

## 環境データ解析用表計算マクロの紹介と解析例 (II)

—CMB 法解析マクロ—

早 狩 進・花 石 竜 治

Introduction of Spreadsheet Application Macros for Environmental  
Data Analysis and Analytical Examples (II)

—Macro for CMB Analysis—

Susumu HAYAKARI\* and Ryuji HANAISHI\*\*

インターネットの個人用ホームページ<sup>1)</sup>にフリーウェアとして公開している Excel のマクロを解析例も含めて紹介する。CMB 法解析マクロ (CMB8J) は、米国製のフリーウェアを ExcelVBA に移植したもので、発生源を絞り込む過程を示す詳細情報と、最適発生源選択のための指標を新たに提案するなどの特徴を有する。

### 1. はじめに

前回紹介した AndersenAnalyzer に引き続き、CMB8J を紹介する。

CMB (Chemical Mass Balance) 法は、主に粉じん中の重金属汚染の発生源究明に威力を発揮する手法として広く利用されている<sup>1)</sup>。目的とする環境サンプルの重金属組成と、その環境に影響したと考えられる複数の発生源サンプルの重金属組成を元に、環境に寄与した発生源を特定し、その寄与率を算定する手法である<sup>2)</sup>。発生源から環境に到達する過程で変質する可能性がなければ、重金属以外にも適用可能であるため、今後、PM-2.5 の解析や有機塩素系化合物の複合汚染解析にも有効な手法として、利用が期待されているものである。

一般に、CMB 法は、互いに無相関な発生源の線形加法性を前提としており、各発生源寄与率合計は必ず 100% を示すが、それは必ずしも発生源の組み合わせが説明変数として十分であることを示しているわけではないことに注意を払う必要がある。また、互いに無相関であれば各寄与率はすべて正の値を持つが、組成

が類似した発生源を用いると、例えば、一方が 30% でもう一方が -30% というように、互いに相殺しあうような値を持つことが知られている。このような場合を多重共線性が強いといい、一方を除去して解析し直す必要がある。

CMB 法の計算方法としては、指標元素法、線形計画法、最小自乗法、有効分散最小自乗法などが知られている<sup>3)</sup>。そのうち、最小自乗法は (1) 式で示される  $\chi^2$  を最小とする発生源寄与を求める<sup>4)</sup> のに対し、有効分散最小自乗法は (2) 式に示される  $\chi^2$  を最小とするものであり、各測定値の誤差を考慮に入れるという特徴がある<sup>5)</sup>。

$$\chi^2 = \sum_{i=1}^n \frac{\left(C_i - \sum_{j=1}^p a_{ij} S_j\right)^2}{C_i^2} \quad (1)$$

$$\chi^2 = \sum_{i=1}^n \frac{\left(C_i - \sum_{j=1}^p a_{ij} S_j\right)^2}{\sigma_{C_i}^2 + \sum_{j=1}^p \sigma_{a_{ij}}^2 S_j^2} \quad (2)$$

ここで、 $C_i$  は環境試料における化学種  $i$  の濃度、 $a_{ij}$  は発生源  $j$  の試料における化学種  $i$  の濃度、 $S_j$  は発

\* 青森県環境保健センターむつ環境管理事務所 〒035-0073 青森県むつ市中央 1-1-8

\*\* 青森県環境保健センター 〒030-8566 青森県青森市東道 1-1-1

生源  $j$  の寄与である。  $n$  は化学種の数であり、  $p$  は発生源の数である。  $\sigma_{c_i}$ 、  $\sigma_{a_{ij}}$  はそれぞれ  $C_i$ 、  $a_{ij}$  の測定に伴う誤差の標準偏差である。(2) の最小解は、次の繰り返し計算によって求めている。

$$S = ({}^t\text{AVA})^{-1}{}^t\text{AVC} \quad (3)$$

ここで  $A$  は要素が  $a_{ij}$  の行列、  $V$  は対角要素が (2) 式の分母の逆数にあたる対角行列、  $C$  は要素が  $C_i$  のベクトル、  $S$  は要素が  $S_j$  のベクトルである。  ${}^tA$  の  $t$  は行列  $A$  の転置行列の意味である。

