Comparative Study on the CMB-8 and PMF Models for a Coastal Industrial Area

Shin'ichi Okamoto[*1], Arpa Wangkiat[2], Prapat Pongkiatkul[3], Chaloemchai Nakkhwan[4], Nguyen Thi Kim Oanh[5]

[1]Tokyo University of Information Sciences, Onaridai, Chiba, Japan
[2]Rangsit University, Lakhok, Pathumthani, Thailand
[3]Center of Excellence for Environmental and Hazardous Waste Management,
Department of Environmental Engineering, Faculty of Engineering,
King Mongkut's University of Technology Thonburi,
Thungkru, Bangkok, Thailand
[4]Team Consultant, Bangkok, Thailand
[5]Asian Institute of Technology, Klong Luang, Pathumthani, Thailand
[1]okamoto@rsch.tuis.ac.jp; [2]arpa@rsu.ac.th; [3]prapat.pon@kmutt.ac.th; [4]chaloemchai_n@team.co.th; [5]kimoanh@ait.ac.th

Abstract- **The positive matrix factorization (PMF) and the chemical mass balance (CMB) models were applied to the source apportionment of ambient PM study for a coastal industrial area. The PMF result for coarse fraction was similar to that of the CMB model. As for coarse fraction, correlation coefficients for the calculated contributions by the PMF and CMB models for individual samples were larger than 0.6 for the source categories of sea salt, soil and steel mill. As for the fine fraction, the PMF could also identify the following emission sources: steel mill, soil, Cl and NO_3 rich secondary particle, aged sea salt, SO_4 rich secondary particle, and refuse burning. However the CMB result for fine fraction was not satisfactory due to higher ratio of secondary particles.**

Keywords- ***Particulate Matter; Source Apportionment; Receptor Model; CMB and PMF***

I. INTRODUCTION

Recently air pollution of fine particle became much concerned due to serious health problems. Source contribution to ambient PM in a location is required for development of efficient air quality control strategies. Lack of emission data introduces to high uncertainties of source apportionment results. Receptor modelling has been introduced into the field, which assesses the contributing sources based on the observations at a receptor. The model is classified into 2 main branches, i.e. chemical mass balance (CMB) and multivariate approaches. Each of them has its own advantages which compensate each other. CMB relies on source composition profiles, while multivariate assess the source contribution based on factor analysis without advance information on source composition profiles. The Positive Matrix Factorization (PMF) is one of multivariate receptor model softwares that use non-negativity constraint. The major advantage of PMF model is to utilize a point-by-point least squares minimization scheme [1].

An air quality study has been conducted by the Industrial Pollution Control Association of Japan (IPCAJ) including fine fraction of PM and its composition (i.e. ions and elements) collected at Kashima area, Japan (one of the largest industrial complexes, 80km ENE of Tokyo). The brief summary had already been presented [2]. A previous source apportionment study using an ordinary multivariate model such as Varimax–rotated factor analysis (FA-MR) has already been conducted but it does not be able to work well. The reason may be the negative correlation between the sulfate (SO_4^{2-}) and nitrate (NO_3^-) concentrations for the entire four-season data set. Generally speaking, a factor analysis substantially assumes that all of the data have been extracted from one and only population. However, as for the fine fraction of the Kashima data set, this assumption can not be applied. It might be extracted from two populations, namely summer and another three seasons.

Recently, the Positive Matrix Factorization (PMF) technique was introduced on this multiple-site Kashima data set for source apportionment study. It could successfully estimate the emission source profile and the contribution of different sources to both coarse and fine fractions [3].

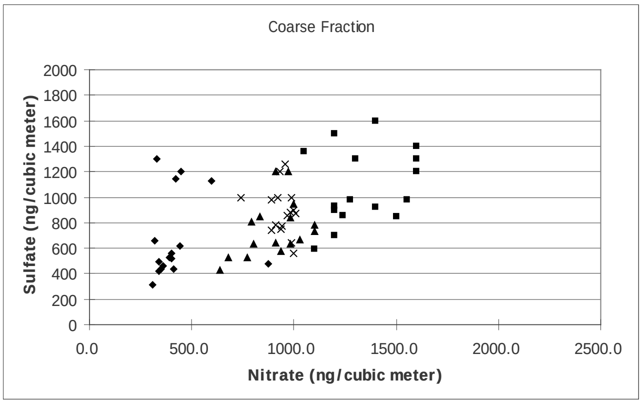
The aim of this study is to examine the performance of the PMF for the Kashima dataset in which some chemical components show a negative correlation, as compare with the source contribution result from the CMB model.

