# ベイズ統計手法に基づきケミカルマスバランス計算を行う表計算マクロの作成：時系列データ処理

花石竜治

筆者は、ベイズ法によりケミカルマスバランス(CMB)計算を行う Microsoft 社 Excel のマクロを 2021 年度に作成した。本報告では、それを時系列データ解析に用いるための時系列データ作成法とマクロの動作、出力値について述べ、出力値に関係する未知発生源の扱いや寄与の比較について報告する。また本研究では、初期値乱数設定および乱数処理法の改良を行い、ベイズ法の応用に用いるマルコフ連鎖モンテカルロ計算の計算条件について知見が得られた。解析例として、従来法で解析が困難だった底質中のダイオキシン類の発生源解析を取り上げ、ベイズ法が安定した寄与率を与えることを示した。以上の検討により、本研究では、ベイズ法による CMB 計算をより精度よく可能にすることができたと考える。

Key words：Bayesian statistics, Chemical mass balance, Time-series data, Dioxins

## １．はじめに

微小粒子状物質としての PM2.5 の削減のシナリオは、成分測定の結果から発生源を推定し、対策を行うことである。発生源推定には、元素濃度などの化学的な情報が用いられる。これまで、浮遊粉じんの発生源推定には、リセプターモデルの中で、発生源から観測点まで化学種が組成を変えずに到達するという仮定を用いた Chemical Mass Balance(CMB)[1] 法や、発生源の情報なしにその化学組成およびその寄与を算出する Positive Matrix Factorization(PMF)[1] 法が用いられてきた。

CMB 法は発生源情報を必要とするが、この情報が確かな場合には環境汚染の観測濃度があれば、少ない計算機資源で発生源寄与を求められるという長所がある。この方法が提唱 [2] されてから半世紀が経過したが、近年は電子計算機の発達で高度な数値計算が可能となり、古典的・頻度主義的な確率論・統計学に基づいた最小二乗法はもちろん、マルコフ連鎖モンテカルロ (MCMC) 法によるベイズ統計手法も計算手法となっている。

筆者は、2021 年度、柏木らのベイズ統計手法 [3] を用いる CMB 法を簡略化して、演算中心部を C 言語で記述した DLL を用いるマイクロソフト社 Excel の VBA マクロとしてコードし、データ表から直接的に CMB 計算をできる環境を整えた [4]。このマクロは、内海ら [5] により PM2.5 の発生源推定に用いられた。

今回、筆者が作成したマクロのうち、内海らが利用した時系列データへの応用の部分を中心に、その後の改良も含めて報告する。

## ２．昨年度開発した Excel マクロについて

計算の流れを以下に簡単に記述する。

発生源寄与および未知発生源の化学種組成をパラメータとし、これが確率変数となるベイズ型モデルとする。柏木は、パラメータが取るベイズ更新についての完全条件付き確率分布である、切断正規分布の期待値と分散の表式を与えた [3]。

MCMC 法を行うには数種類の方法がある [6] が、ギブズサンプラの方法では、パラメータの初期値を乱数で与え、パラメータについて順番にこのベイズ更新を行う。すべてのパラメータについて更新が終了したら、また最初のパラメータセットからベイズ更新を行う。これを繰返して実行すれば、定常的なマルコフ連鎖となる。多数回の実行結果（マルコフ連鎖を多数回行う単一連鎖法および異なった乱数初期値のセットを用いる多重連鎖法による）の期待値を取り（モンテカルロ積分）、モデルパラメータの推定値とする。

## ３．時系列データの設定と解析の概要

図 1 に示すように、時系列データを、おおもとの CMB 計算表の化学種濃度の順にしたがい、記述する。データ名（日付）の列には分析値(μg/m³)、その右の列には標準偏差(μg/m³)を記入する。行による化学種の記述後、一行開けて、質量濃度を記入する。その下には、CMB 計算では用いないが、寄与計算には用いる化学種の濃度(μg/m³)を記述する。これを、列を連続として、右に追加し、時系列データとする。空白列があれば、マクロはそこを時系列データの終わりと認識する。

このワークシートに TimeSeriesDefine という文字列を含む名前をつけ、CMB ダイアログを立ち上げると、プログラムが自動認識し、TimeSeriesDefine がワークシート名の先頭につくシート名の一覧を表示する。図 2 にそのダイアログを示した。

次にこのダイアログで、CMB 計算そのものの設定（例えばベイズ法であれば、繰返し回数、burn-in 回数、多重連鎖法回数、間引き回数などの設定）をしてから、設定した TimeSeriesDefine*を選択し、「時系列データ解析」をクリックする。マクロは最初に化学種濃度の並びのチェックを行ってから、時系列データを次々に読み込み、CMB 計算を行う。原則として、時系列データ解析では、解析に使用した発生源プロファイルのワークシートの表の下に結果を出力する。