有効分散最小自乗法による CMB 法解析ソフトは、既にアメリカの J. G. Watson ら (Desert Research Institute) によってインターネット上にフリーウェア「CMB8」<sup>(6)</sup> としてソース・リストも公開されているが、言語は Pascal、C および FORTRAN で書かれた単独で機能するプログラムであるため、独自のデータファイルを必要とするなど使い勝手があまり良くなかった。そこで、日本語バージョンの作成に当たり、表計算アプリケーション上で解析できるよう、原作者らの了解を得て、Excel VBA マクロへの移植を行い「CMB8J」と名付けたものである。

## 2. CMB8J の特徴

CMB8J には以下の特徴がある。

- (1) CMB8 を Excel VBA マクロでほぼ再現。
- (2) マニュアルレスをめざしたヘルプ機能。
- (3) Excel のデータシートから直接解析可能。
- (4) Excel のグラフ作成機能でグラフ作成。
- (5) Excel の日本語版、英語版に自動対応。
- (6) 環境、発生源の標準偏差が得られない場合既定値で処理 (CMB8 では解析不能)。
- (7) 環境の各成分濃度は mg/kg でも解析可能 (CMB8 では mg/m<sup>3</sup> のデータが必要)。
- (8) 詳細情報として、負の寄与を示す発生源の除去過程および最適発生源選択におけるベスト 10 等の表示機能 (CMB8 では表示されない)。

- (9) 最適発生源選択に用いる 4 指標の基準化を行い、できるだけ対等な評価ができるよう改良。

などが挙げられる。

それらの機能を含めて使い方を以下に紹介する。

## 3. CMB8J の使用方法

CMB8J は、マクロとしてテストデータを含む Excel ファイル「CMB8J.xls」に含まれている。そのファイルとマニュアルは自己解凍型 EXE ファイルに圧縮して提供されるため、最初にハードディスクの空のフォルダにコピーした後、ダブルクリックで解凍する。Excel ファイルを開くと、メニューの右端に「CMB」というメニューが追加され、マクロが利用可能となる。実際に使用する際は、あらかじめテストデータ表 (Table 1) をシートごとコピーし、解析に供する発生源の数になるよう表の発生源列を挿入又は削除し、同じく解析に供する化学種の数に応じて化学成分行も挿入・削除した後、データを記入する。なお、発生源の列は実測値と測定誤差の標準偏差を記入するため、2列ずつ必要である。また、入力するデータの単位は mg/kg のみならず mg/m<sup>3</sup> などにも設定可能であるが、発生源間の単位はそろえる必要がある。標準偏差の実測値が無い場合には、その列を空欄とすると、既定値として 5% を採用する。表が完成したら黄色で示した範囲を反転表示で指定してから、「CMB」をクリックすると、各種設定条件を決める Fig. 1 に示すダイアログが表示される。その中の各種設定を行ってから「OK」をクリックすると、解析結果を表示し、グラフ作成をするかどうか聞いてくるので、必要なら「OK」を押すとグラフ作成後終了する。

## 4. ダイアログによる設定機能

ダイアログ (Fig. 1) の主なボタンの意味するところを以下に示すが、ダイアログを表示直後にヘルプボタンを押すと CMB8J の概略説明を、各種ボタンを押

Table 1. A sample data set for CBM8J.

