

**Study on**

**Integrated Data Analysis and Characterization  
of Particulate Matter in Hong Kong**

**Final Report**

**February 2012**

Submitted to the  
Environmental Protection Department, HKSAR  
for Provision of Service  
under  
Tender Reference AS09-056

by

**Alexis Lau and Zibing Yuan**  
Institute for the Environment  
Hong Kong University of Science & Technology

**Jay Turner and Varun Yadav**  
Department of Energy, Environmental and Chemical Engineering  
Washington University in St. Louis

## Table of Contents

Acronyms.....	5
<b>Executive Summary .....</b>	<b>7</b>
<b>1. Introduction .....</b>	<b>12</b>
<b>1.1 Background.....</b>	<b>12</b>
<b>1.2 Objective of this Project.....</b>	<b>12</b>
<b>2. Data Analysis Methodology.....</b>	<b>13</b>
<b>2.1 Receptor Modeling .....</b>	<b>13</b>
2.1.1 Positive Matrix Factorization.....	13
2.1.2 Principal Component Analysis / Absolute Principle Component Scores .....	14
2.1.3 Unmix .....	15
2.1.4 Chemical Mass Balance.....	15
<b>2.2 Data Quality Assessment Using Collocated Data .....</b>	<b>16</b>
2.2.1 Motivation.....	16
2.2.2 Methodology and Results for PM <sub>10</sub> Analysis Data .....	17
2.2.3 Methodology and Results for PM <sub>2.5</sub> Data .....	20
2.2.4 Recommendations.....	21
<b>2.3 Air Mass Transport Pattern Analysis .....</b>	<b>21</b>
<b>3. Source Apportionment of PM<sub>10</sub> in Hong Kong .....</b>	<b>25</b>
<b>3.1 Data Collection, Analysis and Pre-processing .....</b>	<b>25</b>
<b>3.2 Source Apportionment by PMF .....</b>	<b>25</b>
3.2.1 Source Contributions and Their Seasonal and Spatial Variations .....	29
3.2.2 Annual Trends of Source Contributions over the Eleven Years .....	30
3.2.3 General Stations vs. Roadside Station .....	32
3.2.4 Comparison with Results in Previous Studies .....	33
3.2.5 Summary.....	36
<b>3.3 Source Apportionment by APCA and Unmix .....</b>	<b>38</b>
3.3.1 Implementation of APCA .....	38
3.3.2 Implementation of Unmix.....	39
3.3.3 Factor Source Profiles and Explained Mass Profiles .....	40
3.3.4 Comparison of APCA and Unmix Source Contribution Estimates.....	42
3.3.5 Key Trends for the Source Contributions Estimates .....	42
3.3.6 Summary and Recommendations .....	44
<b>3.4 Source Apportionment by CMB .....</b>	<b>45</b>
3.4.1 Source Profiles – Selection and Modification.....	45
3.4.2 CMB Modeling and Diagnostics .....	47
3.4.3 CMB Results for the Base Case Modeling .....	49
3.4.4 Summary and Recommendations .....	51

<b>3.5</b>	<b>Comparison of Source Contribution Estimates by PMF, APCA, Unmix and CMB</b>	<b>51</b>
3.5.1	PMF, APCA, and Unmix	51
3.5.2	PMF, APCA, Unmix, and CMB	52
<b>4.</b>	<b>Source apportionment of PM<sub>2.5</sub> in Hong Kong</b>	<b>53</b>
4.1	PM <sub>2.5</sub> versus PM <sub>10</sub> Species Concentrations	53
4.2	PMF Modeling Strategy	54
4.3	Characteristics of the PMF-Resolved Factors	55
4.4	APCA Modeling	57
4.5	Unmix Modeling	58
4.6	Comparison of Source Contribution Estimates for PM <sub>2.5</sub> and PM <sub>10</sub>	59
4.7	Summary and Recommendations	60
<b>5.</b>	<b>Conceptual Model for Ambient Particulate Matter over Hong Kong</b>	<b>62</b>
5.1	Introduction	62
5.2	Interaction of Meteorology and Air Quality in Hong Kong and the Pearl River Delta	63
5.3	Air Quality Conditions – Local and Non-local Impacts	65
5.4	Local and Non-local Pollutant Sources	69
5.5	Relationship between Pollutant Level, Source Contributions and Air Mass Transport Patterns	71
5.5.1	Approach	71
5.5.2	Results for PMF-Modeled PM <sub>10</sub> Source Contribution Estimates	73
5.5.3	Results for PM <sub>10</sub> TEOM Mass Concentration Data	75
5.5.4	Air Mass Transport Pattern Analysis Summary	76
5.6	A Preliminary Separation of PRD and Non-PRD Source Contributions using Air Mass Back Trajectories	77
5.6.1	Data and Methodology	77
5.6.2	Results	79
5.6.2.1	<i>Mapping of Hourly PM<sub>10</sub> Data at TW onto Classified Scales of Spatial Transport</i>	79
5.6.2.2	<i>Mapping of Hourly PM<sub>10</sub> Data at CW onto Classified Scales of Spatial Transport – High Mass Concentration Hours</i>	80
5.6.2.3	<i>Comparison with 3-D Air Quality Modeling</i>	82
5.6.3	Discussion and Recommendations	82
5.7	A Compact Summary of the Conceptual Model for Ambient Particulate Matter over Hong Kong	84
<b>6.</b>	<b>Conclusions and Recommendations</b>	<b>85</b>