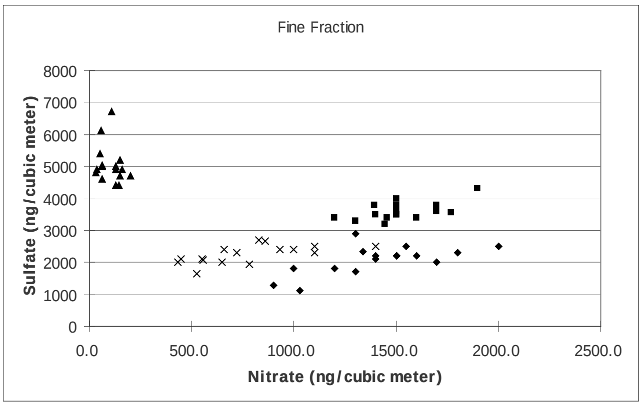
II. DATA DESCRIPTION

The data used in this study were chemical composition data in fine and coarse fraction of PM obtained in the Kashima area. This project was conducted by the Industrial Pollution Control Association of Japan (IPCAJ). The Kashima city contains a number of industries in the industrial complexes. A large steel mill is located in the northern area of this industrial complex, while the southern area consists of an oil refinery, petrochemical industries and a 4400MW oil-fired power plant.

Particulate matter (PM) samples were collected using Anderson size-selective impactors (8 stages with cut-off points, i.e. 11.0, 7.0, 4.7, 3.3, 2.1, 1.1, 0.65, 0.43 μm) at 16 sampling sites. The sampling duration for each site was 14 days for four seasons. PM analysis was divided into 2 groups. The first one is of stages 1-4 (aerometric diameter between 2.1-11 µm), which is considered as coarse fraction, whereas PM in the stages of 5-7 and the backup filter (aerometric diameter under 2.1 µm) was considered as fine fraction. A total of 64 pairs of fine and coarse samples were used for PMF and CMB calculations. More details on monitoring program are described in our previous paper [2].

Particulate matter in this Kashima area contained very high concentrations of Na^+ and Cl^- (sea salt). This is because this area faced to the Pacific Ocean, and the wind frequently came from the ocean (NE wind direction). Compositional concentration of industry-related elements such as Cr, Ni, and V was not so high. Sulfate and nitrate (major component of secondary PM) were the most important members of the fine fraction of PM. In this area, SO_4^{2-} in spring and summer was high, whereas the NO_3^- was much higher in winter. Correlation between the SO_4^{2-} and NO_3^- was quite different in fine and coarse fractions as shown in Fig. 1. Good correlation between the SO_4^{2-} and NO_3^- in the coarse fraction has been observed for the entire four-season dataset, while two groups of data were classified in the fine fraction dataset. The correlations were positive and seemed to be similar for the three seasons (winter, spring, and autumn) except summer [2].





◆Winter ■Spring ▲Summer ×Autumn

Fig. 1 Scatter diagram between SO_4^{2-} and NO_3^- concentrations for fine and coarse fractions

III. CHEMICAL MASS BALANCE MODEL (CMB-8)

The CMB-8 model (U.S. EPA's CMB version 8) uses least squares estimation method to solve chemical mass balance equations [4]. According to the model basic assumptions, the equations will yield a unique solution only when the number of species is equal to or greater than the number of sources. The basic equations can be solved by an effective variance weighted least squares estimation method [5]. The CMB-8.2 software is the latest version of CMB-8 currently in use. It employs an effective variance solution, developed and tested by Watson, et al. [6]. Hereafter this model is called CMB model in this paper.