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | | | | | | | |
| 2 | | 2019/5/8 | | 2019/5/9 | | 2019/5/10 | |
| 3 | | | | | | | |
| 4 | Na | 0.457 | 0.0014 | 0.359 | 0.0014 | 0.255 | 0.0014 |
| 5 | Al | 0.0974 | 0.00098 | 0.562 | 0.00098 | 0.229 | 0.00098 |
| 6 | K | 0.0963 | 0.0002 | 0.201 | 0.0002 | 0.139 | 0.0002 |
| 7 | Ca | 0.073 | 0.0013 | 0.257 | 0.0013 | 0.12 | 0.0013 |
| 8 | Sc | 0.00002 | 2.2E-06 | 0.000135 | 2.2E-06 | 0.000049 | 2.2E-06 |
| 9 | Ti | 0.0059 | 0.000067 | 0.0265 | 0.000067 | 0.0113 | 0.000067 |
| 10 | V | 0.00107 | 5.4E-06 | 0.00224 | 5.4E-06 | 0.00188 | 5.4E-06 |
| 11 | Cr | 0.00031 | 0.000048 | 0.00103 | 0.000048 | 0.0005 | 0.000048 |
| 12 | Mn | 0.00308 | 6.6E-06 | 0.0105 | 6.6E-06 | 0.00538 | 6.6E-06 |
| 13 | Fe | 0.077 | 0.0014 | 0.341 | 0.0014 | 0.143 | 0.0014 |
| 14 | Co | 3.03E-05 | 1.4E-07 | 0.000115 | 1.4E-07 | 0.0000549 | 1.4E-07 |
| 15 | Ni | 0.0004 | 0.00013 | 0.0011 | 0.00013 | 0.0008 | 0.00013 |
| 16 | Cu | 0.000742 | 8.1E-06 | 0.00208 | 8.1E-06 | 0.00123 | 8.1E-06 |
| 17 | As | 0.00195 | 2.5E-06 | 0.00266 | 2.5E-06 | 0.00222 | 2.5E-06 |
| 18 | Se | 0.000673 | 2.2E-06 | 0.00127 | 2.2E-06 | 0.00083 | 2.2E-06 |
| 19 | Rb | 0.000321 | 7.1E-07 | 0.000836 | 7.1E-07 | 0.000483 | 7.1E-07 |
| 20 | Mo | 0.000177 | 1.2E-06 | 0.000496 | 1.2E-06 | 0.000227 | 1.2E-06 |
| 21 | Sb | 0.000304 | 1.2E-06 | 0.000613 | 1.2E-06 | 0.000489 | 1.2E-06 |
| 22 | Cs | 3.62E-05 | 2.8E-07 | 9.23E-05 | 2.8E-07 | 0.000048 | 2.8E-07 |
| 23 | Ba | 0.00116 | 4.8E-06 | 0.00378 | 4.8E-06 | 0.00233 | 4.8E-06 |
| 24 | La | 5.64E-05 | 5.2E-07 | 0.000221 | 5.2E-07 | 0.000105 | 5.2E-07 |
| 25 | Ce | 0.000107 | 1.2E-06 | 0.000488 | 1.2E-06 | 0.00021 | 1.2E-06 |
| 26 | Sm | 7.6E-06 | 5.2E-07 | 3.53E-05 | 5.2E-07 | 0.0000164 | 5.2E-07 |
| 27 | W | 0.000044 | 1.8E-06 | 0.000358 | 1.8E-06 | 0.000071 | 1.8E-06 |
| 28 | | | | | | | |
| 29 | mass | 12.9 | | 23.5 | | 15.7 | |
| 30 | | | | | | | |
| 31 | NH4+ | 1.08 | 0.002 | 1.86 | 0.002 | 1.45 | 0.002 |
| 32 | NO3- | 0.59 | 0.013 | 0.64 | 0.013 | 0.39 | 0.013 |
| 33 | SO42- | 3.19 | 0.0022 | 5.04 | 0.0022 | 3.74 | 0.0022 |
| 34 | Cl | 0.078 | 0.0017 | 0.065 | 0.0017 | 0.05 | 0.0017 |

