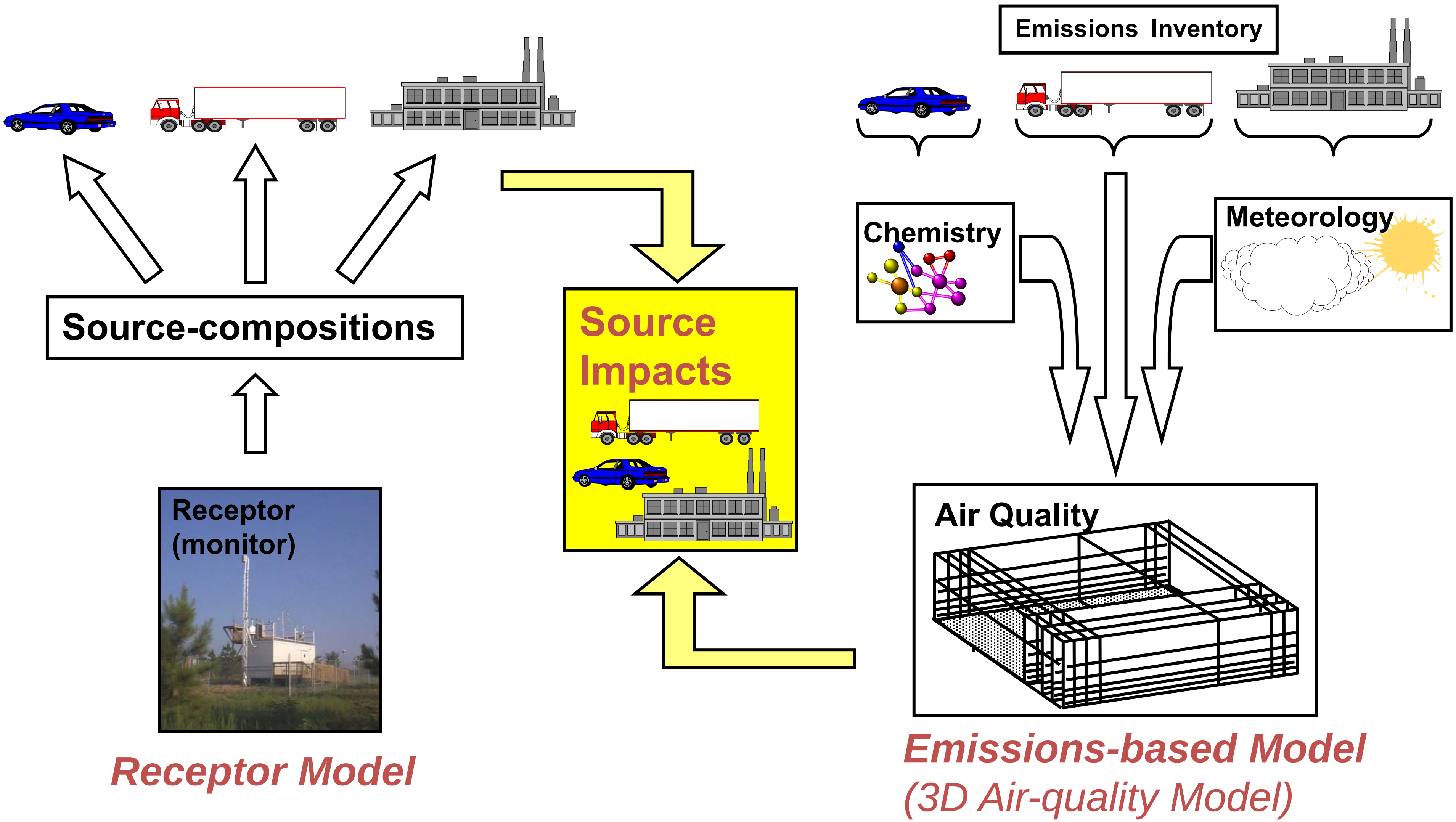
Receptor vs. Emissions-Based Models



Source oriented: date le caratteristiche della sorgente e i dai meteorologici, stima della concentrazione degli inquinanti nei siti recettore

Receptor oriented: date le misure delle concentrazioni degli inquinanti nel sito recettore, stima del contributo delle differenti sorgenti e *source apportionment*

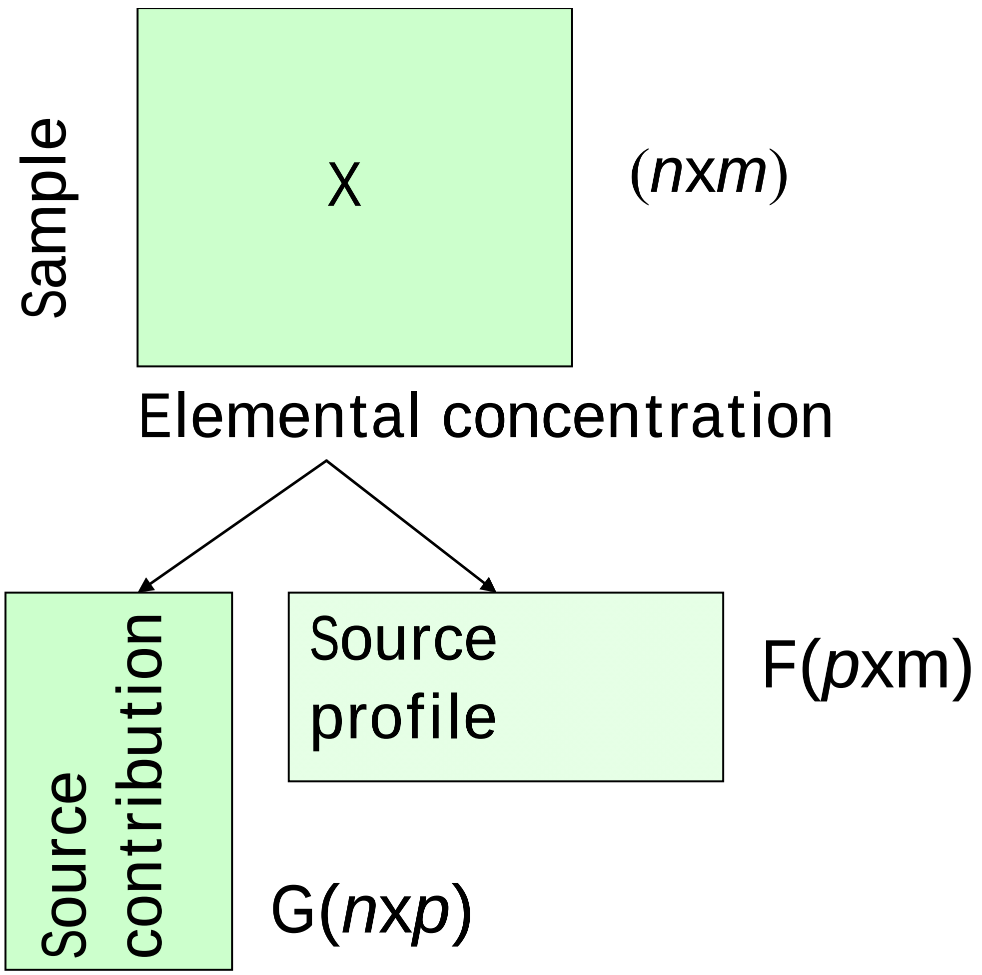
Receptor Models

CMB
known sources

MULTIVARIATE METHODS
unknown sources
SA
- PCA
- TTFA
- Cluster analysis
- PMF
- SAFER…

$$X = GF + E$$

Receptor Model



X = GF+E

THERE ARE THREE MAIN GROUPS OF SOURCE APPORTIONMENT TECHNIQUES

(a) Methods based on the evaluation of monitoring data. Basic numerical data treatment is used to identify sources.

Examples are:

(1) correlation of wind direction with levels of measured components to identify source locations;

(2) the correlation of gaseous pollutants with PM components to identify source associations;

(3) subtraction of levels measured at regional background from those obtained at urban background and/or roadside levels to identify the contributions from the regional background, the city background and the monitored street;

(4) quantification of natural PM contributions (e.g., African dust) by subtracting PM levels at regional background sites from those at urban background locations for specific days.

The main advantage is the simplicity of the methods and the consequent low impact of mathematical artefacts due to data treatment.

(b) Methods based on emission inventories and/or dispersion models to simulate aerosol emission, formation, transport and deposition. These models require detailed emission inventories that are not always available, and they are limited by the accuracy of emission inventories, especially when natural emissions are important. A significant advantage of these methods is that they may be used in scenario studies to evaluate the impact of emission abatement strategies on the anthropogenic contribution to ambient PM concentrations.

(c) Methods based on the statistical evaluation of PM chemical data acquired at receptor sites (receptor models). The fundamental principle of receptor modelling is that mass and species conservation can be assumed and a mass balance analysis can be used to identify and apportion sources of airborne PM in the atmosphere.

OVERVIEW OF THE WIDE RANGE OF STATISTICAL MODELS AND MODELLING APPROACHES WHICH ARE CURRENTLY AVAILABLE IN THE LITERATURE.