<b>6.1</b>	<b>Key Findings</b> .....	<b>85</b>
<b>6.2</b>	<b>Recommendations</b> .....	<b>87</b>
<b>6.3</b>	<b>The Path Forward</b> .....	<b>90</b>
<b>7.</b>	<b>References</b> .....	<b>94</b>
	<b>APPENDIX A – Samples Eliminated from the Modeled Data Set (Cd &gt; 0.01</b> <b>µg/m<sup>3</sup>)</b> .....	<b>97</b>
	<b>APPENDIX B: Steps of Performing PMF, APCA and Unmix</b> .....	<b>98</b>
	<b>APPENDIX C: PM<sub>10</sub> TEOM Mass Partitioning into Base and Excess</b> <b>Contributions</b> .....	<b>101</b>

## Acronyms

APCA	Absolute Principal Component Analysis (PCA/APCS)
APCS	Absolute Principal Component Scores
AQMS	Air Quality Monitoring Station
AQO	Air Quality Objectives
CMB	Chemical Mass Balance
CW	Central/Western
EC	Elemental Carbon
HK	Hong Kong
HKEPD	Hong Kong Environmental Protection Department
HKIA	Hong Kong International Airport
HKO	Hong Kong Observatory
HT	Hok Tsui
HYSPLIT	Hybrid Single Particle Lagrangian Integrated Trajectory Model
IC	Ion Chromatography
KT	Kwun Tong
LOQ	Limit of Quantification
MDL	Method Detection Limit
MK	Mong Kok
OC	Organic Carbon
OM	Organic Matter (assumed OC×1.6)
PATH	Patterns in Atmospheric Transport History
PCA	Principal Component Analysis
PM	Particulate Matter
PMF	Positive Matrix Factorization
PRD	Pearl River Delta
PSCF	Potential Source Contribution Function
QTBA	Quantitative Transport Bias Analysis
SCE	Source Contribution Estimates
SOA	Secondary Organic Aerosol
SSP	Sham Shui Po
TC	Tung Chung
TEOM	Tapered Element Oscillating Microbalances
TOR	Thermal-optical Reflectance
TOT	Thermal-optical Transmittance
TS	Transitional Season

---

TW	Tsuen Wan
WD	Wind Direction
WHO	World Health Organization
XRF	X-ray Fluorescence
YL	Yuen Long

## Executive Summary

Effective air quality management and planning takes advantage of observational air pollutant data to not only assess air quality conditions but to also evaluate the impact of air quality control measures. For more than a decade the Hong Kong Environmental Protection Department (EPD) has operated a network of ambient particulate matter (PM) real-time monitors and time-integrated samplers that provide rich data sets to assess spatial and temporal trends in PM mass and its chemical constituents. The core objective of this project was to utilize these data to determine the emission source categories that impact PM levels in Hong Kong and how these patterns have changed over the previous decade. Emphasis was placed on a *weight-of-evidence approach* using independent data sets and analyses to assess the reliability of key findings. This work has culminated in the development of a conceptual model for ambient PM over Hong Kong with the aim of separating sources from different spatial scales under different administrations. This conceptual model is grounded in the observational data and future work should refine the model by incorporating information from emissions inventories and chemical transport modeling. Analyses conducted for this project included data collected between 1998 and 2008. The conceptual model should be considered a “living document” to be periodically updated as more data are collected.

### Data and Analyses

This project was conducted using routine monitoring data collected by EPD, including: 24-hour integrated PM<sub>10</sub> samples collected every 6<sup>th</sup> day at seven sites from 1998 to 2008 and analyzed for PM mass and species; 24-hour integrated PM<sub>2.5</sub> samples that were collected every 6<sup>th</sup> day at three-to-four sites for three one-year periods and were analyzed for PM mass and species; and hourly PM<sub>10</sub> mass collected by Tapered Element Oscillating Microbalance (TEOM) monitors at 14 sites from 1998-2008. Air mass back trajectories were obtained from [www.sharedair.org](http://www.sharedair.org) which includes an archive of air mass back trajectories for several cities worldwide as compiled by The University of Michigan for the U.S. Environmental Protection Agency. The Hong Kong data set includes four trajectories per day from January 1, 2000 through December 21, 2010. Hourly surface winds data from Hong Kong International Airport and the Hong Kong Observatory was also used.

Source apportionment modeling was performed on the PM<sub>10</sub> and PM<sub>2.5</sub> speciation data sets using Positive Matrix Factorization (PMF), Unmix, and Principle Components Analysis with Absolute Principle Components Scores (PCA-APCS, hereafter termed APCA). Chemical Mass Balance (CMB) modeling was also performed on the PM<sub>10</sub> speciation data. Each of these models has strengths and limitations. Therefore, modeling was performed using each of these models to: evaluate the consistency of results across the models; and recommend a preferred model for future analyses on the Hong Kong data sets which are routinely updated to include the most-recently collected data.