**図 1　時系列データのスプレッドシート表の一部**

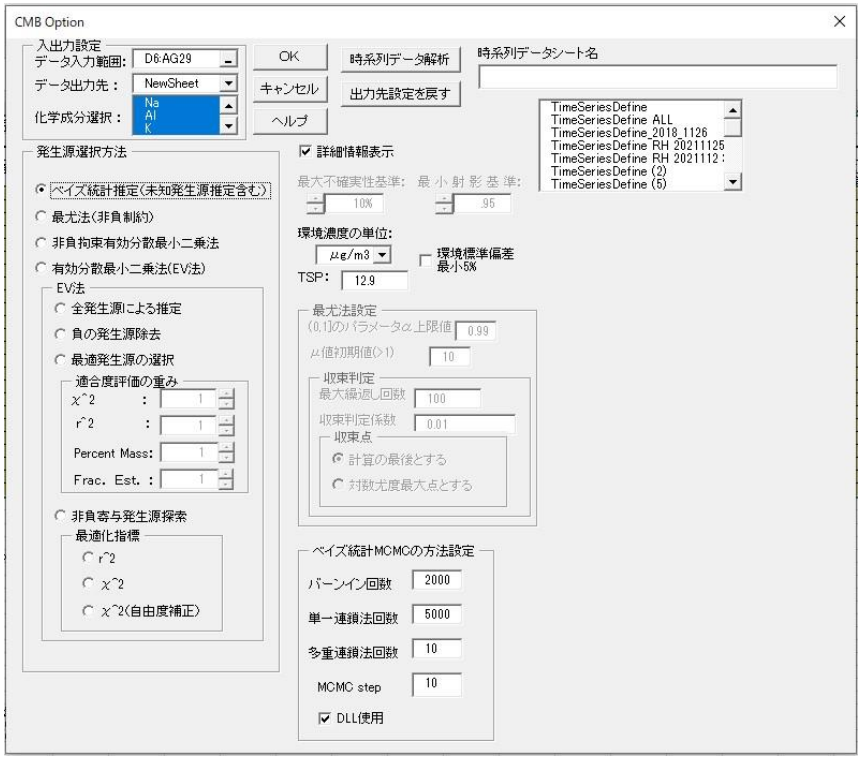



**図 2　CMB 計算マクロの起動時のダイアログ**

## ４．CMB 計算および解析理論の改良

### 4.1　時系列データ解析における出力

時系列データ解析で出力される量は、発生源寄与、CMB 計算では用いないが寄与の計算を行う化学種の寄与、割り当てられなかった発生源寄与($UK$)、不確実さ($UC$) である。以下に説明する。

### 4.2　CMB 計算では用いないが寄与の計算を行う化学種の寄与

CMB 計算で求められた発生源の寄与を $\theta_j$ とすると、柏木による CMB 方程式[3] は、

$$\zeta_i = \sum_{j=1}^{n_j} \xi_{ij}\theta_j + \omega_i\left(1 - \sum_{j=1}^{n_j}\theta_j\right) \qquad (1)$$

ここで、

$\zeta_i$：$i$ 番目の化学種の観測サイトにおける環境濃度（計算時の単位：g 化学種/g 粉じん）。大気中濃度が $z_i$ (μg 化学種/m$^3$)で、質量濃度が $m_{total}$ (μg 粉じん/m$^3$)のとき、

$$\zeta_i = \frac{z_i}{m_{total}} \qquad (2)$$

$\xi_{i,j}$：$i$ 番目の化学種の発生源プロファイルの $j$ 番目の発生源を示す。計算時の単位：g 化学種/g 粉じん

$\theta_j$：$j$ 番目の発生源寄与（無次元）

$\omega_i$：未知発生源における $i$ 番目の化学種濃度。計算時の単位：g 化学種/g 粉じん

$n_j$：発生源の数

である。

次に、CMB 計算に考慮していない化学種を考える。その化学種が $n_k$ 種あり、$k$ 番目の化学種の大気中における濃度が $z'_k$ (μg 化学種/m$^3$)であるとする。これを質量濃度 $m_{total}$ (μg 粉じん/m$^3$) で除すると、その商 $\eta'_k = z'_k/m_{total}$ は、粉じん中の化学種の濃度を表し､単位は g化学種/g粉じんとなる。この化学種が観測地点に届くとすれば、寄与が $\theta'_k$ である仮想的な発生源を考えて、プロファイルが $\xi'_{k,k}$ （単位：g 化学種/g 粉じん）である場合、

$$\eta'_k = \xi'_{k,k}\theta'_k \qquad (3)$$

となる。

以上から、CMB 計算に考慮していない化学種も含めて CMB 方程式をベクトル･行列表記で書くと、

$$\begin{pmatrix} \zeta_1 \\ \vdots \\ \zeta_{n_i-1} \\ \eta'_1 \\ \vdots \\ \eta'_{n_k} \end{pmatrix} = \begin{pmatrix} \xi_{1,1} & \cdots & \xi_{1,n_j} & & & \\ \vdots & \ddots & \vdots & & O & \\ \xi_{n_i-1,1} & \cdots & \xi_{n_i-1,n_j} & & & \\ & & & \xi'_{1,1} & 0 & \cdots & 0 \\ & O & & 0 & \xi'_{2,2} & & \vdots \\ & & & \vdots & & \ddots & 0 \\ & & & 0 & \cdots & 0 & \xi'_{n_k,n_k} \end{pmatrix} \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_{n_j} \\ \theta'_1 \\ \theta'_2 \\ \vdots \\ \theta'_{n_k} \end{pmatrix} \qquad (4)$$