The source profile used in this calculation was the same as previous IPCAJ study [7]. Eight source profiles were used in the calculation including soil, sea salt, steel mill, refuse incineration (biomass burning), fuel oil burning, gasoline automobile, other industries (metal works, glass works, and so on) and limestone (assumed as paved materials in this study). Totally 13 marker elements were selected for the CMB calculation i.e. Al, Na, Mn, Zn, V, Pb, Sb, Ca, Cr, Sc, Ti, Fe, and Br. Note that the source profiles used in this study were prepared for a total mass. Hence, these data were assumed in the calculation for both of fine and coarse fractional data. The diesel profile was not prepared in this study. This is because a half of data set did not contain carbon analysis. A cement industry appeared in the previous dataset seemed to be inappropriate because there were no significant cement industries and seemed to be very similar to paved materials of the roads. Therefore, the calculated contribution for its dataset was considered as a road dust.

IV. POSITIVE MATRIX FACTORIZATION (PMF) MODEL

Positive Matrix Factorization (PMF) developed by Paatero and Tapper [8] is one of multivariate models that assess the contribution using a factor analysis. Non-negativity constraint for both loadings and scores is a key concept in the PMF. In addition, the PMF utilizes a point-by-point least-squares minimization scheme in the calculation that is a major different from other multivariate approaches. PMF2 [9] was used in this study, and it is called as the PMF in this paper.

The fine and coarse PM data from 16 sampling sites were used in the calculation. By a preliminary screening, 12 elements by INAA were deleted, because most data were under detection limits. Elements and ions used in this analysis are shown in Table 1. PM mass and their uncertainty (10% of mass were assigned) were included in the input data in order to constrain the model to mass. The missing values in this study were replaced with the mean of measured concentration of those species, while 4 times the average concentration were assigned for uncertainty as an estimate of missing values, according to the Polissar et al. [10]. More detail procedures were shown in Pongkiatkul et al. [3].

V. RESULTS AND DISCUSSIONS

A. PMF Results

The PMF could identify six source categories for the coarse fraction; soil, sea-salt, industry (oil burning), steel mill, biomass or refuse burning and secondary particle (including aged sea-salt, $NaNO_3$). As for the fine fraction, the PMF could also identify the following emission sources; steel mill, soil, Cl^- and NO_3^- rich secondary particles, aged sea salt, SO_4^{2-} rich secondary particles, and refuse burning.

TABLE I ELEMENTS AND IONS USED IN THIS ANALYSIS

Coarse Fraction
Al, Br, Cl^-, Cr, Fe, K, Mn, Na, NO_3^-, Sc, Si, SO_4^{2-}, Ti, V, Zn

Fine Fraction
Al, As, Br, Cl-, Cr, Fe, K, Mn, Na NO_3^-, Sb, Se, Si, SO_4^{2-}, Ti, V, Zn

The overall average (64 data = 16 sites x 4 seasons) contributions for each source category are shown in Table 2. The Cl^- and NO_3^- rich secondary particle in winter season was about 25% and the highest in fine fraction, but its contribution in summer season was very low (less than 1%) as shown in Fig. 2. Four-season average for this secondary particle was about 20% as shown in Fig. 2. The SO_4^{2-} rich secondary particle in the fine fraction was highest in summer and its contribution was about 40%. The estimated source profile for this SO_4^{2-} rich secondary particle contained higher vanadium (V) ratio, and it seemed to come from oil burning.

TABLE II COMPARISON BETWEEN THE CMB AND PMF RESULTS

Coarse Fraction

CMB	%	PMF	%
Soil	57	Soil	21
Sea Salt	14	Sea Salt	18
Steel Mill	2	Steel Mill	11
Refuse Burning	2	Refuse Burning	9
Fuel Burning	1	Industry	16
Automobile	3	Secondary Particle	25
Road Dust	10		
Unknown	11		

Fine Fraction

CMB	%	PMF	%
Road Dust	27	Soil	19
Steel Mill	1	Steel Mill	6
Refuse Burning	14	Refuse Burning	20
Fuel Burning	6	SO_4^{2-} Rich Secondary	25
Automobile	5	Cl^- Rich Secondary	20
Sea Salt	0.1	Aged Sea Salt	10
Unknown	47		

B. CMB Results

Overall average contributions for designated emission source categories by CMB calculation are also shown in Table 2.