A comprehensive analysis of the speciated PM data sets quality was performed using collocated sampler data. The error structures (uncertainty as a function of concentration) from this analysis are preferred over the uncertainty estimates provided by the respective analytical laboratories for the PM<sub>10</sub> data and especially for the PM<sub>2.5</sub> data. These error structures were used for the PMF and CMB modeling, and in general should be considered when performing any analyses on the PM speciation data.

PM<sub>10</sub> modeling by PMF yielded a nine factor solution with these factors assigned to the following sources (in order of decreasing contributions at the six general stations): secondary sulfate (22%), vehicle exhaust (15%), secondary nitrate (15%), biomass burning admixed with coal combustion (13%), aged sea salt (13%), crustal soil / resuspended dust (12%), trace metals processing such as zinc smelting (5%), fresh salt (3%), and residual oil combustion from marine vessels (0.4%). The PMF modeling was performed to apportion the gravimetric mass measured on the quartz filter. While there are concerns about artifacts potentially associated with the quartz filter gravimetric mass measurement, we have confirmed its reliability as a total variable for the PMF modeling including good agreement with the PM<sub>10</sub> TEOM data.

PM<sub>10</sub> modeling by APCA and Unmix could not resolve a distinct nitrate factor and the measured nitrate was distributed over several of the factors. This issue might arise from the artifacts inherent to nitrate measurement using Hi-Vol samplers and quartz fiber filters; the artifact is influenced by gaseous nitric acid mixing ratios, temperature and humidity and is likely reproducible on a day-specific basis leading to high collocated precision. PMF picks up on the combination of high precision and a lack of covariance with other species to assign nitrate as a distinct factor. Also, APCA and Unmix do not force the species loadings and source contribution estimates to be positive and indeed significant negative loadings onto some factors were observed for some species. Primarily for these reasons, *PMF is deemed the most suitable model for the Hong Kong PM<sub>10</sub> data set.*

PM<sub>2.5</sub> modeling by PMF yielded an eight factor solution with the same source category assignments as the PM<sub>10</sub> modeling with fresh and aged seas salt combined into a single factor. PM<sub>2.5</sub> and PM<sub>10</sub> source contribution estimates showed excellent quantitative agreement for secondary sulfate and vehicle exhaust which is consistent with PM from these sources being in the fine PM size range. Soil/dust source contributions estimates were much higher for PM<sub>10</sub> compared to PM<sub>2.5</sub> which is consistent with PM from soil/dust being primarily in the coarse PM size range. Residual oil source contribution estimates were significantly higher for PM<sub>2.5</sub> than for PM<sub>10</sub>; this is physically unrealistic and suggests considerable uncertainty in the residual oil combustion source contribution estimates. Factors could not be adequately resolved for the APCA modeling of the PM<sub>2.5</sub> data set. Unmix modeling was generally consistent with PMF results. Unmix does not appear to have significant advantages over PMF for the PM<sub>2.5</sub> modeling. Therefore, to provide consistent with the PM<sub>10</sub> modeling, *PMF is deemed the most suitable model for the Hong Kong PM<sub>2.5</sub> data set.*

CMB modeling of the Hong Kong data sets suffer from a lack of representative source profiles. Nonetheless, modeling was performed on the PM<sub>10</sub> data and generally consistent results were obtained compared to the other methods although to conduct the comparisons it was necessary to exclude 17% of all samples which failed the CMB modeling mass closure criterion. These rejected samples tended to occur on low mass concentration days and therefore skew the study-average source contribution estimates. *The utility of CMB modeling on the Hong Kong data could be significantly increased by obtaining representative emission source profiles.*

### Key Findings

The PMF-modeled relative contributions from the various emission source categories were reported above. Temporal trends over the eleven year period 1998-2008 were examined for all source categories. The most remarkable temporal trends were a *54% reduction in motor vehicle contributions (presumed to be dominated by local sources) and a 87% increase in secondary sulfate contributions (presumed to be from non-local sources)*. Similar patterns were observed in the PM<sub>2.5</sub> modeling results. A weight-of-evidence approach was taken to evaluate the robustness of these key findings. In this study, “local sources” are defined for those factors exhibiting spatiotemporal characteristics consistent with what would be expected for emissions from within the Hong Kong territory, while “non-local sources” refers to those factors with spatiotemporal characteristics consistent with to transport the HKSAR from more-distant sources. In the “non-local” source sector, we also classified stratified contributions into “PRD” and “non-PRD” sources according to the presumed geographic locations of the emissions based on a preliminary analysis using the concentration and back trajectory data.