となる。ここで、再掲するものもあるが、

$\zeta_i$ : $i$ 番目の化学種の環境濃度（g 化学種/g 粉じん）

$\xi_{i,j}$ : $j$ 番目の発生源における $i$ 番目の化学種のプロファイル（g 化学種/g 粉じん）

$n_i$：CMB 計算を行う化学種の数+1

$n_j$：発生源の数

$\theta_j$ : $j$ 番目の発生源の寄与（無次元）

$\eta'_k$：CMB 計算には考慮しない $k$ 番目の化学種の環境濃度（g 化学種/g 粉じん）

$\xi'_{k,k}$：考える係数（g 化学種/g 粉じん）

$\theta'_k$：CMB 計算には考慮しない $k$ 番目の化学種の寄与（無次元）

である。

(3)式で、共通の因子として

$$\xi' \equiv \xi'_{k,k} \qquad (5)$$

として、この係数を決定する。

いま、CMB 計算に用いない化学種を含めて、質量均衡が成り立つとすると、

$$\sum_{i=1}^{n_i-1}\zeta_i + \sum_{k=1}^{n_k}\eta'_k = 1\ \text{g 化学種/g 粉じん} \qquad (6)$$

$$\sum_{j=1}^{n_j}\theta_j + \sum_{k=1}^{n_k}\theta'_k = 1 \qquad (7)$$

となる。これらが成り立つためには、$\xi' = 1$ g 化学種/g 粉じん とおけばよい。すなわち、

$$\eta'_k = \theta'_k \times 1\ \text{g 化学種/g 粉じん} \qquad (8)$$

すなわち、

$$\theta'_k = \eta'_k/(1\,\text{g 化学種/g 粉じん}) \tag{9}$$

と取ればよいことが分かる。よって、ベクトル・行列で表した CMB 方程式(4)式は、

$$\begin{pmatrix} \zeta_1 \\ \vdots \\ \zeta_{n_i-1} \\ \eta'_1 \\ \vdots \\ \eta'_{n_k} \end{pmatrix} = \begin{pmatrix} \xi_{1,1} & \cdots & \xi_{1,n_j} & & & \\ \vdots & \ddots & \vdots & & O & \\ \xi_{n_i-1,1} & \cdots & \xi_{n_i-1,n_j} & & & \\ & & & 1 & 0 & 0 & 0 \\ & O & & 0 & 1 & 0 & 0 \\ & & & 0 & 0 & \ddots & 0 \\ & & & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_{n_j} \\ \theta'_1 \\ \theta'_2 \\ \vdots \\ \theta'_{n_k} \end{pmatrix} \tag{10}$$

となる。

以上から、CMB 計算で考慮しない化学種を含めた場合の未知発生源の寄与 $UK$ は

$$UK = 1 - \left\{ \sum_{j=1}^{n_j} \theta_j + \sum_{k=1}^{n_k} \theta'_k \right\} \tag{11}$$

と定義される。

### 4.3　発生源寄与の不確実さ

モンテカルロ積分から、$j$ 番目の発生源寄与の確率分布の期待値の近似値を求めて発生源寄与とするが、この分散を $\sigma_j^2$ として算出する。この分散を用いて、次式により発生源寄与の不確実さ $UC$ を定義する。

$$UC = \sqrt{\frac{\sum_{j=1}^{n_j} \sigma_j^2}{n_j}} \tag{12}$$

### 4.4　比率モデルの発生源寄与から、比較可能な発生源寄与へ

これについては、内海ら[5]に記述されている。いま、(1)式の右辺第一項を $A_i$ として、既知発生源由来の $i$ 番目の化学種の環境濃度（計算時の単位：g 化学種/g 粉じん）を表すとする。すなわち、

$$A_i = \sum_{j=1}^{n_j} \xi_{ij}\theta_j \tag{13}$$

となる。

次に次式を考える。

$$m_{total} \cdot A_i = m_{total} \sum_{j=1}^{n_j} \xi_{ij} \cdot \theta_j = \sum_{j=1}^{n_j} \xi_{ij} \cdot m_{total} \cdot \theta_j \tag{14}$$

$m_{total}$ は、大気中の単位体積あたりの質量濃度（単位：μg 粉じん/m³）で、$A_i$ は、その質量を与える粉じん中の $i$ 番目の化学種の濃度（単位：g 化学種/g 粉じん）であるから、その積 $m_{total} \cdot A_i$ は、大気中の単位体積あたりの化学種の濃度（単位：μg 化学種/m³）となる。

無次元の発生源寄与 $\theta_j$ は、質量濃度との積 $m_{total} \cdot \theta_j$ を与えて、この単位は μg 粉じん/m³ となり、粉じん濃度と寄与の両方の情報を持つ。