Most dominant members obtained by the CMB calculations were soil and road dust (source profile derived by cement) for both of fine and coarse fractions. The contributions by CMB model did not contain the secondary particles, therefore unknown parts were large, especially as for the fine fraction, the unknown parts were nearly a half.

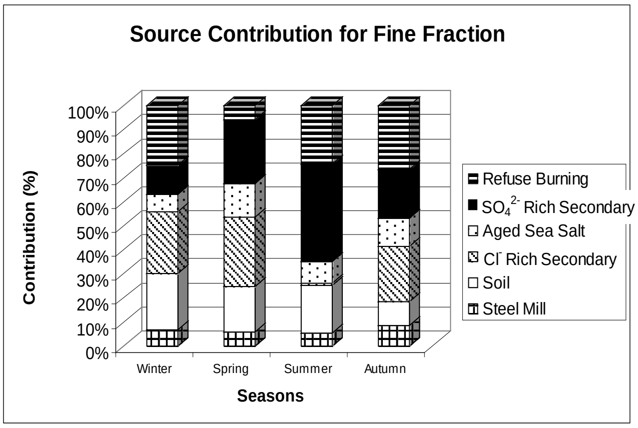


Fig. 2 Seasonal variations of fine fraction 16-sites average source contributions by the PMF model

C. Comparison by Individual Data

Comparison by individual data points obtained by the PMF and CMB models was conducted. Since the CMB model can not identify the secondary particles, a limited number of source categories were compared. Scatter diagrams for some selected sources are shown in Fig. 3, while Table 3 shows the correlation coefficients between the results from CMB and PMF models. As for the coarse fraction, the estimated contribution for sea salt was almost comparable with high correlation ($R^2 = 0.71$).

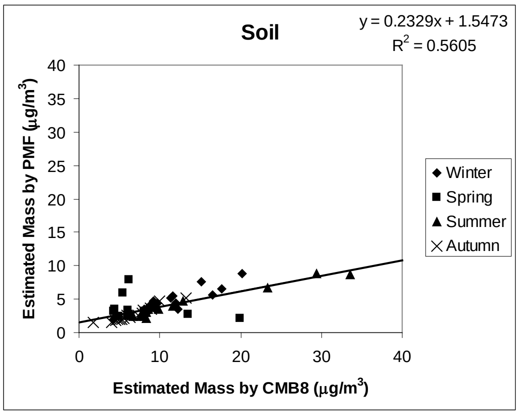
TABLE III CORRELATION COEFFICIENTS FOR THE RELATED SOURCE CATEGORIES PRODEUCED BY CMB AND PMF

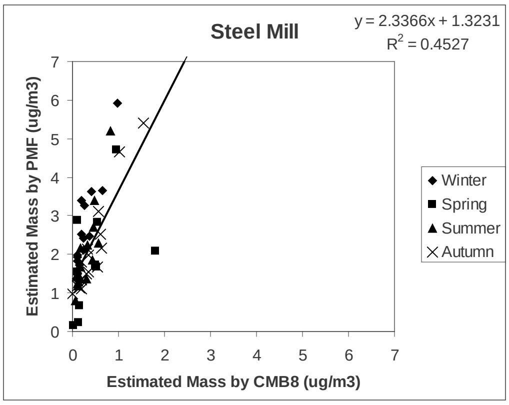
	Coarse Fraction	Fine Fraction
Soil (and road dust)	0.749	0.835
Steel mill	0.673	0.640
Refuse burning	0.486	0.564
Sea salt	0.844	-

