデル独立 (model independent もしくは model-free)” とされる。すなわち、個別のモデル形式や構造、モデル表現法 (例えば、線型システムか否か、加法性や独立性)、パラメータ間の相互干渉の有無などによらず一般的に適用可能である。以下の各節ではモデル独立なメタモデルの視点から、グローバル感度解析について考察する。

グローバル感度解析は、現象に対して演繹的に導かれた状態方程式や数理モデルに基づくのではなく、対象への入出力関係などのシステム機能に着目し、現象が内包する不確かさを確率分布の形で組み込んだ、メタモデルからの帰納的アプローチと考えることができ

る。メタモデルの構築は、大規模複雑系において典型的なように、事前には十分な情報や知識が存在しない状況からスタートして、いわばブラックボックスについての入出力関係に基づき行われることも多い。

こうしたモデル独立な方法論としてのメタモデルに関しては、カーネル回帰など数多くの回帰分析手法や、Spline 近似などの平滑化 (smoothing) 法、さらには GMDH (Group Method of Data Handling) や最近では RBF (Radial Basis Function) ネットワークなどがよく知られている。ここでは、グローバル感度解析で開発されたメタモデルのアプローチとして HDMR (High Dimensional Model Representation) [9]~[11]を取り上げる。

システムについての入出力関係を抽象化して、一般化された入力変数をモデル因子と呼び、出力変数に対応するモデル関数がスカラー関数の場合を考える。簡単のため、正規化により  $k$  次元のモデル因子を  $0 < x_i < 1 (i=1, 2, \dots, k)$  とし、モデル因子から構成される  $k$  次元の単位超立方体  $\Omega^k = [0, 1]^k$  で一様分布をすると仮定する。また、モデル関数を  $f = f(x_1, x_2, \dots, x_k) \in L_2$  とする。

HDMR においてはモデル因子の次元、すなわち座標系がモデル構築の基本となる。HDMR はモデル因子の次元の増加に従って、階層的な座標系で作られる部分空間へのモデル関数の展開を次のように与える。

$$f = f_0 + \sum_{i=1}^k f_i(x_i) + \sum_{1 \leq i < j \leq k} f_{ij}(x_i, x_j) + \dots + f_{12\dots k}(x_1, x_2, \dots, x_k) \quad (10)$$

ここに、積分の評価を  $\Omega^k$  上で行い  $f_0 = \iint \dots \int f(x_1, x_2, \dots, x_k) dx_1 dx_2 \dots dx_k$  は平均値である。上式右辺中の  $f_{ijs}$  などは、添字で示されたモデル因子を独立変数とする 2 乗可積分関数で  $2^k$  個の項数から成り、モデル因子の次元の増加 (最大  $k$  次元) に従い、定数項  $f_0$  からスタートして順次再帰的に計算できる。ただし、HDMR はモデル因子に関する有限次元の展開であるため、一般の完備な展開、例えば Taylor 展開や Fourier 展開などの無限級数とは異なり、ユニークに決定できるとは限らない。

モデル因子から構成される空間  $\Omega^k$  について(10)式を積分し、すべての単独モデル因子に関して  $\int f(x_i) dx_i = 0$  であれば、 $\iint f(x_i) f(x_j) dx_i dx_j = 0 (i \neq j)$  となることが導かれる。すなわち、異なる 2 つのモデル因子に

ついでモデル関数の内積が直交する。詳細については省略するが、この直交性を仮定したとき、(10)式右辺中の各項は下のように一意に計算可能である [10]。

$$f_0 = E(f) = \iint \dots \int f(x_1, x_2, \dots, x_k) dx_1 dx_2 \dots dx_k$$

$$f_i(x_i) = E(f|x_i) - f_0$$

$$f_{ij}(x_i, x_j) = E(f|x_i, x_j) - f_i - f_j - f_0, \text{ etc.}$$

ここに、 $E(f|\bullet)$  はモデル因子の情報に基づくモデル関数の条件付き期待値を表す。このとき、条件付き期待値をモデル因子についての最小分散 (回帰) 推定値と考えれば、上式はモデル因子の階層的な座標系が作る部分空間上にモデル関数を射影していく演算として解釈できる。詳しくは、本稿の第 6 節で述べる。