	Source(mg/kg)						Ambient(mg/kg)					
	A-1		A-2		A-3		Measured(M)		Calculated(C)		C/M (%)	
Species	Conc.	Std. Err.	Conc.	Std. Err.	Conc.	Std. Err.	Conc.	Std. Err.	Conc.	Std. Err.	Conc.	Std. Err.
Ni	15,100	332	12,100	305	95.0	2.7	19,400	243				
Fe	146,000	2810	85,100	2580	170,000	2290	230,000	2060				
Pb	282	10.7	4,770	98.7	29,100	87.3	246	78.8				
Zn	1,470	0.8	22,300	6.8	312,000	60.4	982	5.6				
Mn	2,780	6.3	2,330	5.7	15,600	50.8	3,400	4.6				
Cu	52.0	2.3	57.0	2.1	2,340	18.6	35.1	16.8				
Contribution(R)												
Ratio of R(%)												

Execute the Macro after activate the range colored with yellow

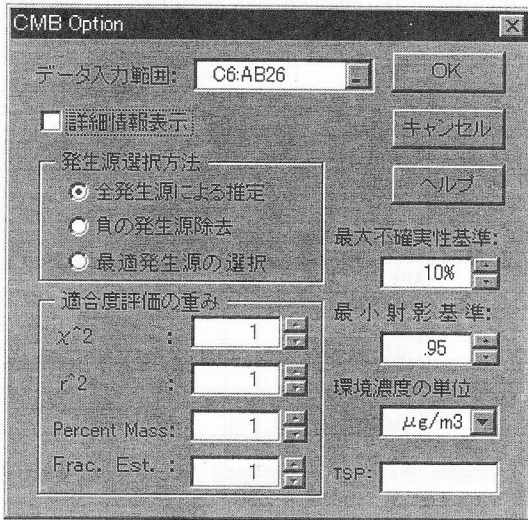


Fig. 1. Dialog display.

してからヘルプを押すとそのボタンの説明を表示する。

#### 4.1 全発生源による推定

すべての発生源を取り込んで環境濃度を再現することを目的に、最小自乗計算を行うものである。ただし、結果がそのまま利用できることはまれで、どの発生源が負の寄与率を持つか把握するために用いることが多い。もしも、負の寄与率を持つ発生源が存在する場合、次の「負の発生源除去」で取り除くことを奨める。

#### 4.2 負の発生源除去

負の寄与率を持つ発生源がなくなるまで、負の値の大きい変数を順次推計計算から除去していくものである。詳細情報にその除去過程を印字してくれる。組成の類似した発生源の組を独立変数としてデータシートに加えた場合に、一方の寄与率が負の大きな値を持つてしまうことが多い。これは、いわゆる多重共線性が災いしているので、多重共線性がなくなるまで発生源を絞り込む必要がある。この処理では、負の寄与率がなくなるまで順次発生源を除去してくれるが、負の値の小さいものは、必ずしも多重共線性に起因するとは限らないので、この詳細情報を元に負の大きな値を示す発生源だけをデータシートから取り除くという一次スクリーニングを行ってから、次の「最適発生源の選択」に持ち込むことを奨める。

#### 4.3 最適発生源の選択

その下の「適合度評価の重み」に示す  $\chi^2$ ,  $r^2$ , Percent Mass, Frac. Est. という4つの指標の値に、そ

れぞれ設定した重みを掛けたものの合計を総合指標 (FitMeasure) として、各発生源の全組み合わせの中から、その総合指標の最大ケースとなる組み合わせを求めるものである。FitMeasure の計算式は、次のとおりである。

$$\text{FitMeasure} = \text{Wt1} \times \chi_0^2 / \text{chisqr} \quad (4) \\ + \text{Wt2} \times (\text{rsqvar} / 0.8)^3 \\ + \text{Wt3} \times (5 / |100 - \text{pcmass}|)^{1/5} \\ + \text{Wt4} \times (\text{FracEst} / 0.8)^3$$

ここで、 $\chi_0^2$  は危険率5%の  $\chi^2$  分布値、chisqr は  $\chi^2$  計算値、rsqvar は  $r^2$  計算値、pcmass は Percent Mass, FracEst は Fraction Estimable, Wt1~4 は Fig. 1 の Dialog で設定されるウェイトを意味する。それら4指標については4.4~4.7で詳述する。