Annual-average vehicle exhaust source contribution estimates were strongly correlated with the Hong Kong annual PM emission inventory from road transport. This trend is very compelling and demonstrates that *vehicle emission control programs implemented in Hong Kong have resulted in significantly lower vehicle exhausted-related PM ambient burdens* at not only the roadside stations but also the general stations. The temporal trend in PMF-modeled vehicle exhaust was further examined by clustering air mass back trajectories into five categories representing synoptic-scale air mass transport nominally from the east coast of mainland China (fast and slow moving air masses were separately resolved), from the south/southwest (corresponding to monsoon conditions), from the east, and also stagnant air masses. PMF-modeled vehicle exhaust source contribution estimates were stratified according to these air mass patterns and the annual average contribution by air mass pattern was regressed on time. The rates of change in the motor vehicle contribution were statistically significant (95% confidence level) for all five air mass transport patterns. However, these rates of changes were statistically indistinguishable across the five air mass transport patterns as would be expected for a local source. In contrast, the increase in PMF-modeled secondary sulfate contributions was observed only for the two air mass transport patterns nominally from the east coast of mainland China with statistically insignificant changes when the air masses were from the south/southwest,

east, or are stagnant. This behavior is consistent with increased transport to Hong Kong from locations to the north.

Given the important implications of these findings, complementary analyses were conducted on daily average  $PM_{10}$  mass from the network of TEOM monitors reporting hourly  $PM_{10}$  mass data. For each site,  $PM_{10}$  mass was stratified by air mass transport pattern and the annual median concentrations for each air mass transport pattern were regressed on time. While these  $PM_{10}$  mass data cannot be used to *a priori* distinguish contributions from different emission source categories, the data set has high daily completeness which leads to robust statistics even after stratifying the data according to air mass transport patterns and year. The regression slopes, which represent the average rate of change in  $PM_{10}$  mass concentration over time, were statistically indistinguishable from zero for most sites and air mass transport pattern. This is consistent with the offsetting effects of reduced vehicle exhaust contributions and increased transport contributions. However, some statistically significant temporal changes were observed. Sites outside the Hong Kong core urban area (Tap Mun, Tung Chung, and Yuen Long) experienced statistically significant increases in  $PM_{10}$  over time but only for air mass transport patterns from the east coast of China. Conversely, the three roadside sites (Causeway Bay, Central, and Mong Kok) and the strongly vehicle influenced general stations of Eastern, Kwun Tong, Sham Shui Po) experienced statistically significant decreases in  $PM_{10}$  over time but only for air mass transport patterns from the south/southwest, east, or stagnant conditions. This is consistent with the increase in transported PM offsetting the vehicle emission reductions when winds are from the east coast of China.

Concerning the long terms trends, *the weight-of-evidence using independent data sets and analyses all lead to the same conclusion that the lack of significant change in Hong Kong ambient  $PM_{10}$  levels over the previous decade arises from a significant increase in non-local sources offsetting the significant reductions in vehicle exhaust PM from Hong Kong emission control measures.*

Given the importance of transport, a preliminary analysis was conducted to estimate relative contributions from local sources (i.e. within the Hong Kong SAR), PRD sources (i.e. the greater Pearl River Delta (PRD) region excluding the Hong Kong SAR), and non-PRD sources including but not limited to the East Asian Continent. Air mass back trajectories were classified into these three spatial categories based on their presence and residence time in these geographic zones.  $PM_{10}$  and  $PM_{2.5}$  mass data from 2006 was used for this study, with most analyses performed on the Tsuen Wan site data. During the summertime, which is dominated by monsoon conditions,  $PM_{10}$  contributions are greatest from non-PRD sources (27~56%), followed by local (35%) and PRD sources (8~38%). For the scenario of high PM hours in the summertime, however, PRD contributions (16~81%) increase with a corresponding drop in local contributions (10%). In the wintertime, contributions are greatest from transport beyond the PRD (45~67%), followed by local contributions (33~36%) and finally PRD contributions (0~19%). On high PM days in the wintertime, however,

the contributions are quite different with PRD (0~37%) and non-PRD (33~76%) transport dominating in some cases with slightly reduced local sources (24~29%). Similar percentages were derived for PM<sub>2.5</sub> contributions, suggesting PM<sub>10</sub> the core elements of the PM<sub>10</sub> conceptual model for Hong Kong also apply to PM<sub>2.5</sub>. Source apportionment conducted within the chemical transport model CAMx for November 2006 yielded splits between local/PRD/non-PRD contributions that were generally consistent with the trajectory-based analysis. While these analyses are preliminary and subject to considerable uncertainty, they do suggest that all three spatial scales of emissions – local, PRD, and non-PRD – have significant impacts on Hong Kong PM air quality. Differential impacts occur during the summertime and wintertime, and on high PM days compared to average PM days. *While Hong Kong air quality would certainly benefit from additional controls within the Hong Kong SAR, emissions reductions from both the PRD and more distant sources are critical to improving air quality in Hong Kong with controls within the PRD particularly important for reducing PM levels on high PM days.*

# 1. Introduction

## 1.1 Background

Particulate matter (PM) is one of the major air pollutants in Hong Kong (HK) which impairs visibility, imposes health risk and leads to regional climate change. Spatiotemporal investigation of PM<sub>10</sub> (PM with aerodynamic diameters  $\leq 10\mu\text{m}$ ) source apportionment during 1998 - 2005 by Huang *et al.* (2009) has highlighted the importance of anthropogenic primary sources from the Pearl River Delta (PRD), especially the more urbanized areas to the northwest of HK, affecting local air quality. Although similar patterns were observed in 12-month PM<sub>2.5</sub> (PM with aerodynamic diameters  $\leq 2.5\mu\text{m}$ ) studies in 2000/2001 and 2004/2005 (e.g. Wang 2003, Louie *et al.* 2005, Guo *et al.* 2009), differences in the receptor modeling techniques used across these studies prevents an internally consistent comprehensive synthesis of results.