最後に“オッカムの剃刀 (Occam's razor)”，つまり解き明かしたい機能や現象を説明するものの中では、そのレベルで最もシンプルなモデルが最も優れている、とされることに注意されたい。HDMR モデルにおける(10)式右辺の全項数はモデル因子のべき乗  $2^k$  となり、いわゆる“次元の呪い (curse of dimensionality)”をもたらす可能性がある。すなわち、モデルは詳細で複雑なものが一概に良いわけではなく、メタモデルを単純化する最適化手法の開発が今後の課題となる。

## 5. ANOVA-HDMR

分散分析 (ANOVA) を(10)式に適用して得られる展開は ANOVA-HDMR [10] と呼ばれる。Sobol' 法 [11] は ANOVA-HDMR の代表例であるが、これは超一様分布準乱数列 (minimum discrepancy sequence) で有名な Sobol' により提案された手法である。Sobol' 法については、本特集号の本間氏による第 3 論文を参照されたい。また、本稿前半の第 1~第 3 節で取り上げた FAST も ANOVA-HDMR に分類される。

モデル関数  $f$  の総分散は次式で定義される。

$$D = \iint \dots \int f(x_1, x_2, \dots, x_k)^2 dx_1 dx_2 \dots dx_k - f_0^2 \quad (11)$$

ここに、積分はモデル因子が定義する単位超立方体  $\Omega^k$  上で評価するものとする。さらに HDMR の(10)式を 2 乗して  $\Omega^k$  で積分すれば、(11)式を考慮することによりモデルの総分散に対する次の ANOVA-HDMR を得る。

$$D = \sum_{i=1}^k D_i + \sum_{1 \leq i < j \leq k} D_{ij} + \dots + D_{12\dots k} \quad (12)$$

上式右辺中の  $D_{ijs}$  などは、添字で示されたモデル因子の座標系におけるモデル関数の偏分散 (partial

variance) を表し  $(2^k - 1)$  個の項数となるが、モデル因子の次元の増加に従って順次再帰的に計算できる。

また、感度指標を  $S_{ijs} = D_{ijs}/D$  などと相対的な分散比で定義し、(12)式の両辺を総分散で除せば次式を得る。

$$1 = \sum_{i=1}^k S_i + \sum_{1 \leq i < j \leq K} S_{ij} + \dots + S_{12 \dots k} \quad (13)$$

上式は、モデル因子の各次元における感度指標の総和が1となるので便利である。特に、上式の右辺第1項は個別モデル因子についての1次オーダーの主感度を表す。また、(13)式の右辺中の第2項以降の各項は、モデル関数の偏分散に対する感度指標を、モデル因子の座標系の次元の増加に従って階層的に分解していったものである。ANOVA-HDMRの(12)や(13)式の評価に必要な、モデル因子の空間  $\Omega^k$  についての多重積分の計算には、モンテカルロ法を使うのが一般的である。

## 6. 単位の分解

HDMR や ANOVA-HDMR に対応する射影演算子  $P$  を考えることができる。明らかに  $P^2 = P$  および  $P(1-P) = 0$  となる。モデル関数の空間に導入された適当な測度に基づいて射影演算子を定義することにより、HDMRの展開をユニークに決定できることが知られている。このときHDMRの(10)式に対応して、射影演算子は次式を満足する[12]。

$$1 = P_0 + \sum_{i=1}^k P_i + \sum_{1 \leq i < j \leq k} P_{ij} + \dots + P_{12 \dots k} \quad (14)$$

上式左辺の1は単位作用素であり、右辺における  $P_{ijs}$  などは、添字で示されたモデル因子の座標系が作る部分空間に対する射影演算子を表す。(14)式は、モデル因子の次元の増加に従って、単位作用素をモデル因子の座標系が作る、より高い次元の部分空間への射影に順次階層化していくことにより、“単位の分解 (resolution of the identity)” を与えるものである。