処理速度の限界から、発生源の個数を最大14個と限定しているため、なおさら「負の発生源除去」を元にした一次スクリーニングは欠かせないものである。

#### 4.4 $\chi^2$

$\chi^2$  (chisqr) は、目的とする環境データの化学種毎の実測値と推定値の差の重み付けした二乗和で、それ自体は値が小さいほど適合度が良好と判断されるが、(4)式右辺第1項に示すように、他の指標と合わせるために危険率5%の  $\chi^2$  分布値を  $\chi^2$  計算値で割ったものを指標として用いている。したがって、その項が1.0以上なら良好と評価される。

原典のCMB8では、計算上の分子に当たる  $\chi^2$  分布値の代わりに近似的に自由度を用いている。

#### 4.5 $r^2$

$r^2$  (rsqvar) は、重回帰式における決定係数に当たり、実測環境濃度の分散の内、推定計算でどれだけ説明できるかを示している。値はゼロから1.0までの値も持ち、ここでは0.8を超えれば良く説明できていると評価するように、(4)式右辺第2項に示すとおり0.8で除して更に若干の加工を行い、その項が1.0以上なら良好と判断できるようにしている。

#### 4.6 Percent Mass

$\chi^2$  は化学種毎の適合度をまとめたものとなっているのに対し、Percent Mass (pcmass) は、化学種合計濃度の適合度を評価するもので、実測環境濃度の和を計算された環境濃度の和で割った百分率で現し、 $100 \pm 5\%$ 以内を良好と判断するが、他の指標とそろえるために、(4)式右辺第3項に示す加工を行い、その項が1.0以上なら良好と見なせるようにした。

原典のCMB8では、分母に計算された環境濃度ではなく総粉じん濃度を用いるようになっている。例えば重金属に絞った解析では、粉じん全体の組成を把握できないため、原典のままではPercent Massの評価が良くなることはない。

#### 4.7 Frac. Est.

Frac. Est. (FracEst: Fraction Estimable) は、推定計算に採用された発生源数の内、評価に耐える発生源の割合とした。原典のCMB8では、分母にデータ表の発生源数を用いているが、最初に範囲指定する発生源数はユーザーの作為で増減可能であるため、CMB8Jでは推計計算に採用された発生源の精度のみに着目したものである。評価に耐える発生源であるかどうかの判断は、以下に示す最大不確実性基準と最小射影基準によってなされる。FracEstは0~1の値をとるので、(4)式右辺第4項に示すとおり0.8で割って若干の加工を行い、その項が1.0以上なら良好と見なせるようにした。

#### 4.8 最大不確実性基準

上記のFracEstの評価に用いる基準の一つである。発生源プロファイルを特異値分解したときに得られる特異値の逆数と、ここで設定される最大不確実性基準とを比較し、逆数が基準より小さい特異値だけに対応した固有ベクトルで張られる線形空間を適格空間という<sup>7)</sup>。最大不確実性基準を小さくすれば、逆数がそれより大きい特異値が切り捨てられ、適格空間の次元が減り、評価に耐える発生源の数が制限を受ける。CMB8では、不確実性基準は、Percent Massと同じように総粉じん濃度に対して経験的に20%という既定値を持つが、ここでは各成分の合計濃度に対する評価として5%を既定値としている。評価を緩めるなら5%より大きな値に設定すればよい。

#### 4.9 最小射影基準

4.8の最大不確実性基準によって評価された適格空間への寄与ベクトルの射影について、ここで設定した最小射影基準より小さな場合に、その発生源は評価に耐え得ないものと判断される。射影値の最大値は1で、発生源の間に過度の共線性がある場合や、個々の発生源プロファイルに不確実性が大きい場合に射影値は1より小さな値となる。通常、経験的に0.95という値を既定値としているが、評価を緩くしたければそれより小さな値に設定すればよい。

Table 2. Detailed information displayed in the elimination of negative contributions.