To date, the HK Environmental Protection Department (HKEPD) possesses validated PM<sub>10</sub> speciation datasets from 1998 to 2010 and three 12-month PM<sub>2.5</sub> speciation datasets of 2000/2001, 2004/2005 and 2008/2009. Updating the previous PM source apportionment results with the latest speciation data and performing integrated receptor modeling techniques, with interpretation aided by detailed meteorological information, can increase confidence in the source apportionment results including but not limited to the identifying changes in PM trends. As such, a robust "conceptual model" (i.e. Qualitative description of physical processes and chemical characteristics governing PM formation; Turner and Garlock 2007) can be developed for more effective revision and reinforcement of emission control strategies of PM sources.

## 1.2 Objective of this Project

The objectives of this study were:

- to characterize the contributing PM sources in Hong Kong using various receptor modeling techniques, including: Positive Matrix Factorization (PMF), Principle Component Analysis (PCA), Unmix and Chemical Mass Balance (CMB);
- to provide insights into the transport history of the air masses by integrating the source apportionment results with meteorological data through one-dimensional non-parametric wind regression (or related methods) and air mass back trajectory analysis (e.g. The Hybrid Single Particle Lagrangian Integrated Trajectory Model - HYSPLIT);
- to evaluate the relationship between PM<sub>2.5</sub> and PM<sub>10</sub> source apportionments in Hong Kong; and
- to develop a conceptual model based on the findings from the above objectives. Specifically, the observed ambient PM data was worked backwards in time, by means of combining emission source characterization and physical transport processes analysis, to identify and quantify the precursors and emission sources (see Pun and Seigneur 1999, Watson *et al.* 2002).

## 2. Data Analysis Methodology

### 2.1 Receptor Modeling

The fundamental principle of receptor modeling is to identify the source to receptor relationship by mass balance and conservation. Multivariate receptor models are based on the analysis of the correlation between measured concentrations. The assumption is that highly correlated compounds come from the same source and the chemical characteristics of the sources are constant throughout the measurement period. Receptor models can apportion the measured speciated data into groups – generally referred to “sources” – via identifying intrinsic statistical features of the data and imposing reasonable constraints, and estimate the contribution from each source to the ambient particulate matter loading collected on each sample. In this study, a series of receptor models, including Positive Matrix Factorization (PMF), Principal Component Analysis / Absolute Principal Component Scores (PCA / APCS), and Unmix were applied to the speciated PM<sub>10</sub> and PM<sub>2.5</sub> data collected at multiple stations in Hong Kong. The Chemical Mass Balance (CMB) was also applied to the PM<sub>10</sub> data.

#### 2.1.1 Positive Matrix Factorization

PMF is an advanced receptor modeling technique that apportions the measured concentrations to a set of factors which are inferred to represent emission source categories. The model utilizes nonnegative factor elements and is a weighted least square fit. Each individual measurement is weighted by an estimate of its uncertainty. Any matrix  $X$  can be factorized into two matrices by mass conservation as

$$X_{ij} = \sum_{p=1}^n G_{ip} \bullet F_{pj} + E_{ij} \quad (2-1)$$

where  $X$  is the matrix of measurement data,  $G$  and  $F$  are matrices generated by the factor analysis, and  $E$  is the residual matrix for the unexplained part of  $X$ . Thus,  $G$  and  $F$  represent almost all the primary variation of  $X$  and leaves  $E$  as secondary noise.

The objective of PMF is to minimize the sum of squares,  $Q$

$$Q = \sum_{i=1}^r \sum_{j=1}^c \left[ \frac{e_{ij}}{s_{ij}} \right]^2 \quad (2-2)$$

where  $e_{ij}$  is the element of  $E$ ,  $s_{ij}$  is the uncertainty of the corresponding element in  $X$ . Equation 2-2 is solved iteratively, proceeding until the difference in the  $Q$  value from one iteration to the next is less than a prescribed tolerance. This process is repeated multiple times using different initial values for elements of  $G$  and  $F$  in order to obtain the optimum solution. Minimization of the sum of squares on residual error,

weighted inversely with the uncertainty, allows flexibility in treating outliers, missing data and below method detection limit (MDL) data. An approach to estimating the uncertainty of each element will be described in section 2.2.