HDMR は、単位作用素を用いて  $1 \cdot f = f(x_1, x_2, \dots, x_k)$  としたときの、単位の分解についての視点から統一的に理解できる。一般に、(14)式のような単位の分解には複数の作用素展開があり得る。例えば、適当な直交関数系 (一様分布に対する Legendre 多項式など) についてのテンソル積で生成することができる。こうした場合、モデル関数の完備な表現を与えるが、無限次元の展開となる。これに対して(14)式は、モデル因子の次元の増加に従って順次得られる、より大きな次元の部分空間への射影により、単位作用素を階層的に有

限次元で分解するものである。また、(14)式の右辺中の高次の作用素展開は、高次元部分空間におけるモデル因子間相互の相関作用に対応する。部分空間の次元の増加に従いモデル因子の座標間の高次偏相関係数が急速に減衰するため、(14)式の収束が早いことが知られている。

2乗可積分関数空間  $L_2$  における Lebesgue 測度に基づき、第5節で導出した ANOVA-HDMR に対しても同様な射影演算子を定義できる[12]。このとき、射影演算子は着目するモデル因子の座標系が作る部分空間に関する偏分散のみを抽出して、総分散についての情報を階層的に縮約していくものと解釈できる。この意味では、ANOVA-HDMR は適当な射影を行うことにより、モデルの総分散をモデル因子の各次元の偏分散に情報縮約することに他ならない。ただし、射影を行うことはモデルを粗視化 (coarse graining) することでもあり、モデル因子間に無視できない相関がある場合などには注意を要する。

## 参考文献

- [1] R. I. Cukier, C. M. Fortuin, K. E. Schuler, A. G. Petschek and J. H. Schaibly: "Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients I," J. Chemical Physics, vol. 59, pp. 3873-3878, 1973.
- [2] R. I. Cukier, R. B. Levine and K. E. Schuler: "Non-linear sensitivity analysis of multiparameter model systems," J. Computational Physics, vol. 26, pp. 1-42, 1978.
- [3] H. Weyl: American J. Mathematics, vol. 60, pp. 889-896, 1938.
- [4] M. Koda, G. J. McRae and J. H. Seinfeld: "Automatic sensitivity analysis of kinetic mechanisms," Int. J. Chemical Kinetics, vol. 11, pp. 427-444, 1979.
- [5] A. Saltelli, S. Tarantola and K. Chan: "A quantitative, model independent method for global sensitivity analysis of model output," Technometrics, vol. 41, pp. 39-56, 1999.
- [6] Y. Lu and S. Mohanty: "Sensitivity analysis of a complex, proposed geologic waste disposal system using the Fourier Amplitude Sensitivity Test method," Reliability Engineering and System Safety, vol. 72, pp. 275-291, 2001.
- [7] R. May: "Simple mathematical models with very complicated dynamics," Nature, vol. 261, pp. 459-467, 1976.

- [8] M. Koda: "Uncertainty and sensitivity analyses based on chaos search in high-dimensional parameter space," 高次元データの統計学—理論・方法論・関連分野への応用—(科学研究費によるシンポジウム論文集 2009年12月14日-16日 筑波大学), pp. 35-45, 2009.
- [9] H. Rabitz, O. Alis, J. Shorter and K. Shim: "Efficient input-output model representations," Computational Physics Communication, vol.117, pp. 11-20, 1999.
- [10] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana and S. Tarantola: Global Sensitivity Analysis The Primer, John Wiley & Sons, 2008.
- [11] I. M. Sobol': "Sensitivity analysis for non-linear mathematical models," Mathematical Modelling and Computational Experiment, vol. 1, pp. 407-414, 1993.
- [12] H. Rabitz and O. Alis: "Managing the tyranny of parameters in mathematical modelling of physical systems," in A. Saltelli, K. Chan and E. M. Scott (Eds), Sensitivity Analysis, John Wiley & Sons, pp. 199-223, 2000.