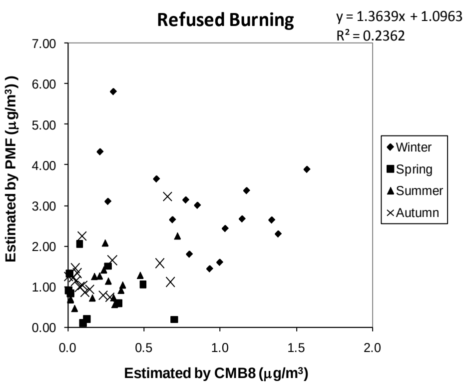
The estimate contributing mass-produced by the both models for soil and steel mill also showed high correlation (R^2=0.56 and 0.45, respectively). However, the CMB produced higher soil contribution (about 4 times) than the one estimated by the PMF, whereas the PMF produced almost 2.5 times higher mass contribution for steel mill (Fig. 3). This is because of the difference in steel mill profiles. PMF produced high Ti fraction in the profile, whereas Fe was the major element in the profiles used in the CMB calculation. Note that the source profile used in the CMB calculation was derived by some limited source data in particulate facilities such as a coke oven, a smelter, and a blast furnace, therefore, its data may not represent the entire steel mill emissions. Since the CMB relies mostly on the input source profiles, different contribution for both sources might be created. However, they were together correlated. Similar trends were also observed in the fine fraction for soil and steel mill (Fig. 3).

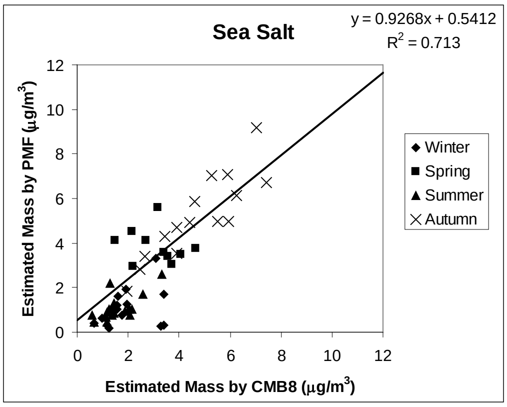
The tendency of scatter diagram for refused burning was similar for coarse and fine fractions. They exhibited quite week correlation (R^2=0.24 and 0.32, respectively). Note that the Kashima area is surrounded by rice paddy field and forest, open burning including biomass might be significant.