(14)式の CMB 方程式は、大気中の化学種濃度（単位：μg 化学種/m³）を、g 化学種/g 粉じんの次元の発生源プロファイルで CMB 解を求めるものと等価であり、この方程式で得られる寄与は、g 粉じん/m³ の次元を持つ。

以上から、$m_{total} \cdot \theta_j$ は、質量濃度が比率モデルの寄与に乗算された発生源寄与と解釈される。この発生源寄与 $m_{total} \cdot \theta_j$ は、異なる検体間で比較が可能である。

### 4.5　乱数初期値

次式により初期乱数を設定するよう改良した。ここで、$\theta_j^0$ は $\theta_j$ について、$\omega_i^0$ は $\omega_i$ についての乱数初期値である。

$$\theta_j^0 = \frac{(r_{rand}\ \%\ 100) + 1}{101(n_j + 1)} \times 0.01 \tag{15}$$

$$\omega_i^0 = \frac{(r_{rand}\ \%\ 100) + 1}{101 n_i} \times 0.01 \tag{16}$$

演算子 % は、整数の除算の剰余を表し、$r_{rand}$ は電子計算機が与える符号なし整数型の乱数を示す。これにより、$N$ を変数の次元として、区間 $[0, 1/100(N+1)] \cong [0, 1 \times 10^{-3}]$ の範囲の乱数初

期値を生成した。

この乱数初期値は、パラメータ $\theta_m$ で区間内 $[0,1-\sum_{j\neq m}\theta_j]$、パラメータ $\omega_m$ で区間内 $[0,1-\sum_{i\neq m}\omega_i]$ にある。

### 4.6　乱数の更新法

区間 $[0,1]$ に入る一様乱数 $U$ を、上下端間の区間 $[a,b]$、期待値 $\mu_t$、分散 $\sigma_t^2$ の切断正規分布にしたがう乱数 $x_t$ を変換する式として、

$$x_t=\phi^{-1}\{\phi(\alpha)+U(\phi(\beta)-\phi(\alpha))\}\sigma_t+\mu_t \quad (17)$$

ここで、

$$\alpha=\frac{a-\mu_t}{\sigma_t} \quad (18)$$

$$\beta=\frac{b-\mu_t}{\sigma_t} \quad (19)$$

とした [4]。なお、$\phi(x)$ は標準正規分布の累積関数である。

この式は、入力の $U$ が区間 $[0,1]$ にある場合には有効であるが、区間が狭い場合に修正が必要である。

その区間を $[a,b]$ とし、特に今回の場合、寄与および未知発生源の化学種濃度の切断正規分布の下端はゼロであるから、$a=0$ として、区間 $[0,b]$ を考える。

次に、係数 $k_t$ を仮定し、

$$\psi=\phi(\alpha)+k_tU(\phi(\beta)-\phi(\alpha)) \quad (20)$$

とする。$U=a=0$ のとき、および $U=b$ のとき、(20)式から、

$$\psi=\phi(\alpha) \quad (21)$$

$$\psi=\phi(\alpha)+k_tb\left(\phi(\beta)-\phi(\alpha)\right) \quad (22)$$

(22)式の右辺が $\phi(\beta)$ に等しい条件を課すと、

$$\phi(\alpha)+k_tb\left(\phi(\beta)-\phi(\alpha)\right)=\phi(\beta) \quad (23)$$

(23)式を解くと、

$$(k_tb-1)\{\phi(\beta)-\phi(\alpha)\}=0 \quad (24)$$

ここで、

$$\phi(\beta)-\phi(\alpha)>0 \quad (25)$$

であるから、

$$k_t=\frac{1}{b} \quad (26)$$

よって(17)式は、

$$x_t=\phi^{-1}\left\{\phi(\alpha)+\frac{U}{b}(\phi(\beta)-\phi(\alpha))\right\}\sigma_t+\mu_t \quad (27)$$

となる。(27)式が、区間 $[0,b]$ に入る切断正規分布に従う乱数を入力とした乱数変換の式である。

## ５．時系列データ解析における繰返し回数の検討

2019 年 5 月 8 日から 21 日において弘前市立文京小学校でサンプリングした PM2.5 測定結果（時系列データ）の CMB 計算を、内海ら [5] が用いた発生源プロファイルおよび環境データのもとで多重連鎖法を 10 回行い、単一連鎖法における繰返し回数を 250 回、500 回、1000 回、5000 回および 10000 回として、その回数まで 200 回のデータを、間引きステップ 10 回でサンプリングし、得られた 200 個のデータから寄与の期待値と分散 $\sigma_j^2$ を求め、それから(12)式により不確実さ $UC$ を計算した。