### 2.1.2 Principal Component Analysis / Absolute Principle Component Scores

PCA is a multivariate technique to reduce the dimensionality in a data set by identifying a new set of fewer independent variables which explain most of the covariance in the original data set. The new variables are linear combinations of the original variables. The general approach to multivariate modeling of air quality data is to decompose a concentration matrix  $\mathbf{C}$  into matrices of factor loadings ( $\mathbf{F}$ ) and factor scores ( $\mathbf{G}$ ),  $\mathbf{C} = \mathbf{GF}$ . In the ideal case, the factors represent emission source categories and scores can be used to compute source contributions to each sample.<sup>1</sup> In this project, PCA with APCS, hereafter called Absolute Principal Components Analysis (APCA), was used to obtain source profiles and source contribution estimates. The method is described by Thurston and Spengler (1985) and consists of the following steps.

1. Starting with a concentration matrix  $\mathbf{C}$  consisting of  $n$  species and  $m$  observations, the vector of data for each species is standardized to have a mean of zero and standard deviation of unity. The  $n \times m$  matrix of Z-scores is termed  $\mathbf{Z}$ .
2. The  $\mathbf{Z}$  matrix is decomposed into  $\mathbf{W}$  ( $n \times p$ ) and  $\mathbf{P}$  ( $p \times m$ ) matrices by principal components analysis where  $\mathbf{W}$  is the component coefficients matrix and  $\mathbf{P}$  is the component scores matrix. The dimension  $p$  is the number of components (maximum value of  $n$ , the number of latent roots). This decomposition step also generates the set of  $n$  eigenvalues,  $\lambda$ , which are screened as described in step 4 below. The decomposition was performed using MYSTAT 12 software package (SYSTAT Software, Inc., Chicago, IL/USA) although many statistical analysis software packages can be used to perform this analysis.
3. A scoring matrix  $\mathbf{B}$  is calculated from  $\mathbf{W}$  and the eigenvalues  $\tilde{\lambda}$ .
4. After choosing the number of components  $p$  to be retained in the analysis, which in source apportionment studies is deemed the number of emission source categories impacting the receptor sites(s), the PCA solution is rotated. In this analysis, Kaiser's criterion was applied which retains the  $p$  components with eigenvalues greater than unity (Kaiser, 1960). Following standard modeling practice, a varimax rotation was used to obtain the rotated scoring matrix  $\mathbf{B}^*$  and rotated principal component scores  $\mathbf{P}^*$ .<sup>2</sup> Steps 2-4 were performed using

---

<sup>1</sup> The terms "components", "factors", and "source profiles" are used interchangeably in this report.

<sup>2</sup> There is no unique solution to the over-determined system and the solution is typically rotated. A varimax rotation is commonly used to generate orthogonal (i.e. Non-correlated) sources although in reality the emission sources might be correlated to some degree, e.g., if the emission sources are

- the MYSTAT software package although many statistical analysis software packages can perform this analysis.<sup>3</sup>
5. The “absolute zero” component score is calculated for each component by calculating the Z-scores at zero concentration  $Z_0$  and then calculating the rotated absolute component scores  $P_0^*$ .
  6. A matrix of Absolute Principal Components Scores (APCS) is calculated by subtraction  $P^* - P_0^*$ .
  7. The vector of total particulate mass  $M$  is regressed on APCS and the regression coefficients together with APCS are used to generate the source contribution estimates (SCE) to ensure mass conservation. For this analysis and consistent with the PMF analysis, total particulate mass was obtained by summing the measured species in each sample. The multivariable regressions were also performed using MYSTAT.

Additional details are provided in the paper by Thurston and Spengler (1985) which includes an appendix with an example that can be used to learn the mechanics of the method. Note that this method does not guarantee non-negativity for the factors loadings or source contribution estimates.

### 2.1.3 Unmix

The Unmix model provides another solution to Equation 2-1 using a self-modeling curve resolution algorithm. This algorithm searches for “edges” in the data that define  $F$ . These edges exist because some samples are devoid of contributions from at least one source or are dominated by contributions from one source for a particular group of species. The number and direction of the edges derived from Unmix depend on the set of species used. In contrast, PMF solution is not as sensitive to the choice of input species. Unmix incorporates the algorithm “NUMFACT” that estimate the number of factors in the data using PCA on randomly sampled subset of the original data. In this way, Unmix can guide the choice of the number of factors used in the PMF, but is not good at identifying minor factors with relatively smaller contributions.

### 2.1.4 Chemical Mass Balance

CMB is also a typical tracer model, in which a specific compound, that is associated with a particular type of source, is used to quantify the contribution of each source. CMB is also essentially based on Equation 2-1. However, the fraction of each species of the total mass known as source profile and the receptor concentrations, each with realistic uncertainties estimate serve as input data to the CMB model. Source contributions are then quantified by a least-square multiple regression analysis

---

located in the generally same spatial zone and have similar temporal emission patterns. In this case, the sources might be admixed into a single factor or, if there is sufficient differences in covariance, split into two factors that are correlated.

<sup>3</sup> The same commands are used to perform this analysis in MYSTAT and SYSTAT.

on the total PM mass.