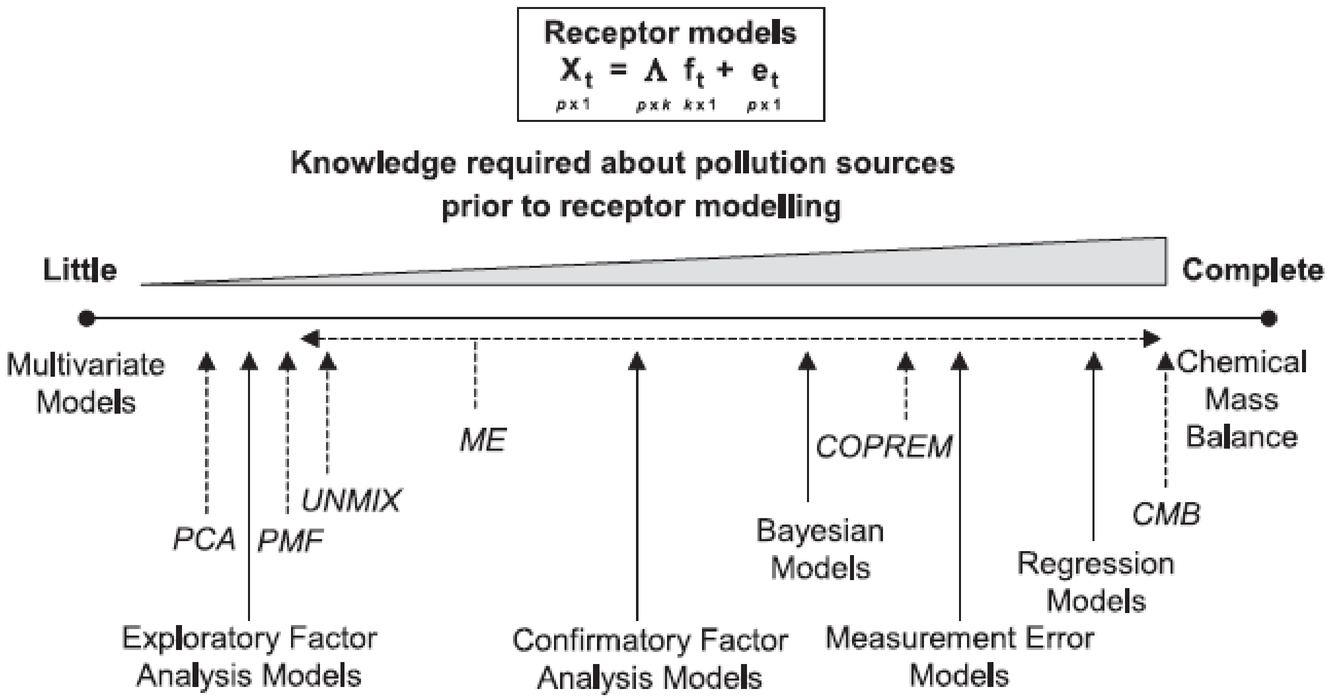
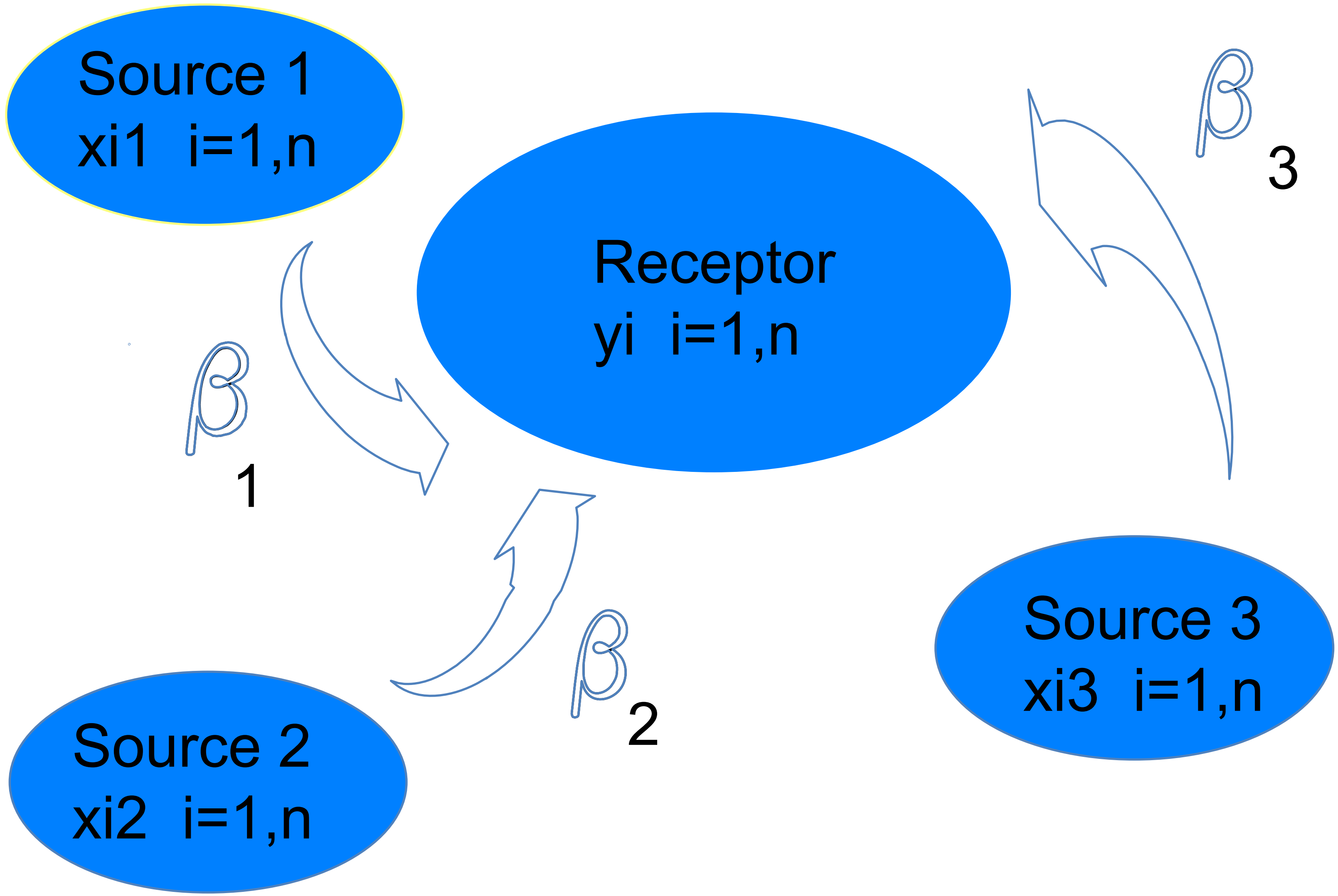


Fig. 1. Approaches for estimating pollution source contributions using receptor models (modified from Schauer et al., 2006). Specific models are shown in italics and with dotted arrows.

Chemical Mass Balance Receptor Modelling



PARTICULATE MATTER

European standard for the environmental air quality and for a cleaner air in Europe (2008/50/CE)

Standard limits for **PM_{10}**:

- 40 µg/m^3 as yearly average
- 50 µg/m^3 as daily average (not to be exceeded more than 35 times a year)

PM_{10} concentrations to reduce within 2010 (1999/30/EG and 96/62/EG)