Coarse fraction

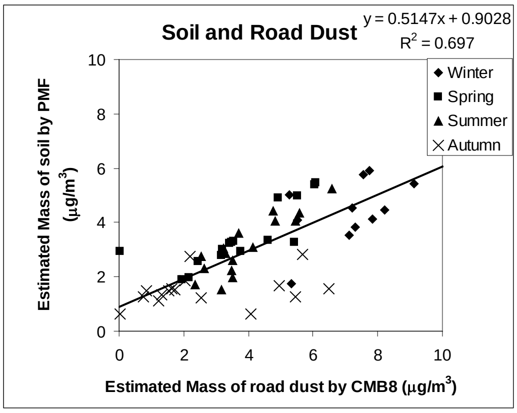


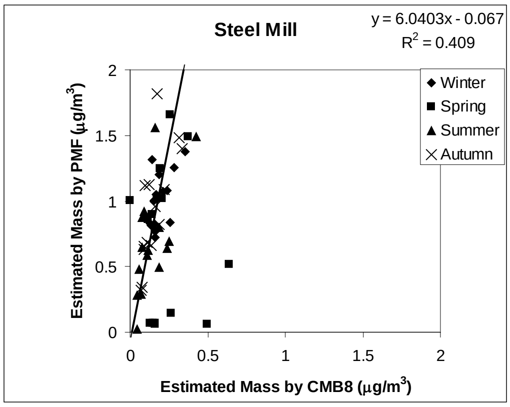


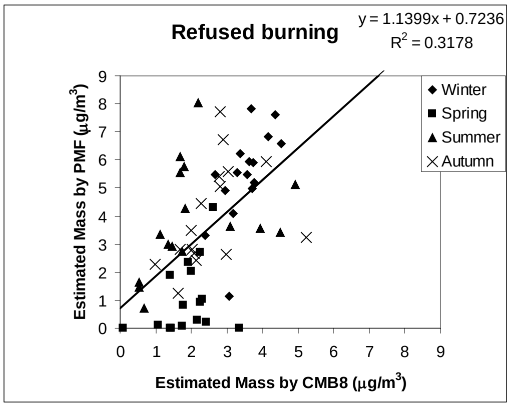




Fine fraction







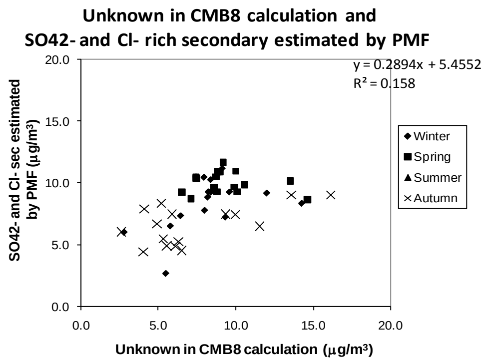


Fig. 3 Comparison between the CMB and PMF calculations

Therefore, the source data for straw burning were selected as the profile for refuse burning in this CMB calculation. The scatter diagram for this refuse burning suggested that the selection of the source profile seemed to be correct.

Since the SO_4^{2-} rich and Cl^- rich secondary particles produced in the fine fraction cannot be estimated by the CMB model due to no source profile available. Comparison between a combination of SO_4^{2-} rich and Cl^- rich secondary particles produced by the PMF and unknown particles from the CMB was also conducted. The scatter diagram showed that there was a weak correlation (Fig. 3). Therefore almost sulfate might come from outside or more upstream distant areas.

A part of this dataset did not contain carbon analytical data, and contributions for diesel exhaust could not be estimated by the CMB calculations. The Kashima area had three major towns, not the cities at the moment of this study, and its population was less than several tens thousands each. Therefore, traffic-related pollution was not so severe at that moment. A large amount of unknown parts by CMB results might contain the diesel exhaust. However, the PMF result exhibited nearly zero contribution for the diesel exhaust. The result suggested that important marker elements should be included even in the PMF calculations.

Comparison between source profiles obtained from the PMF and the one used in the CMB model was also performed (Fig. 4). However, only 2 sources show almost similar profiles. Major elements in soil profile are Al and Na, which are found in both profiles (PMF and CMB). However, the soil profile obtained from PMF for fine fraction was lack of Na, and contained high concentration of Zn and Fe. Ti and Fe were also observed in the soil profile collected in the IPCAJ study, whereas the profile obtained by the PMF model does not contain those elements for coarse and fine fractions. This is because the profiles collected under the IPCAJ were estimated from the PM total mass. High Na and Br were observed in the seasalt profile for both CMB and PMF. Note that the PMF did not produce soil profile for the fine fraction.

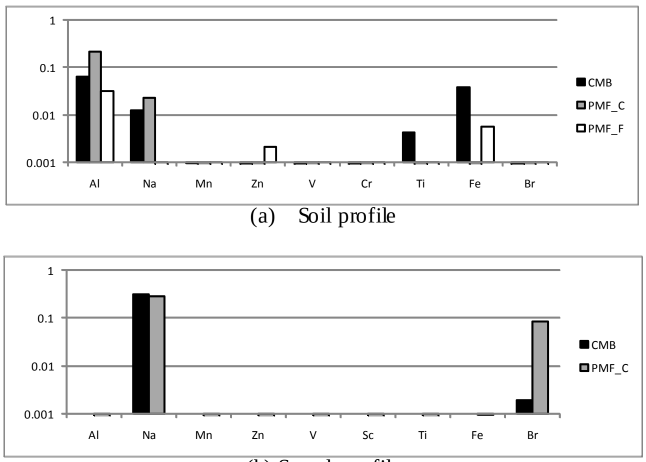


Fig. 4 Comparison between source profiles used in CMB model and source profiles produced by PMF model

VI. CONCLUSIONS

The positive matrix factorization (PMF) and the chemical mass balance (CMB version8) models were applied to the source apportionment of ambient PM study for Kashima area. As for coarse fraction, high correlations between the PMF and CMB calculations were obtained.

As for the fine fraction, the PMF could identify the six emission sources: steel mill, soil, Cl^- and NO_3^- rich secondary particle, aged sea salt, SO_4^{2-} rich secondary particle, and refuse burning. The CMB result for fine fraction was not satisfactory due to higher ratio of secondary particles. However, some limited emission sources both results showed higher correlations.

ACKNOWLEDGMENT

A part of this paper had already presented at the 14th IUAPPA World Clean Air Congress 2007.