ここで、繰返し回数は、パラメータセット内におけるギブズサンプラ更新の回数を含まず、このセット内の更新を一巡したときに、繰返し回数が 1 増えるという意味である。

図 3 に不確実さ $UC$ のプロットを示した。ここで、iter とは繰返し回数を指す。

図 3 から、不確実さ $UC$ が低くなり下げ止まる繰返し回数は、当然、検体にもよるが、概ね 5000 回であることが分かった。

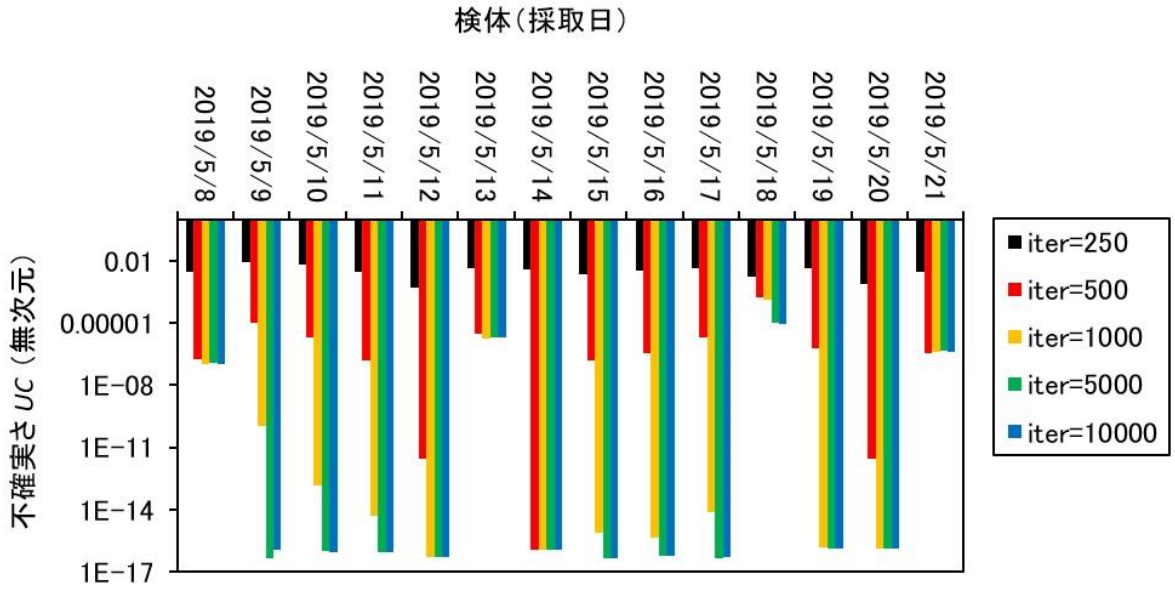



**図 3　弘前市立文京小学校でサンプリングした PM2.5 解析における MCMC 計算の繰返し回数と得られた寄与の不確実さとの関係**

### ６．CMB 計算についての話題としてのダイオキシン類データ解析

ここでは、従来法（有効分散最小二乗法、非負拘束有効分散最小二乗法、非線形計画化法）では発生源寄与の計算が困難であったが、ベイズ法の適用により発生源が推定された一例として、底質のダイオキシン類の例を挙げる。

この事例は、花石ら[7]が非負拘束有効分散最小二乗法の応用例として報告した。ここでは、発生源と仮定するのが過去の農薬で、環境データは 2008 年の東京湾の st.5 における底質中のダイオキシン類[8]である。農薬中のダイオキシン類のうち、ポリ塩素化ジベンゾジオキシン(PCDD)およびポリ塩素化ジベンゾフラン(PCDF)の 17 種類について、農薬中の含有量[9]を発生源プロファイルとして用いた。また、環境濃度における標準偏差は測定濃度の 5 %と仮定した。

表 1 に発生源プロファイルおよび環境データを、表 2 に各 CMB 法での寄与率%を示す。ここでは、有効分散最小二乗法(EVLSQ-ALL)、負の寄与の発生源を除去する有効分散最小二乗法(EVLSQ-NE)、非負拘束有効分散最小二乗法(NN-EVLSQ)、非線形最適化法(MLM)、およびベイズ法(Bayesian)を用いた。

ベイズ法では、繰返し回数 5000 回、burn-in 回数 2000 回、多重連鎖法回数 10 回、間引きステップ数 10 回の条件で計算した。

花石ほか[7]では、CMB 計算に比率モデルを用いず、寄与の相対値を計算して議論した。本研究では比率モデルを、ベイズ法はもちろん、ほかの手法でも用いたが、寄与の上限に制限を設けて計算するベイズ法以外では、100 %を超える寄与率が与えられており、これらの手法による結果は解釈が困難であった。

一方、ベイズ法では、寄与率は 100 %を超えずに、解釈可能であった。この結果、未知発生源寄与も 11.6 %と計算された。ベイズ法により、発生源 H (CNP1987) が 70.0 %、発生源 G (CNP1986) が 12.6 %の寄与率を与え、これらが底質中のダイオキシン類の主な発生源と推定された。ここで、CNP とは、クロルニトロフェンである。