	Weight				
$\chi^2 =$	60.69	1	$\chi^2(\phi = 5, \alpha = 0.05) = 11.07$		
$r^2 =$	0.9354	1			
% Mass =	67.41	1			
Frac.Est. =	1	1			
Modified pseudo-inverse normalized					
	A-1	A-2	B-1	B-2	B-3
Ni	0.1025				
Fe	0.1174				
Pb	0.0102				
Zn	0.7441				
Mn	1				
Cu	0.0088				
Estimable sources (Max.Src.Unc.= 10%, Min.Src.Proj.= .95)					
	A-1	A-2	B-1	B-2	B-3
Proj.	1				
Estimable	TRUE				
Process of elimination of negative source contribution					
	A-1	A-2	B-1	B-2	B-3
Step0	1.2765	0.1186	-0.2680	0.5365	-0.1660
Step1	1.3501	0.1027		0.1093	-0.1063
Step2	1.2502	-0.0134		-0.0021	
Step3	1.2427			-0.0031	
Step4	1.0336				

#### 4.10 詳細情報表示

モデルの検証を行うために必要な詳細情報を表示するかどうかを設定する。「負の発生源除去」ではその除去過程 (Table 2) を、「最適発生源選択」では最適な発生源の組み合わせベスト10 (Table 3) を表示する。その他共通な情報としては、 $\chi^2$ などの4指標値、規格化され転置された擬似逆行列が表示される<sup>8)</sup>。これは一般にMoore-Penroseの一般逆行列と呼ばれているものである<sup>9)</sup>。この転置擬似逆行列の各列で1又は-1を示す化学種は、その発生源の中で最も影響の強い成分であることを示す。

### 5. CMB8Jの適用例

#### 5.1 粉じん苦情への適用<sup>10)</sup>

八戸市内で平成11年3月に発生した粉じん苦情に際し、CMBを適用した。苦情発生直後に、粉じんが飛散した地点の周辺工場に立ち入りヒヤリングを行ったが、操業状態に異常のある発生源は見あたらなかった。このままでは再発も考えられたため、取り急ぎ苦情者の付近の堆積粉じんと、発生源と考えられる3社の金属精錬工場の集塵機ダストとを採取し、重金属の分析を行った。粉じんおよびダストは、灰化せずに塩酸・過酸化水素分解を施し、抽出液のフレイム原子吸

Table 3. Detailed information displayed in the Best-Fit.

	Weight					
$\chi^2 =$	3.797	1				$\chi^2(\phi = 4, \alpha = 0.05) = 9.488$
$r^2 =$	0.9992	1				
% Mass =	80.76	1				
Frac. Est. =	1	1				
Modified pseudo-inverse normalized						
	A-1	A-2	B-1	B-2	B-3	
Ni	0.1054					-0.0899
Fe	0.1200					-0.1016
Pb	0.0059					0.0015
Zn	-0.0316					1
Mn	1					-0.8224
Cu	0.0080					-0.0044
Estimable sources (Max.Src.Unc.= 5%, Min.Src.Proj.= .95)						
	A-1	A-2	B-1	B-2	B-3	
Proj.	1					1
Estimable	TRUE					TRUE
Best 10 of Fit Measure						
	A-1	A-2	B-1	B-2	B-3	FitMeasure
No.1	1.247					-0.003 1.800
No.2	1.243			-0.003		1.771
No.3	1.241		-0.002			1.769
No.4	1.265	-0.040				1.707
No.5	1.370	0.091	0.069		-0.090	1.704
No.6	1.250	-0.013		-0.002		1.670
No.7	1.252	-0.019	-0.001			1.666
No.8	1.233	0.030			-0.005	1.665
No.9	1.350	0.103		0.109	-0.106	1.339
No.10	1.376			0.094	-0.086	1.306