The data set used here is the outcome of the study conducted by the Industrial Pollution Control Association of Japan (IPCAJ). The name of this association was recently changed to the Japan Environmental Management Association for Industry (JEMAI). We are grateful to the late Mr. K. Kobayashi and staffs of this association for helpful discussions of its use.

REFERENCES

[1] Paatero P. (2007) *User's guide for positive matrix factorization programs PMF2 and PMF3*. Part 1: Tutorial, Part 2: Reference, University of Helsinki, Finland.

[2] Okamoto S., Kobayashi K. (1991) "Characterization of aerosols in the Kashima area–a summary of IPCAJ Kashima SPM study," *Atmospheric Environment*, vol. 25B, pp. 133-142.

[3] Pongkiatkul P., Kim Oanh N.T., Wangkiat A., Okamoto S. (2010) "Further analysis and PMF application to the chemical composition data of IPCAJ's Kashima SPM study," *Int. J. Environment and Pollution*, Vol. 40, No. 4, pp. 337–350.

[4] Coulter C.T. (2004) *EPA-CMB8.2 Users Manual*. EPA-452/R-04-001, U.S.EPA office of air quality planning and standards, emission, monitoring and analysis division, air quality model group.

[5] Watson J.G., Cooper J.A., Huntziker J.J. (1984). "The effective variance weighing for least squares calculations applied to the mass balance receptor model," *Atmospheric Environment*, vol. 18, pp. 1347-1355.

[6] Watson J.G., Robinson N.F., Lewis C., Coulter T. (1997) *Chemical mass balance receptor model version 8 (CMB8) user's manual*, Desert Research Institute, Document, No. 1808. 1D1.

[7] Okamoto S., Nakajima M., Kainuma Y., Shiozawa K. (1987) A study on the estimation of source contribution for the suspended particulate matter – Application of the factor analysis – multiple regression analysis model, *J. Japanese Society for Quality Control*, 17: 294-300, (In Japanese).

[8] Paatero P., Tapper U. (1992) "Positive Matrix Factorization: a non-negative factor model with optimal utilization of error estimates of data values," *Fourth International Conference on Statistical Methods for Environmental Sciences, Environmetrics*, August 17-21, 1992, Espoo, Finland.

[9] Paatero P. (1998) Use's Guide for Positive Matrix Factorization Programs PMF2 and PMF3, Part 1 and Part 2, Available at: ftp://rock.helsinki.fi/pub/misc/pmf/.

[10] Polissar A.V., Hopke P.K., Paatero P., Malm W.C., Sisler J.F. (1998) "Atmospheric aerosol over Alaska 2, Elemental composition and sources, *Journal of geophysical research,* vol. 103, No. D15, pp. 19045-19057.

Shin'ichi Okamoto is a Professor of Department of Environmental Information, Tokyo University of Information Sciences. He received Doctor of Engineering from the Graduate School of Waseda University, Japan, in 1977. His major field is Environmental Management. His is a chair of environmental assessment committee of Chiba prefecture, Japan. He is also a visiting professor of School of Environment Resource and Development, Asian Institute of Technology

Arpa Wangkiat is a Lecturer of Department of Environmental Engineering, Rangsit University, Thailand. She received PhD from the Joint Graduate School of Energy and Environment at King Mongkut's University of Technology, Thonburi, Thailand, in 2003.

Prapat Pongkiatkul is a Lecturer at the Center of Excellence for Environmental and Hazardous Waste Management, Department of Environmental Engineering, Faculty of Engineer, King Mongkut's University of Technology Thonburi, Bangkok, Thailand. He completed his BE in Environmental Engineering from King Mongkut's University of Technology, Thailand, in 1999, and his ME and PhD from Environmental Engineering and Management programme, Asian Institute of Technology, Thailand, in 2001 and 2006.

Chaloemchai Nakkhwan is a staff of Team Consultant in Bangkok. He graduated from Department of Environmental Engineering, Rangsit University.

Nguyen Thi Kim Oanh is a Professor of School of Environment Resource and Development (SERD), Asian Institute of Technology, Thailand. Her research areas include air pollution engineering and management, and air pollution meteorology with the focus on particulate matters, ground-level ozone and trans-boundary air pollution.