## 2.2 Data Quality Assessment Using Collocated Data

### 2.2.1 Motivation

As previously mentioned, uncertainty estimates for the reported concentrations values are required for source apportionment modeling by PMF and recommended for modeling by CMB. Preliminary PMF modeling by Lau *et al.* (2010) on the Hong Kong PM<sub>10</sub> 1998-2008 speciation data set used the relative expanded uncertainty and method detection limit values provided by the HKEPD to generate error structures (uncertainty estimates as a function of ambient concentration values for each species). Concerns over the assumed form of the error structure led us to conduct a detailed data quality assessment to support not only the source apportionment modeling but also other uses for these data. Discussions with HKEPD revealed that collocated PM<sub>10</sub> sampling is routinely performed at two sites (SSP and TW). The collocated sampler data were organized by HKEPD staff and provided to the project team. For each species the scaled relative difference (collocated precision relative to the arithmetic mean concentration for the data) was calculated using data from one site. These precision estimates together with the HKEPD-provided detection limits were used to generate error structures using the method of Polissar *et al.* (1998). Refined PMF modeling was performed and these results have been documented (Interim Report #1, Amended, July 2010). While this *interim approach* was a significant improvement to the PMF modeling strategy, it still had inconsistencies and potential bias. For example, analysis of the collocated data assumed the relative precision was constant over the entire concentration range but the subsequent PMF modeling used the Polissar *et al.* (1998) error structure equation which includes both an additive term (based on the detection limit) and multiplicative term that is the scaled relative difference times the concentration value. Thus, the Polissar *et al.* (1998) approach inherently assumes the relative precision is not constant over the entire concentration range.

The remainder of this section summarizes our detailed data quality assessment using the collocated data that resulted in the development of error structures used for the PM<sub>10</sub> and PM<sub>2.5</sub> modeling presented in this report. Analysis of the collocated data can capture many, but not all, sources of measurement error. For example, if the samples are analyzed in the same batch then it cannot account for systematic errors in the extraction and chemical analysis (e.g., analyzer calibration). That said, it does provide valuable insights into the data quality by quantifying the aggregate errors from field sampling operations and analytical sensitivity. The analysis adopted the methodology of Hyslop and White (2008a,b) with some deviations to account for differences in the information available. Hyslop and White (2008a) used field blanks data to estimate “critical limits” for the concentration values and compared these estimates to the method detection limits reported for the IMPROVE network. Subsequently, Hyslop and White (2008b) examined the collocated precision as a

function of concentration. From these two analyses, key derived data quality metrics were the critical limits and the collocated precision for the subset of data with concentration values at least three times greater than the method detection limit. In our work, field blanks data were not available. Thus, the method detection limit values provided by HKEPD could not be independently examined.

### 2.2.2 Methodology and Results for PM<sub>10</sub> Analysis Data

The original PM<sub>10</sub> data transmission for this project included gravimetric mass and species concentration values for the primary PM<sub>10</sub> high volume sampler at each site. Species data (but not gravimetric mass data) were subsequently provided by HKEPD for the collocated PM<sub>10</sub> high volume samplers at the SSP and TW sites spanning the period 1998 to 2008. In the ideal case, all data for this analysis would be at the measured concentration values even if below the MDL so the precision could be determined for cases where the concentration value from one sampler was above the MDL and from the other sampler was below the MDL. Data for the primary sampler at both sites was already conditioned with concentration values below the method detection limit (MDL) imputed as half the MDL; in contrast, the collocated sampler data for SSP and TW included concentration values below the MDL were marked as "<XXX" where XXX is the sample-specific MDL value. Subsequently, HKEPD provided data for the primary sampler at SSP for the period 2004 to 2008 with concentration values below the MDL marked as "<XXX". Thus, we used the SSP collocated data to generate error structures which was then evaluated using the TW collocated dataset. Well-behaved trends between concentrations above MDL and the estimated precision were observed to break down for concentrations below the MDL, with generally poor and fluctuating precision for concentrations below MDL. Hence, sampling days with one or both concentrations below the MDL were excluded from the error structure estimation *t* (additive and multiplicative error terms of uncertainty). SSP collocated data for 1998 to 2008 was used to generate error structures which were then evaluated using the corresponding TW collocated dataset for 1999 to 2008.

The objective of this data quality assessment was primarily to inform the source apportionment modeling. For this application, we needed to separate any systematic bias between the collocated data from random error. Scatter plots were generated to visually check for measurement bias between the primary and collocated sampler data. Both linear and log scaling were used to emphasize different portions of the concentration range. As expected, no sampler-to-sampler bias could be visually detected. Next, the methodology of Hyslop and White (2008b) was taken to examine the precision and bias as a function of concentration. The mean concentration, scaled arithmetic difference and scaled relative difference were calculated for each data pair. Scaled arithmetic difference is the difference of the two measured concentrations divided by  $\sqrt{2}$ . Scaled relative difference is the scaled arithmetic difference divided by the mean concentration for the data pair. Scaled arithmetic and relative differences were plotted against the mean concentration to examine the concentration dependence of measurement differences; linear and log