<u>1° STEP</u>: actual situation of **PM_{10}** in the Lombardy Region

<u>2° STEP</u>: sources of PM_{10} in the Lombardy Region

PM SOURCE IDENTIFICATION is a need

1° *n*-ALKANES

1. Spatial distribution
2. Seasonal distribution
3. Dimensional distribution

INORGANIC IONS

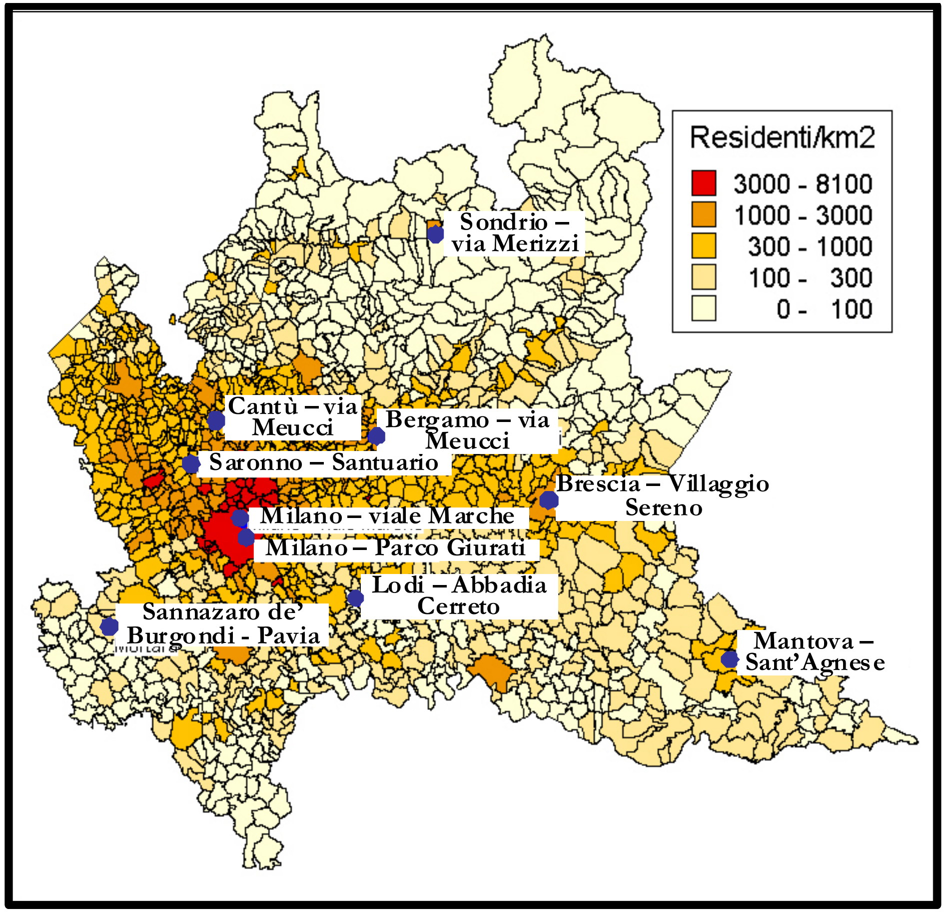
1. Spatial distribution

2° DIESEL EMISSIONS

PM SAMPLING

Ten sites of the Lombardy Region, from urban background to regional (pre-alpine) air quality.

- Low volume sampler, dual channel, Hydra (FAI Instruments); PM_{10} inlet cut
- 24h sampling (0:00 to 24:00)
- Total volume = 27 - 55 m^3
- PTFE filters, diameter of 47 mm



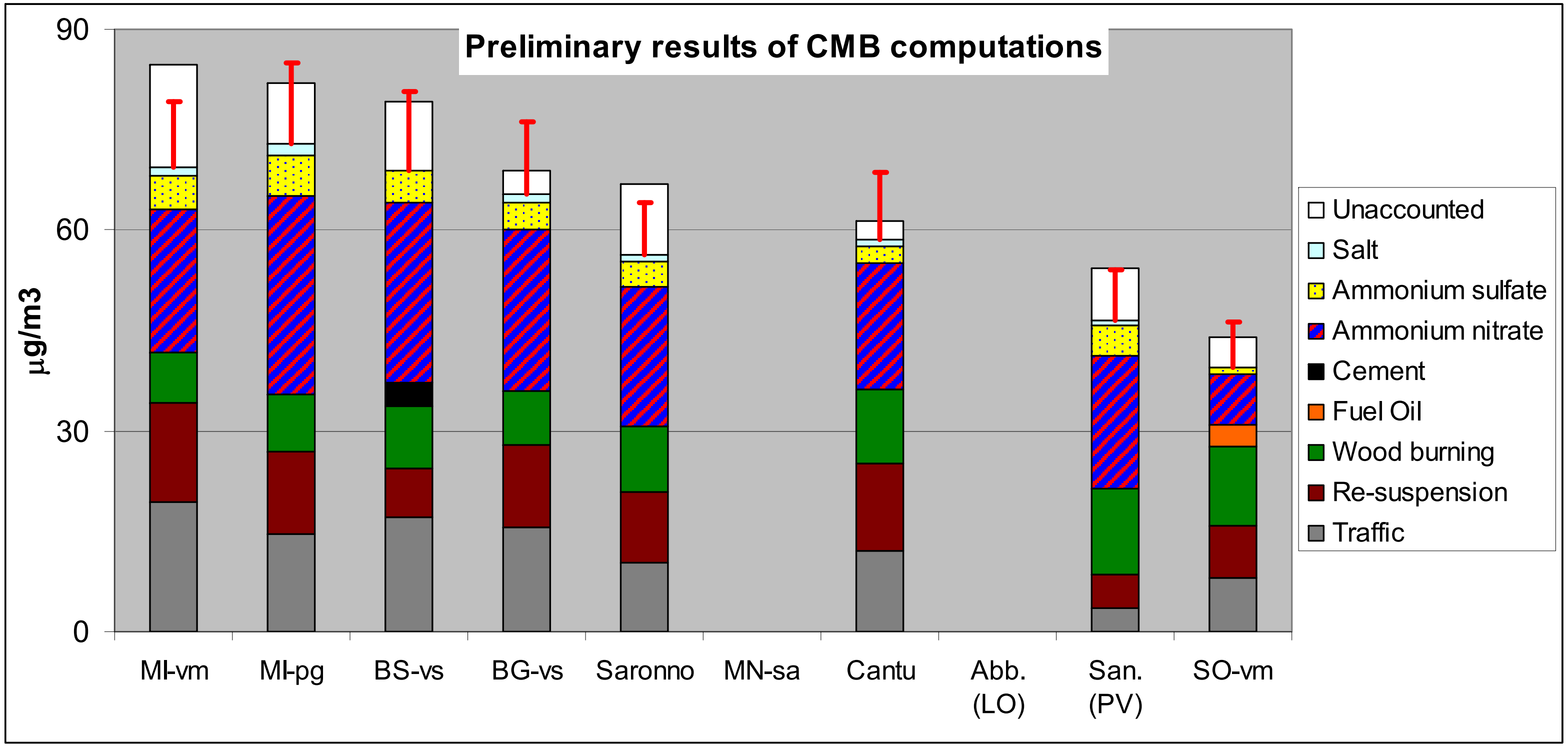
Sampling during **winter (feb-mar) 2007**; 14<n<18 for a single site.

ANALYSIS:

- inorganic ions
- elementar and inorganic carbon
- levoglucosan
- polycyclic aromatic hydrocarbons
- trace elements
- linear alkanes (not included)

SOURCE APPORTIONMENT (Chemical Mass Balance)

PM SOURCE APPORTIONMENT



1. SECONDARY AEROSOL [NH_4NO_3+$(NH_4)_2SO_4$] = 30-45%
2. TRAFFIC+RESUSPENSION = 31-41%
3. WOOD BURNING = 10-27%

Chemical Mass Balance
Receptor Modelling

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + ... + \varepsilon_i$$

y_i : relative abundance of component i at the receptor

x_{i1} : relative abundance of component i at source 1

β_1 : fractional contribution of source 1

ε_i : the difference (error) between measured (y_i) and modelled ($\sum_j \beta_j x_{ij}$) relative abundances of component i

Chemical Mass Balance Receptor Modelling

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \ldots\ldots + \varepsilon_i$$

minimize the sum of the square of the errors:

$$\sum_i \varepsilon_i^2 = \sum_i \left(y_i - \sum_j \beta_j x_{ij} \right)^2$$

by adjusting the fractional contributions β_j

Discrepancy in the total mass balance:

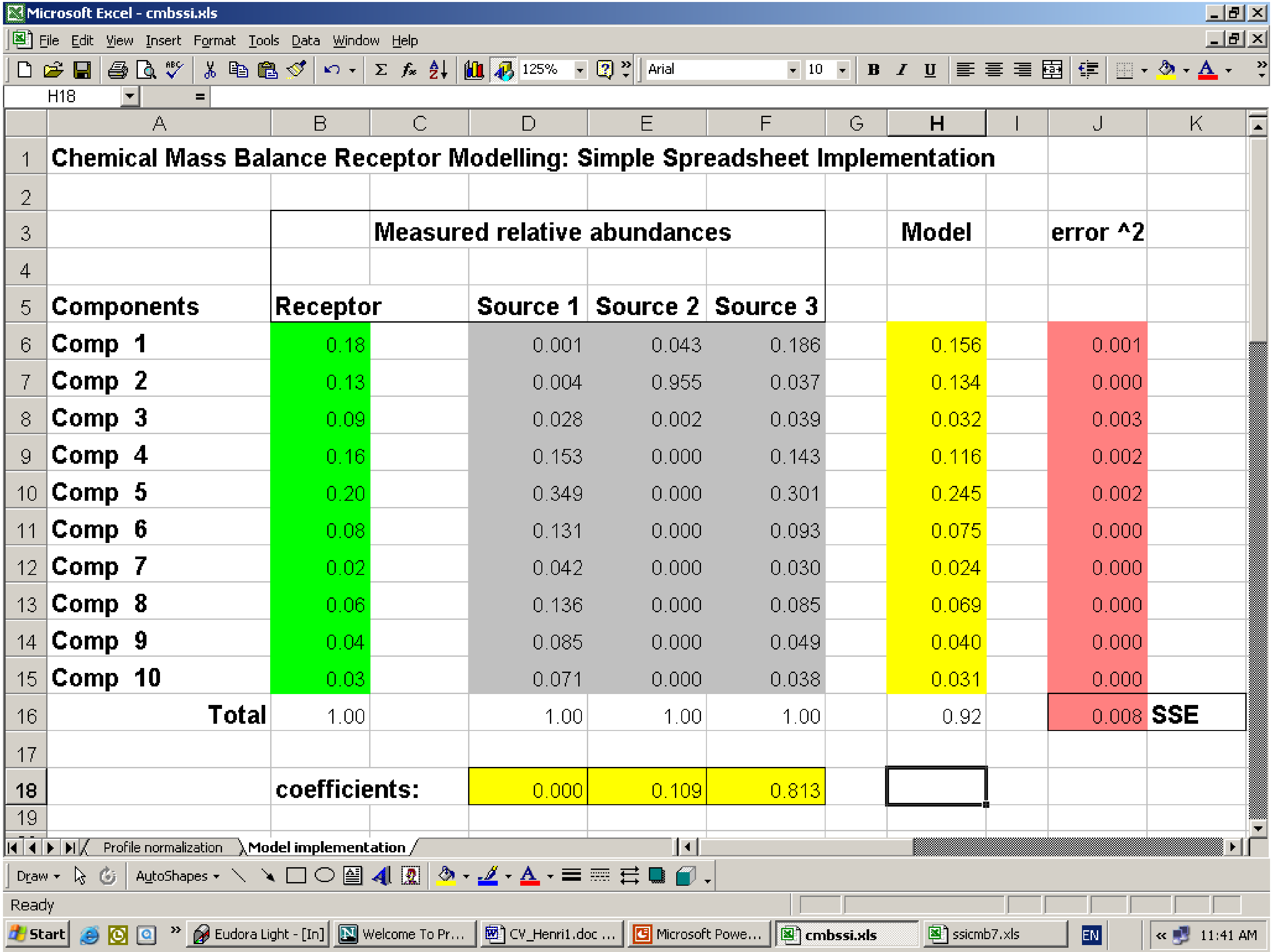
$$1 - \sum_j \beta_j$$

Chemical Mass Balance
Receptor Modelling
Simple Spreadsheet Implementation

- Electronic spreadsheets like Excel make it easy to implement the model in the above form.

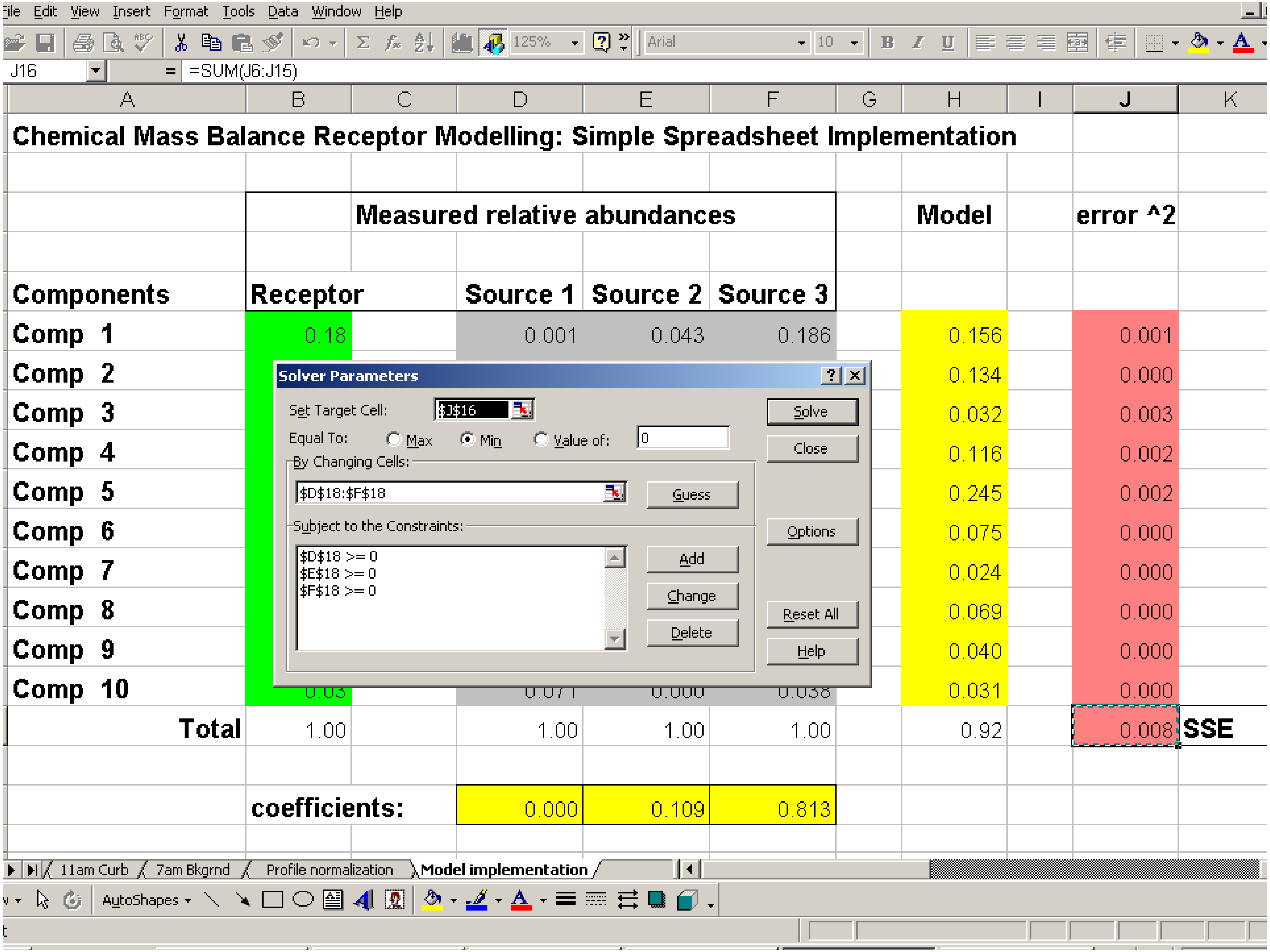
- The “Solver” function in Excel enables the minimization operation.

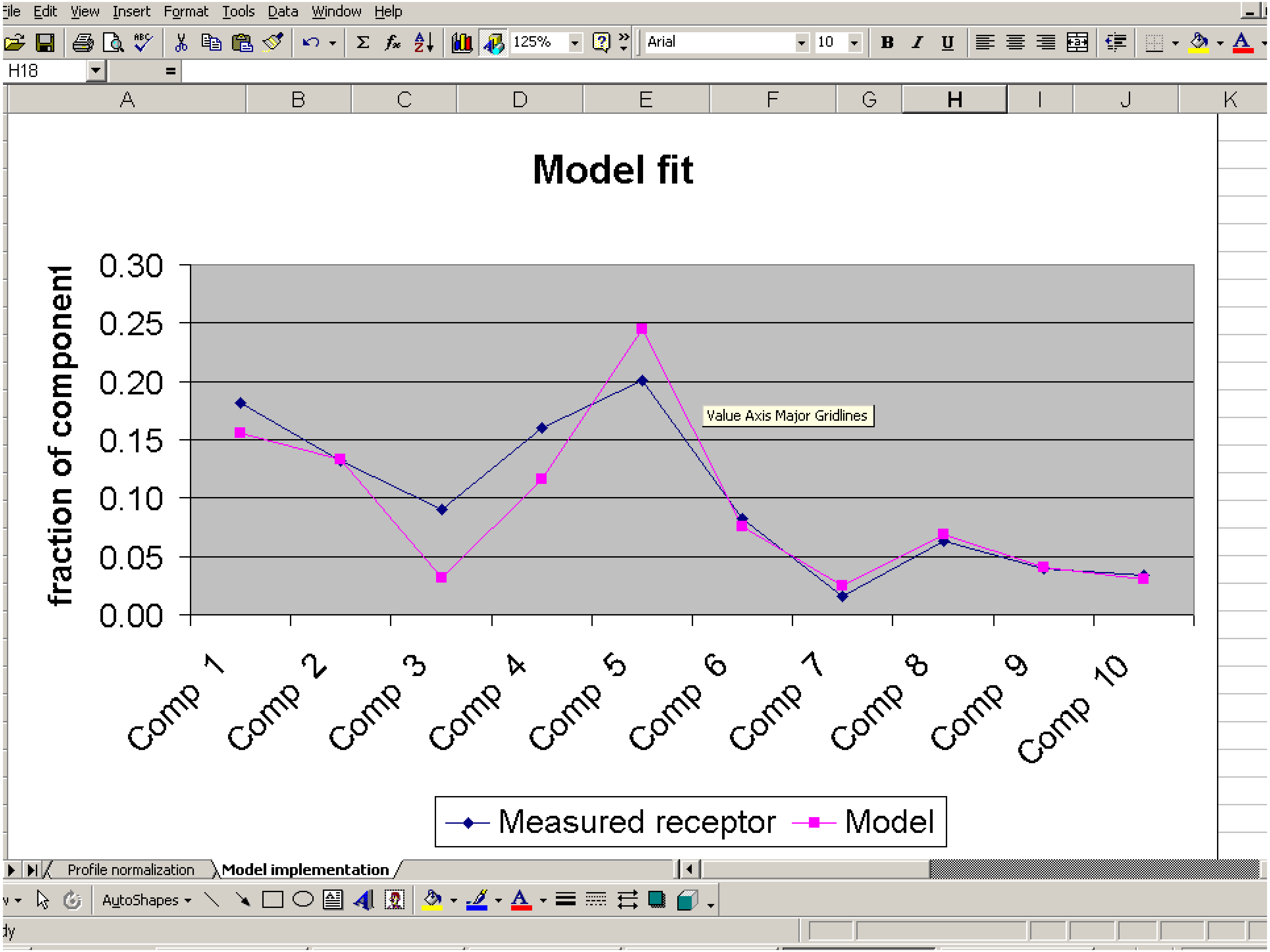
- Graphical tools in Excel enable qualitative assessment of model fit and the inputs.