光分析で Ni, Fe, Pb, Zn, Cd, Mn および Cu を定量した。また、測定値の標準偏差は、ブランクとして用いた希硝酸の繰り返し測定結果の標準偏差を用いた。

解析に CMB8J による有効分散最小自乗法を適用したところ、Table 4 が得られた。その寄与率データと相関行列 (Table 6) から、A-1 と A-2 の間および B-1 と B-2 の間に多重共線性があることが認められた。

これは、粉じん発生工程は異なるものの同じ原料から得られたダストであるため、組成が類似したものと考えられた。共線性の強い A-1, A-2 の内プラスを示す A-1 を残して推定計算し直したところ、Table 5

および Fig. 2 が得られた。この結果を元に A 社に再点検を依頼したところ、A 社では、A-1 の集塵ダストを炉に戻す配管の途中で点検口が開いている箇所を発見し、そこからのダスト飛散を認めた。A 社は、その点検口を密閉する対策を講じ、苦情は解決した。解析によって工場内の発生源をある程度絞り込むことができたのも、CMB 法によるところが大きい。

一般的な環境汚染を CMB 法で解析する場合には、発生源データとして、海塩、土壌、自動車排ガスなどを加える必要があるが、今回粉じん苦情の発生源特定が目的であったため、周囲のばい煙、粉じん発生源に限定してデータを集めて解析を行ったものである。

6. おわりに

今回、自作ホームページ<sup>11)</sup>上で公開している環境データ解析用 Excel マクロの中から、CMB 法解析マクロを紹介した。原典である CMB8 は、他の手法より優れた有効分散最小自乗法を用いているが、原作者によるとその総合指標などの設定は、実験的に導入した

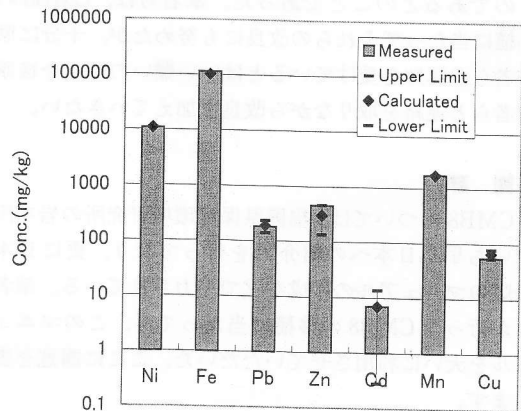


Fig. 2. Comparison of calculated concentrations with measured.

Table 4. Result of CMB apportionment.

Species	Source(mg/kg)					Ambient(mg/kg)		
	A-1	A-2	B-1	B-2	B-3	Measured(M)	Calculated(C)	C/M (%)
	Conc.	Conc.	Conc.	Conc.	Conc.	Conc.	Conc.	Conc.
Ni	15100	12100	110	82.0	54.0	10900	11100	101%
Fe	146000	85100	136000	138000	13600	114000	102000	89.0%
Pb	282	4770	28300	21500	258000	183	188	103%
Zn	1470	22300	365000	273000	388000	459	298	65.0%
Cd	59.0	388	567	376	569	7.00	7.00	99.0%
Mn	2780	2330	11700	9890	312	1720	1720	100%
Cu	52.0	57.0	2260	1630	1020	55.0	67.0	121%
Contribution(R)	0.862	-0.161	0.233	-0.306	0.003			
Ratio of R(%)	136.7%	-25.6%	37.0%	-48.6%	0.4%			

Table 5. Result of CMB apportionment without a source of A-2.