ベイズ法によるダイオキシン類の濃度実測値と CMB 推定値を図 4a に、発生源寄与率の円グラフを図 4b に示した。再現は良好で、また、想定した発生源で環境濃度の約 9 割が占められた。これらの解析結果はベイズ法以外では得られなかったものである。

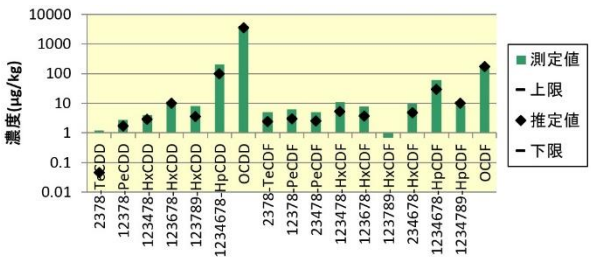
以上から、ベイズ法が、ほかの手法では困難な場合でも、堅牢に CMB 結果を与える方法であり、有用なことが示唆されたと考える。

表 1　東京湾底質のダイオキシン類起源解析の発生源プロファイルおよび環境データ

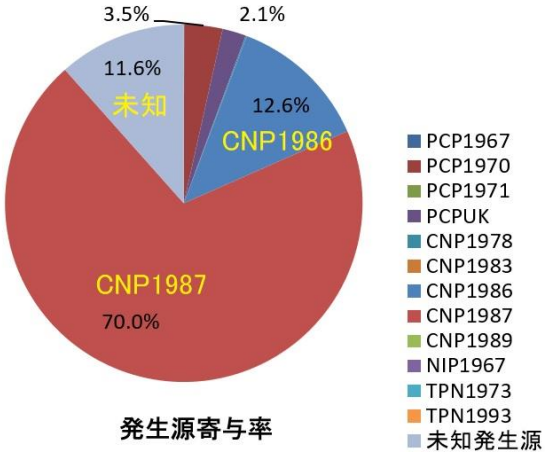
| 異性体 | 発生源濃度(mg/kg) | | | | | | | | | | | | 環境濃度(ng/kg) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | PCP1967 | PCP1970 | PCP1971 | PCPUK | CNP1978 | CNP1983 | CNP1986 | CNP1987 | CNP1989 | NIP1967 | TPN1973 | TPN1993 | |
| | A | B | C | D | E | F | G | H | I | J | K | L | |
| 2378-TeCDD | 0 | 0 | 0 | 0.0019 | 0.0051 | 0 | 0 | 0 | 0 | 0.000083 | 0 | 0 | 1.2 |
| 12378-PeCDD | 0.0058 | 0 | 0 | 0.011 | 0.73 | 0.34 | 0.0019 | 0.00062 | 0.00027 | 0.000064 | 0.0000023 | 0.0000033 | 2.7 |
| 123478-HxCDD | 0.0081 | 0.0022 | 0.0077 | 0.13 | 0.02 | 0.028 | 0 | 0.000023 | 0 | 0 | 0 | 0 | 4 |
| 123678-HxCDD | 0.87 | 0.0046 | 0.12 | 0.0095 | 0.48 | 0.3 | 0.065 | 0.00054 | 0.0014 | 0.000027 | 0.0000051 | 0 | 9.9 |
| 123789-HxCDD | 0.038 | 0 | 0.0046 | 0.008 | 0.17 | 0.13 | 0.024 | 0.00023 | 0.00052 | 0 | 0 | 0 | 7.9 |
| 1234678-HpCDD | 58 | 0.29 | 3.1 | 0.38 | 0.13 | 0.075 | 0.0072 | 0.00019 | 0.00025 | 0.00004 | 0.00003 | 0.000034 | 200 |
| OCDD | 5800 | 34 | 12 | 1.8 | 0 | 0.0016 | 0.00061 | 0 | 0.000083 | 0.00014 | 0.00021 | 0.0002 | 3700 |
| 2378-TeCDF | 0.015 | 0 | 0 | 0.0012 | 0.0048 | 0 | 0.0001 | 0.000015 | 0.000023 | 0.00033 | 0.00001 | 0.0000042 | 4.9 |
| 12378-PeCDF | 0.0039 | 0.0038 | 0.0079 | 0.005 | 0.012 | 0 | 0 | 0.000038 | 0.000078 | 0.000054 | 0.000011 | 0.0000053 | 6.2 |
| 23478-PeCDF | 0.0014 | 0.0018 | 0 | 0.0028 | 0.043 | 0 | 0 | 0.00004 | 0.000029 | 0.000012 | 0.000006 | 0.0000054 | 5.1 |
| 123478-HxCDF | 0.2 | 0.11 | 0.091 | 0.011 | 0.002 | 0 | 0 | 0 | 0 | 0 | 0.00001 | 0.000006 | 11 |
| 123678-HxCDF | 0.031 | 0.018 | 0.033 | 0.004 | 0.0098 | 0 | 0.0055 | 0.000027 | 0.00016 | 0 | 0.0000071 | 0.000006 | 7.7 |
| 123789-HxCDF | 0 | 0 | 0.0026 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.7 |
| 234678-HxCDF | 0.014 | 0.0085 | 0.012 | 0.0061 | 0.82 | 0.25 | 0.0028 | 0.00027 | 0.00028 | 0 | 0.0000055 | 0.0000066 | 9.7 |
| 1234678-HpCDF | 4.6 | 0.64 | 0.96 | 0.052 | 0.018 | 0.0098 | 0.00042 | 0 | 0.000024 | 0 | 0.00014 | 0.000038 | 61 |
| 1234789-HpCDF | 0.6 | 0.28 | 0.26 | 0.0053 | 0 | 0.0002 | 0.000024 | 0 | 0 | 0 | 0.0000075 | 0 | 8.5 |
| OCDF | 47 | 4.4 | 3 | 0.081 | 0 | 0.00063 | 0.00011 | 0 | 0 | 0 | 0.024 | 0.0027 | 170 |