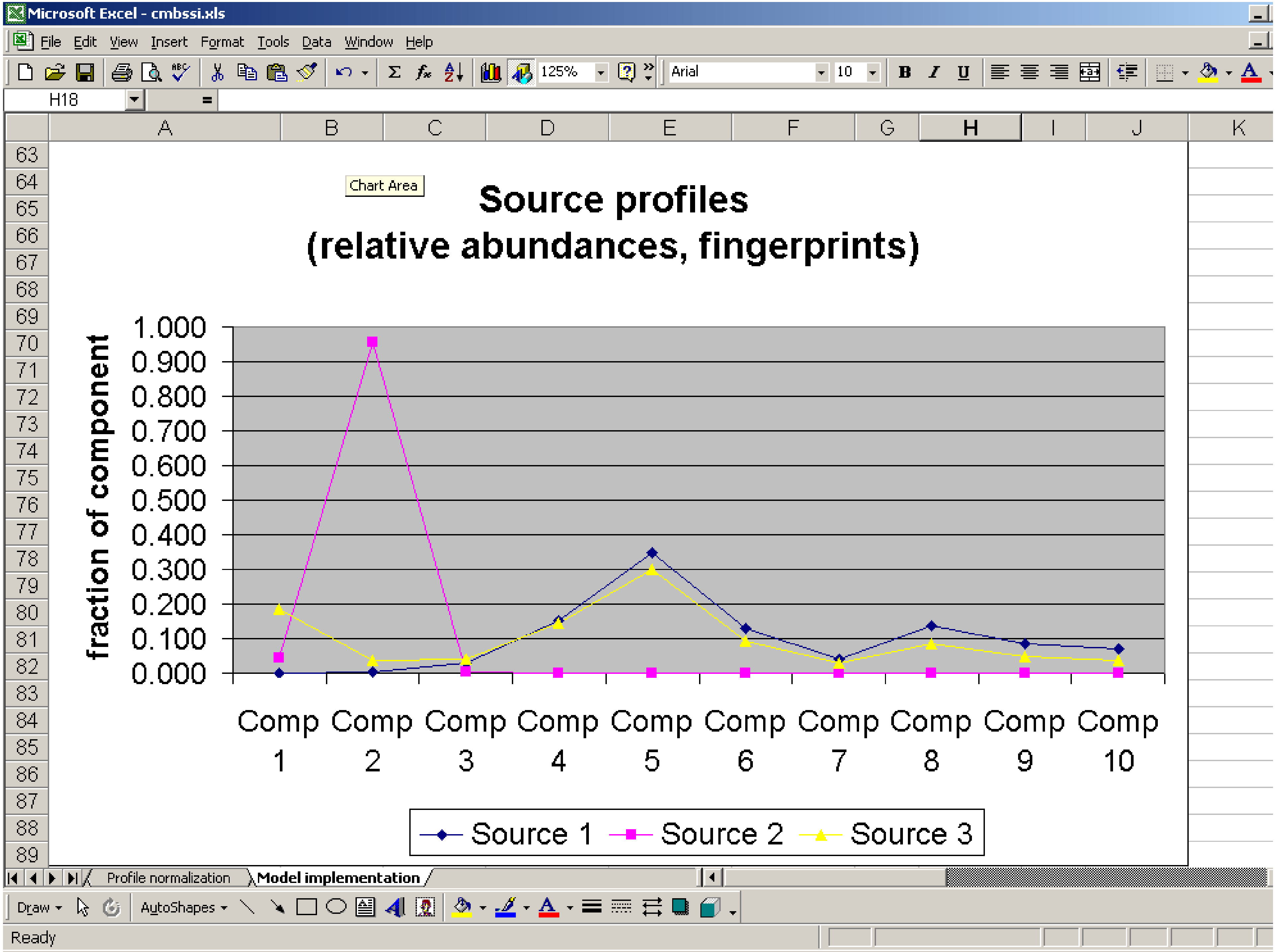


Chemical Mass Balance Receptor Modelling: Simple Spreadsheet Implementation

	Measured relative abundances				Model	error ^2	
Components	Receptor	Source 1	Source 2	Source 3			
Comp 1	0.18	0.001	0.043	0.186	0.156	0.001	
Comp 2	0.13	0.004	0.955	0.037	0.134	0.000	
Comp 3	0.09	0.028	0.002	0.039	0.032	0.003	
Comp 4	0.16	0.153	0.000	0.143	0.116	0.002	
Comp 5	0.20	0.349	0.000	0.301	0.245	0.002	
Comp 6	0.08	0.131	0.000	0.093	0.075	0.000	
Comp 7	0.02	0.042	0.000	0.030	0.024	0.000	
Comp 8	0.06	0.136	0.000	0.085	0.069	0.000	
Comp 9	0.04	0.085	0.000	0.049	0.040	0.000	
Comp 10	0.03	0.071	0.000	0.038	0.031	0.000	
Total	1.00	1.00	1.00	1.00	0.92	0.008	SSE
coefficients:		0.000	0.109	0.813			


File Edit View Insert Format Tools Data Window Help
J16 =SUM(J6:J15)
Chemical Mass Balance Receptor Modelling: Simple Spreadsheet Implementation
Measured relative abundances
Model
error ^2
Components
Receptor
Source 1
Source 2
Source 3
Comp 1
Comp 2
Comp 3
Comp 4
Comp 5
Comp 6
Comp 7
Comp 8
Comp 9
Comp 10
0.18
0.001
0.043
0.186
0.156
0.134
0.032
0.116
0.245
0.075
0.024
0.069
0.040
0.031
0.001
0.000
0.003
0.002
0.002
0.000
0.000
0.000
0.000
0.000
Total
1.00
1.00
1.00
1.00
0.92
0.008
SSE
coefficients:
0.000
0.109
0.813
Solver Parameters
Set Target Cell: $J$16
Equal To: Max Min Value of: 0
By Changing Cells:
$D$18:$F$18
Guess
Subject to the Constraints:
$D$18 >= 0
$E$18 >= 0
$F$18 >= 0
Add
Change
Delete
Solve
Close
Options
Reset All
Help
11am Curb
7am Bkgrnd
Profile normalization
Model implementation
AutoShapes


File Edit View Insert Format Tools Data Window Help
H18
Model fit
fraction of component
0.30
0.25
0.20
0.15
0.10
0.05
0.00
Value Axis Major Gridlines
Comp 1
Comp 2
Comp 3
Comp 4
Comp 5
Comp 6
Comp 7
Comp 8
Comp 9
Comp 10
Measured receptor
Model
Profile normalization
Model implementation
AutoShapes



Microsoft Excel - cmbssi.xls
File Edit View Insert Format Tools Data Window Help
Chart Area
Source profiles
(relative abundances, fingerprints)
fraction of component
1.000
0.900
0.800
0.700
0.600
0.500
0.400
0.300
0.200
0.100
0.000
Comp 1
Comp 2
Comp 3
Comp 4
Comp 5
Comp 6
Comp 7
Comp 8
Comp 9
Comp 10
Source 1
Source 2
Source 3
Profile normalization
Model implementation
Ready

Chemical Mass Balance Receptor Modelling U.S. EPA's CMB Model

- There are uncertainties associated with both the receptor and source profiles. The measured quantities are thus:
- The model results then also have uncertainties associated with them:

$$y_i \pm \Delta y_i \qquad \text{and} \qquad x_i \pm \Delta x_i$$

$$\beta_j \pm \Delta \beta_j$$

U.S. EPA's CMB Model
The "effective variance" method

The CMB model uses an effective variance which looks at variances in both the receptor, and source profiles weighted by the source contributions:

$$V_{ei} = \sigma_{Ri}^2 + \sum_j \beta_j^2 \sigma_{Sij}^2$$

$\sigma_{Ri}^2, \sigma_{Sij}^2$: variances associated with the ith component in the receptor and source profiles

V_{ei} : diagonal elements of the "effective variance matrix" used in the iterative search for β_j

U.S. EPA’s CMB8

- The CMB model incorporates the statistics to improve our estimate of the source contributions by weighting the source and receptor profiles according to the uncertainties associated with them.