Species	Source(mg/kg)				Ambient(mg/kg)		
	A-1	B-1	B-2	B-3	Measured(M)	Calculated(C)	C/M(%)
	Conc.	Conc.	Conc.	Conc.	Conc.	Conc.	Conc.
Ni	15100	110	82.0	54.0	10900	9350	85.8%
Fe	146000	136000	138000	13600	114000	90200	79.1%
Pb	282	28300	21500	258000	183	140	76.3%
Zn	1470	365000	273000	388000	459	459	100%
Cd	59.0	567	376	569	7.00	35.8	512%
Mn	2780	11700	9890	312	1720	1710	99.2%
Cu	52.0	2260	1630	1020	55.0	29.4	53.5%
Contribution(R)	0.619	-0.001					
Ratio of R(%)	100.2%	-0.20%					

Table 6. Correlation matrix between sources.

	A-1	A-2	B-1	B-2	B-3
A-1	1.000				
A-2	0.969	1.000			
B-1	0.171	0.402	1.000		
B-2	0.295	0.514	0.992	1.000	
B-3	-0.250	-0.033	0.760	0.706	1.000

ものであるとのことであった。筆者らは、CMB8の移植に当たってそれらの改良にも努めたが、十分に原作者らの認知を受けているとはいえないので、今後原作者らと連絡を取りながら改良を加えていきたい。

### 謝 辞

CMB8については、福岡県保健環境研究所の岩本氏がいち早く日本への紹介<sup>12)</sup>を行っており、更に日本語版のマニュアルの作成などで尽力されている。筆者らが行ったCMB8の移植に当たっても、このマニュアルを大いに利用させていただいた。ここに謝意を表します。

(受稿 2000.5.1)

### 文 献

- 溝畑 朗他：堺市における大気浮遊粒子状物質中の諸元素の発生源の同定, 大気汚染学会誌, **15**, 198-206 (1980).
- 浮遊粒子状物質対策検討会：浮遊粒子状物質汚染予測マニュアル, 267-269, 東洋館出版社, 1997.
- Henry, R. C. *et al.*: Review of receptor model fundamentals, *Atmos. Environ.*, **18**, 1507-1515.
- 環境庁大気保全局：スパイクタイヤによる浮遊粉じん中の各成分濃度等測定方法指針, 1983.
- Watson, J. G. *et al.*: The effective variance weighting for least squares calculation applied to the mass balance receptor model, *Atmos. Environ.*, **18**, 1347-1355.
- <ftp://eafs.dri.edu/cmb80>
- Henry, R. C.: Dealing with near collinearity in chemical mass balance receptor models. *Atmos. Environ.*, **26A**, 933-8 (1992).
- Watson, J. G. *et al.*: The USEPA/DRI chemical mass balance receptor model, CMB 7.0, *Environmental Software*, 1990, **5**, 38-49.
- 中川 徹他：最小二乗法による実験データ解析プログラム SALS, p. 65, 東京大学出版会, 1982.
- 花石竜治他：青森県環境保健センター研究報告, **10**, 45-48 (2000).
- <http://www.jomon.ne.jp/~hayakari/>
- 岩本真二：CMB8ソフトによる浮遊粒子状物質発生源寄与の推定, 大気環境学会講演要旨集, **39**, 292 (1998).

# Introduction of Spreadsheet Application Macros for Environmental Data Analysis and Analytical Examples (II)

—Macro for CMB Analysis—

Susumu HAYAKARI\* and Ryuji HANAISHI\*\*

\* Aomori Prefectural Institute of Environment and Health-Mutsu Environmental Management Office.  
1-1-8, Chuoh, Mutsu, Aomori 035-0073, JAPAN

\*\* Aomori Prefectural Institute of Environment and Health. 1-1-1, Higashitsukurimichi Aomori, Aomori 030-8566, JAPAN

Introduction of Microsoft Excel macros for environmental data analysis that one (SH) of the present authors has uploaded to web site (<http://www.jomon.ne.jp/~hayakari/>).

CMB8J, a macro for the CMB analysis, was made by translation of a freeware coded in USA for the CMB apportionment and has some advantages. The CMB8J yields detailed information about the processes of choice of sources and it propose a new performance measure for selection of a set of sources that gives a nice fitting.

**Key words:** web site, Excel VBA macro, CMB