表 2　表 1 の CMB 計算結果。寄与率(%)で表示。ベイズ法による解析結果を太字で強調した

| Method | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| EVLSQ-ALL | 0.055 | -1.67 | 5.57 | 3.00 | 0.607 | -0.932 | 2.27 | -598 | 834 | 1080 | -11000 | 94700 |
| EVLSQ-NE | 0.048 | 0 | 5.10 | 3.00 | 0.038 | 0 | 0 | 0 | 755 | 1400 | 0 | 1490 |
| NN-EVLSQ | 0.048 | 0 | 5.10 | 3.00 | 0.038 | 0 | 0 | 0 | 757 | 1400 | 0 | 1490 |
| MLM | 0.055 | 0.121 | 4.17 | 3.02 | 0.119 | 0.045 | 1.05 | 75.7 | 637 | 1400 | 72.3 | 4050 |
| **Bayesian** | **0.040** | **3.46** | **0** | **2.13** | **0.113** | **0** | **12.6** | **70.0** | **0** | **0** | **0** | **0** |




（a）濃度実測値と CMB 推定値




（b）寄与率の表示

**図 4　東京湾底質のダイオキシン類のベイズ法による CMB 計算結果。(a)　濃度実測値と CMB 推定値、(b)　寄与率の表示**

## 7．結論

時系列データを解析するために、CMB マクロをコードした。この過程で以下の成果が得られた。

1) 解析結果の評価に必要な諸量を定義し、これを出力するようにした。併せて、モンテカルロ法で使用する乱数初期値の設定およびそのベイズ更新の式を改良した。

2) CMB 計算結果の発生源寄与の不確実さは、本報告で使用した例（PM2.5 成分測定データ）では、MCMC 計算の繰返し回数を 5000 回以上にすることで減少した。

3) これらを受けて、従来法で発生源寄与の計算が困難であった底質中のダイオキシン類の発生源解析をベイズ法により行った結果、このダイオキシン類が 1980 年代後半の CNP 不純物由来と推定された。

以上から、ベイズ法による CMB 法が環境汚染解析に有用であることが確認された。今後もデータ解析事例を積み重ね、正確で堅牢な CMB 解析システムとして、完成度を上げていきたい。

## 文　献

1) 橋本俊次ほか：ケミカルマスバランス法による環境汚染物質に対する発生源寄与率の推定. ぶんせき,144-151,2014

2) Friedlander, S. K.: Chemical element balances and identification of air pollution sources. *Environmental Science and Technology*,**7**,235-240,1973

3) Kashiwagi, N.: Chemical mass balance when an unknown source exists. *Environmetrics*,**15**,777-796,2004

4) 花石竜治：ベイズ統計手法によりケミカルマスバランス法計算を行う表計算ソフトマクロの作成. 青森県環境保健センター年報,**32**,69-83,2021

5) 内海宣俊ほか:ベイズ統計手法を用いた CMB 法による微小粒子状物質($PM_{2.5}$)成分の発生源解析. 青森県環境保健センター年報,**32**,123-144,2021

6) 中妻照雄：入門　ベイズ統計学. 朝倉書店,東京,2007

7) 花石竜治ほか：非負寄与を与える発生源探索によるケミカルマスバランス法. 青森県環境保健センター年報,**27**,70-73,2016

8) http://www.kankyo.metro.tokyo.jp/chemical/chemical/dioxin/result/index.html
（2022 年 12 月 26 日現在アクセス可能）

9) Masunaga, S. et al: Dioxin and dioxin-like PCB impurities in some Japanese agrochemical formulations. *Chemosphere*,**44**,873-885,2001