- It also reports the uncertainties associated with the source contribution estimates and “goodness of fit” parameters.

Positive Matrix Factorizaton (PMF)

Other receptor models

CMB applies the mass balance principles to *known* receptor and source compositions to arrive at the source contribution estimates.

Another type of source apportionment analyzes multiple measurements at the receptor (i.e. many hours of hourly average data) *without prior knowledge of the sources* but attempts to identify "factors" that explain the variation in the composition. The factors are then interpreted as types of sources.

PMF Characteristics

- Method: Weighted least-squares
- Utilize error estimates of the data to optimum data point scaling
- Non-negativity constraints
- Does not require comprehensive advance information on source compositions
- Incorporate the time variation
- Obtain uncertainties for source composition and source contribution output profiles

PMF Characteristics

- Input:
  - Ambient concentration data
  - Uncertainty of ambient data
- Output
  - Source compositions (F-factor)
  - Source contributions (G-factor)
  - Scaled residuals (e_{ij}/s_{ij})

Methodology Approach

- Data handling/preparation
- Run PMF
- Interpretation of PMF results
- Source identification and source apportionment
- Optional examination of variations of source contributions and meteorological effects

$$X = GF + E$$

X: $n \times m$ data matrix measured at the receptor, of
m compounds and n observations

G: $n \times p$ matrix of the time variation of source contributions,
n periods and p sources

F: $p \times m$ matrix of source (factor) compositions,
p sources, m compounds

E: $n \times m$ matrix of residuals

Given X, PMF attempts to find G and F with all positive elements to minimize the elements of E, i.e, to minimize:

$$Q(E) = \sum_{i=1}^{n}\sum_{j=1}^{m}\left(\frac{e_{ij}}{s_{ij}}\right)^2$$

where e_{ij} are the elements of the E matrix,

s_{ij} are the error estimates for the elements of the X matrix

Data File

Sample No.	Al	As	Br	Ce	Cl
1	1016.86	0.56	-1.29	0.59	675.83
2	853.37	2.61	9.63	0.71	915.45
3	822.65	0.99	8.23	0.31	567.13
4	1574.65	1.68	18.03	1.93	710.76
5	1074.94	0.97	11.23	0.85	693.22
6	2497.35	3.42	38.70	2.03	534.36
7		1.56	8.91	1.16	399.18
8	1057.57	1.02	9.72	1.02	875.76
9	998.97	1.43	40.43	0.93	< 100
10	1719.75	1.85	51.05	1.55	1290.81

Data Handling

	Data	Unc.
• Missing data ➢	$\bar{x}_{ij}$	$4\ \bar{x}_{ij}$
• < DL ➢	1/2 DL	DL
• Negative values ➢	Real data	

PMF Operation

- PMF2 program
  - Dr. Pentti Paatero, University of Helsinki, Finland
- EPA PMF 1.1 using ME2 program
  - Dr. Shelly Eberly, US EPA, U.S.A.

PMF Running: Trial and error

- Varying the number of factor
- Varying parameters/optional functions
- Graphical interpretation of obtained results of G and F factors
- Decision make for the number of source factor to be ratained

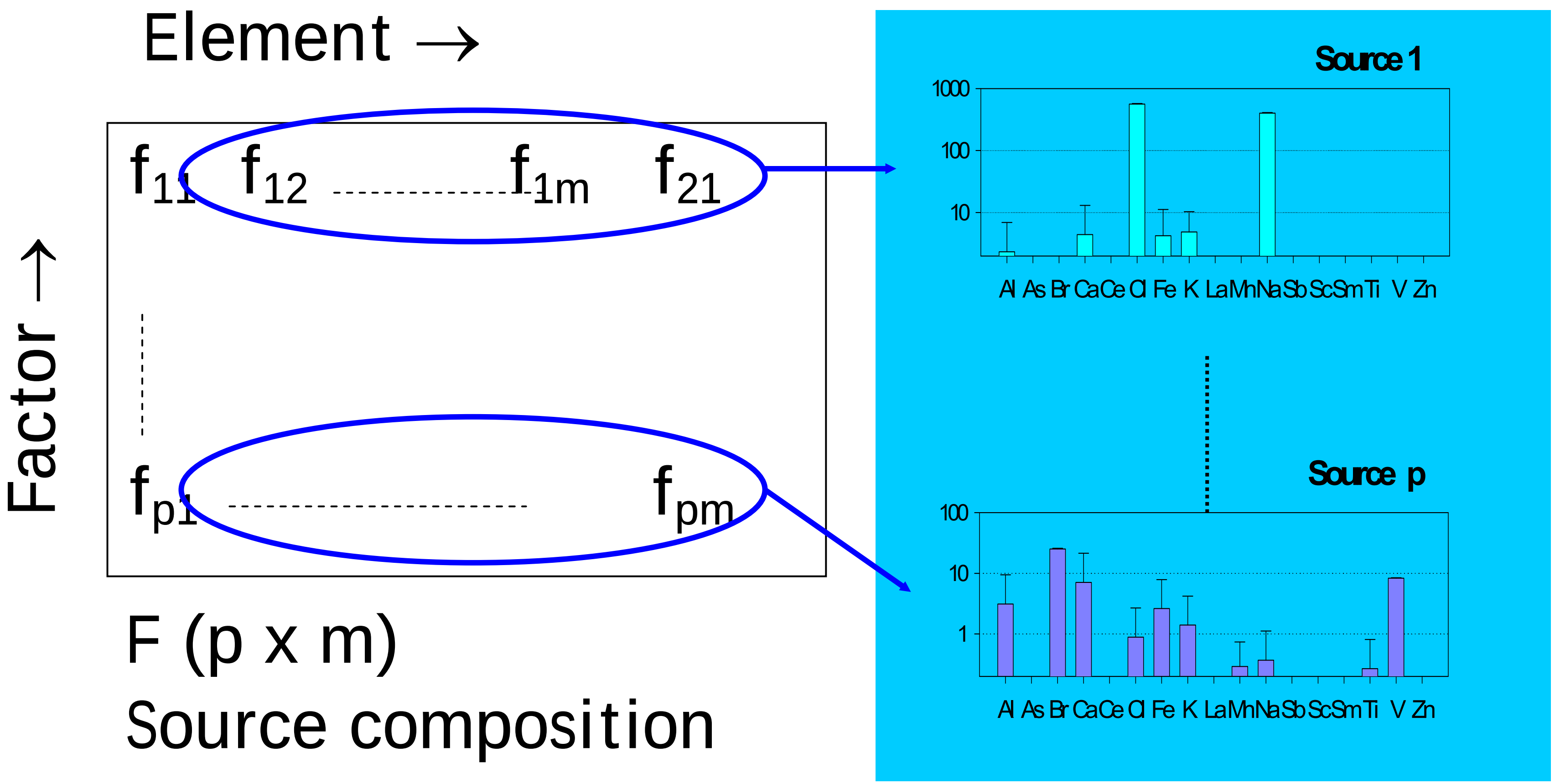
Determination of the Number of Sources

- Too few factors will combine sources of different nature together
- Too many factors will make a real factor dissociate into two or more non-existing sources

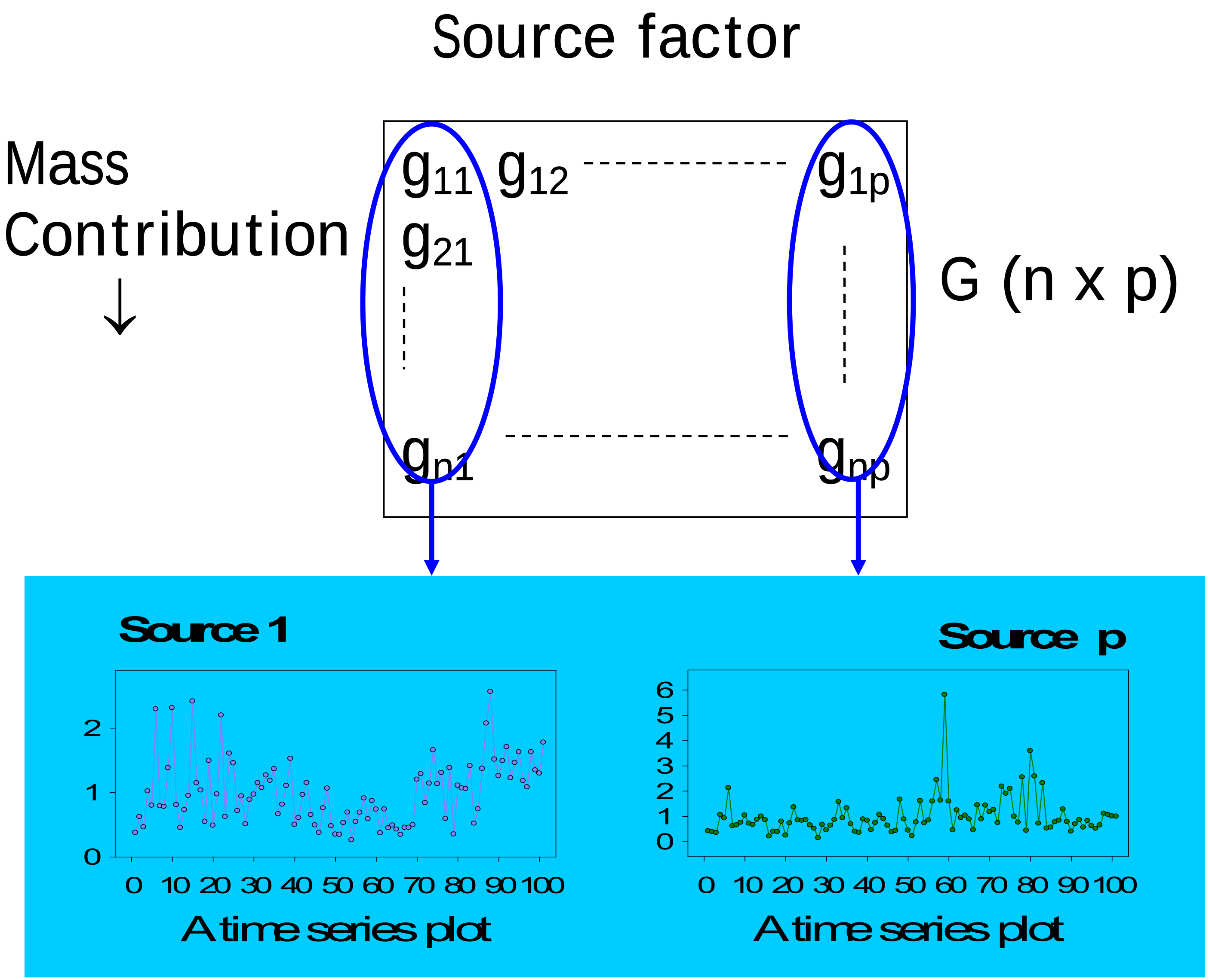
- A good fit of Q value
- The weighted residuals of the model ($e_{ij}/s_{ij} \sim \pm 2$)
- The interpretable and most meaningful factors

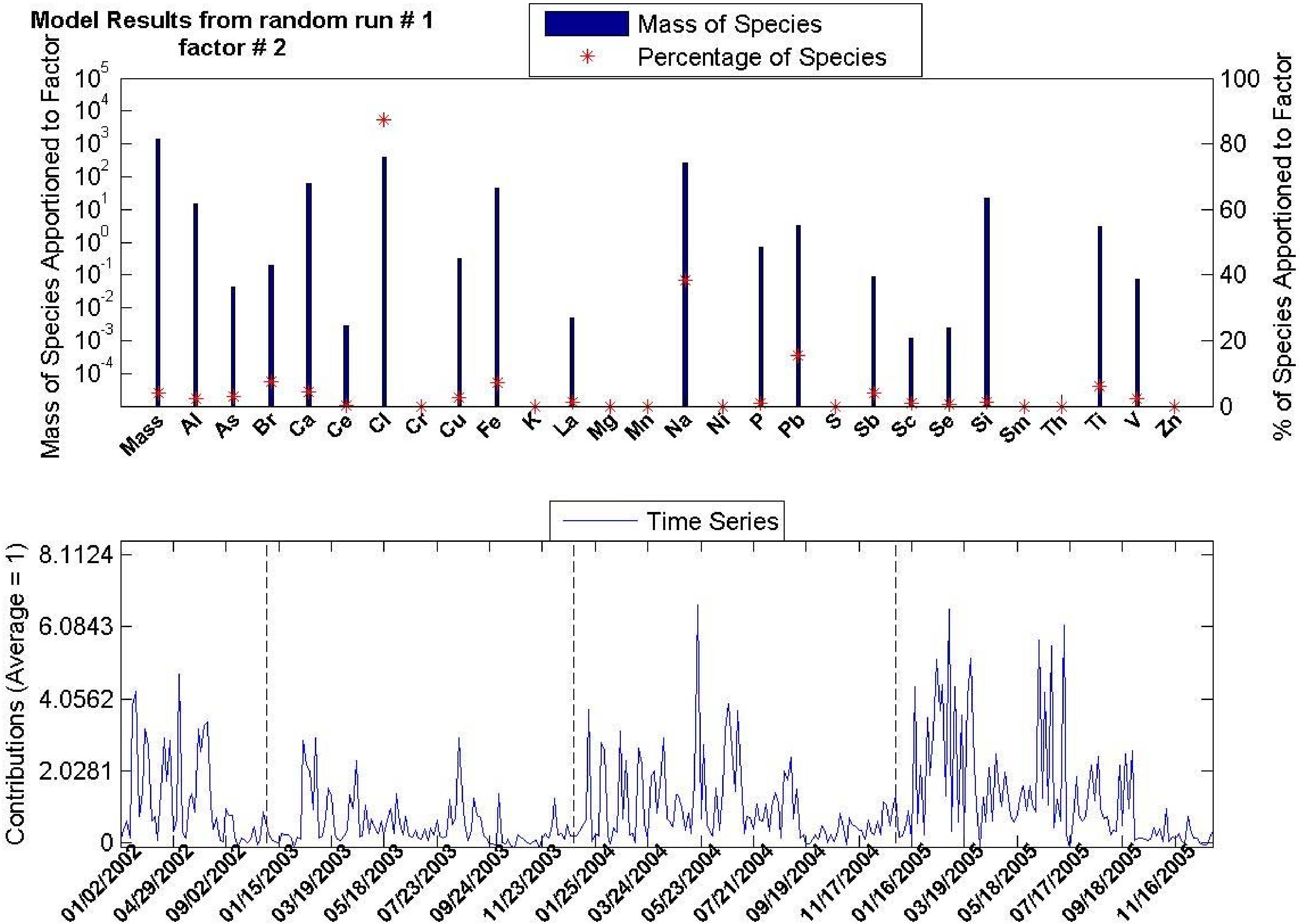
PMF Results :
Source Identification/Fingerprint

Source Identification



Source Contributions to Samples





Model Results from random run # 1
factor # 2
Mass of Species
Percentage of Species
Mass of Species Apportioned to Factor
% of Species Apportioned to Factor
10^5
10^4
10^3
10^2
10^1
10^0
10^-1
10^-2
10^-3
10^-4
100
80
60
40
20
0
Mass
Al
As
Br
Ca
Ce
Cl
Cr
Cu
Fe
K
La
Mg
Mn
Na
Ni
P
Pb
S
Sb
Sc
Se
Si
Sm
Th
Ti
V
Zn
Time Series
Contributions (Average = 1)
8.1124
6.0843
4.0562
2.0281
0
01/02/2002
04/29/2002
09/02/2002
01/15/2003
03/19/2003
05/18/2003
07/23/2003
09/24/2003
11/23/2003
01/25/2004
03/24/2004
05/23/2004
07/21/2004
09/19/2004
11/17/2004
01/16/2005
03/19/2005
05/18/2005
07/17/2005
09/18/2005
11/16/2005

Airborne Contributions of Certain Marker Species

Source	**Elements**
Soil	Al, Si, Ca, Sc, Ti, Fe, Mn, K
Cement/Construction	Ca, Mg
Sea-salt	Na, Cl, Mg
Motor vehicles	Br, Pb, Zn, C
Refuse incineration	Sb, Zn, Cd, Ag, Sn, Pb
Wood burning	K, C
Oil combustion	V, Ni, Rare earths
Coal combustion	As, Se, S, C, K
Sulfide smelters	In, Cd, As, Se, S

PMF Results :
Source Apportionment

Source Apportionment

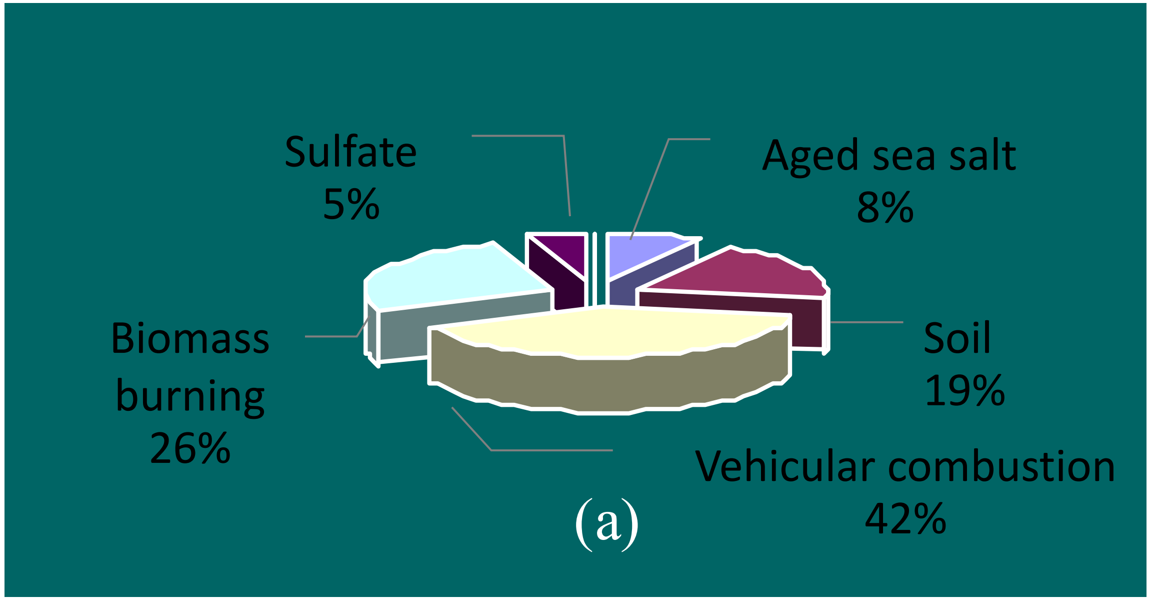
- Source contribution factor (G factor) from PMF result
- MLR analysis program
  - Excel
  - StatGraphics
- Graphical interpretation
  - Excel
  - SigmaPlot
  - StatGraphics

Multilinear Regression

- Calculates the statistics for a line by using the "least squares" method to calculate a straight line that best fits your data
- The equation for the line is:

$$y = mx + b$$

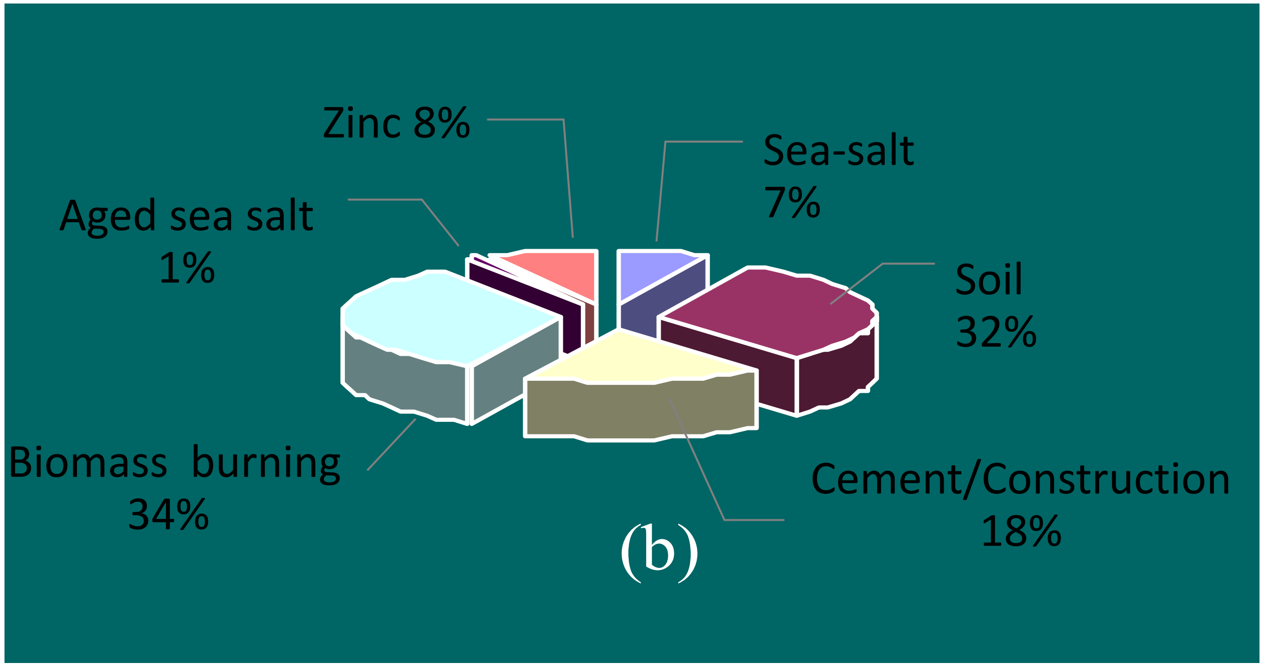
PM mass = $a_1G_1 + a_2G_2 + ... + a_kG_k$



(a)

FPM At Bkk

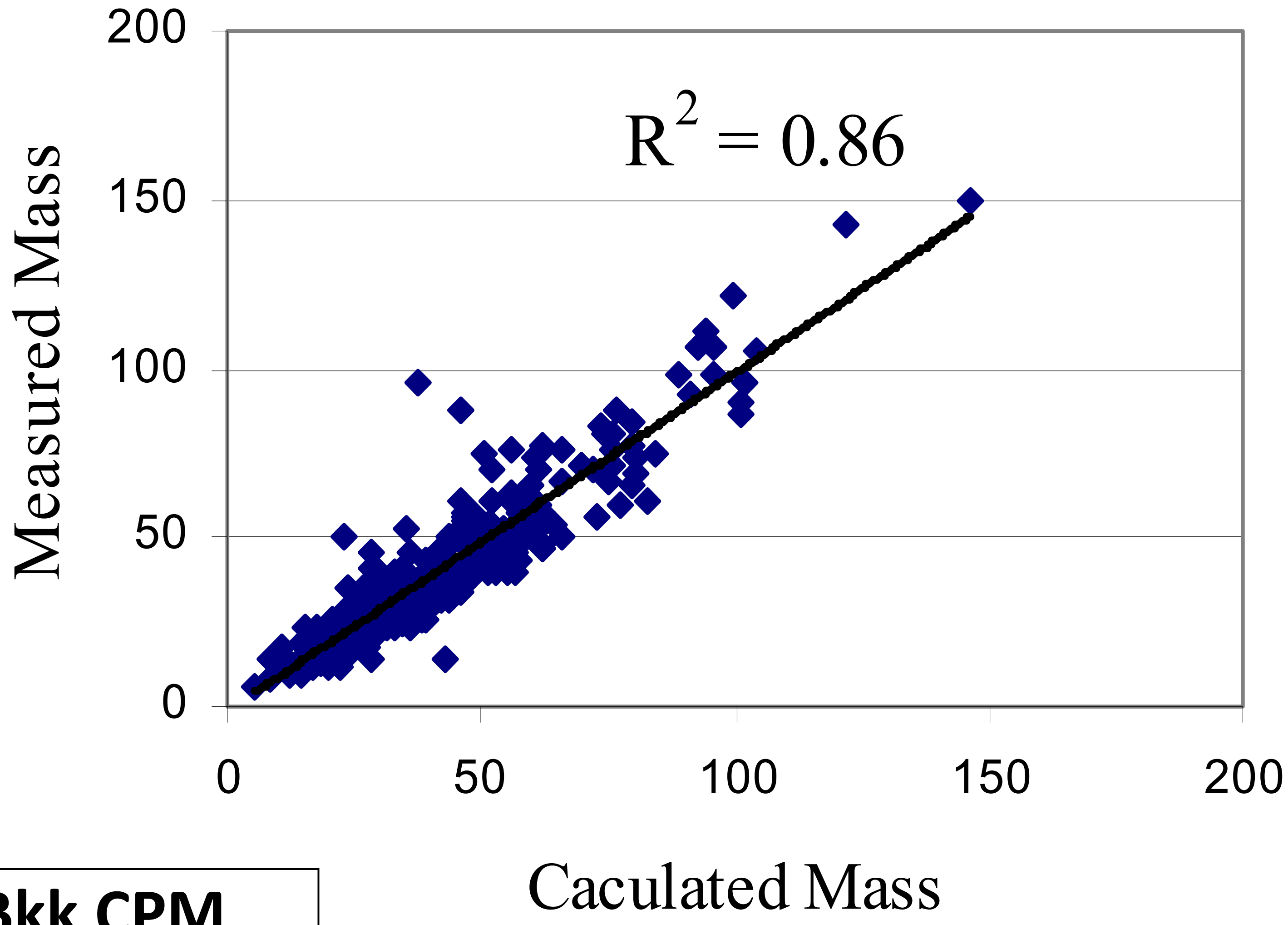
CPM at Bkk



(b)

Model fitting

- Measured values & Model values
- Graphical interpretation
  - Excel



200
150
100
50
0
Measured Mass
$R^2 = 0.86$
0
50
100
150
200
Caculated Mass

#Bkk CPM

In CMB the required data are the source profiles, or fingerprints, of a list of compounds for all potential sources and the corresponding ambient measurements. CMB modelling can be done individually for each ambient measurement (even a single one!) but getting the source profiles for locally relevant sources requires much effort.

In PMF, many ambient measurements are obtained over a long time period. Statistical analysis of the trends in the individual measured components then leads to the identification of “factors" that explain the variation. These factors can then be identified with potential sources. Actual source profile measurements are not required. Instead, collecting sufficiently long time series data on ambient concentrations becomes the critical issue.

CMB – PMF Comparison

Even a single receptor observation can be analyzed by CMB, assuming we have characterized the potential sources.

PMF does not require that we characterize the potential sources but it does require sufficient observations to identify the potential sources.