concentration scales were used to emphasize high and low concentration values, respectively. Examples are shown in **Figures 2.1-2.3** for each of the three classes of species: elements (arsenic, Figure 2.1); ions (nitrate, Figure 2.2); and carbon (elemental carbon, Figure 2.3). In each case panels (a) and (b) are the scaled arithmetic differences using a linear and logarithmic concentration axis, respectively; and panels (d) and (e) are the scaled relative differences using a linear and logarithmic concentration axis, respectively. For all species (including those not shown), as concentration values approach the MDL from above, the relative differences remain within 20%. At concentrations greater than the MDL the relative differences decrease with increasing concentration and for many species approach an asymptotic value. These patterns indicate that a concentration-dependent error structure is justified. However, as demonstrated in Figures 2.1-2.3 the arithmetic and scaled differences tend to stratify into bands for certain concentration ranges. This behavior is particularly evident in the scaled relative difference plots (panels d and e) and arises from the number of significant figures retained in the data set. For each species the laboratory reported the data to two significant digits regardless of the concentration value (e.g. 0.013, 0.13, 1.3  $\mu\text{g}/\text{m}^3$ ). For example, the nitrate data in Figure 2.2e exhibits banding in the region near 1  $\mu\text{g}/\text{m}^3$  which correspond to differences of 0.0,  $\pm 0.1$ ,  $\pm 0.2$   $\mu\text{g}/\text{m}^3$ . As described below, this behavior has implications to the generation of error structure estimates.

The scaled arithmetic and relative differences plots provide insights into the data quality including the concentration dependence of the measurement precision. To generate the error structures which were subsequently used to assign a concentration dependent uncertainty to each concentration value, the data pairs were ordered by mean concentration and the collocated precision was calculated for fifteen bins (each with nominally fourteen data pairs) across the concentration range. Examples for the binned absolute and relative collocated precision are shown in Figures 2.1-2.3, panels (c) and (f), respectively. The trends are not strictly monotonic and this behavior largely arises from artifacts introduced by the aforementioned rounding protocol for the concentration values.<sup>4</sup> The following error estimation procedure would be more robust if more digits were retained in the reported data.

Following Polissar *et al.* (1998) the error structure was assumed to have the form:

$$U_{ij} = a_j + b_j \times C_{ij}$$

where  $U_{ij}$  = uncertainty for species  $j$  in sample  $i$ ;  $a_j$  = additive (constant) error term for

---

<sup>4</sup> For EC and OC the samples were analyzed by different laboratories during the period of the study. Most of the data were reported to two significant digits (e.g. 11  $\mu\text{g}/\text{m}^3$ , 1.1  $\mu\text{g}/\text{m}^3$ ) but some of the data were reported with several significant digits (e.g., ###.#####). These differences dramatically affected the collocated precision estimates. After rounding all concentration values to two significant digits the precision was better behaved. These conditioned data were used to generate the EC plots of Figure 2.3.

species  $j$ ;  $b_j$  = multiplicative error term for species  $j$ ; and  $C_{ij}$  = concentration value for species  $j$  in sample  $i$ .

The first error estimation strategy used the approach of Polissar *et al.* (1998) which sets the additive error term to MDL/3. MDL values were estimated as follows: (i) for species with some concentration values reported as “<XXX”, the arithmetic mean of the XXX values was taken as the MDL; (ii) for species with no concentration values below the MDL (i.e. no data reported as “<XXX”), the average of laboratory-reported MDL range provided by HKEPD was used; and (iii) for aluminum which was the only species with no data below the MDL nor a laboratory-reported MDL, the MDL value assumed by Lau *et al.* (2010) was used. The multiplicative error term was first estimated by taking the mean of the binned relative collocated precision for the six bins with highest concentration values. The additive and multiplicative errors for this approach are presented in the “Unweighted” columns of **Table 2.1**. The regression fits using this approach were poor for most species and thus another approach was taken.

The second approach used a weighted linear regression to estimate both the additive and multiplicative error terms using all fourteen bins. The weight  $w_{kj}$  for species  $j$  and bin  $k$  was taken as the inverse of square of the absolute collocated precision ( $\sigma_k$ ) for bin  $k$  (dropping the  $j$  subscripts on each parameter):

$$w_k = 1/\sigma_k^2$$

with the coefficients for the error structure calculated using:

$$a_j = \frac{(\sum w_k C_k^2)(\sum w_k \sigma_k) - (\sum w_k C_k)(\sum w_k C_k \sigma_k)}{(\sum w_k)(\sum w_k C_k^2) - (\sum w_k C_k)^2}$$

$$b_j = \frac{(\sum w_k)(\sum w_k C_k \sigma_k) - (\sum w_k C_k)(\sum w_k \sigma_k)}{(\sum w_k)(\sum w_k C_k^2) - (\sum w_k C_k)^2}$$

For most species the weighted error estimates, summarized in the “Weighted” columns of Table 2.1, provided better fit than the unweighted error estimates. The weighted error estimates were further refined by a third approach which included removing individual sample pairs that were subjectively deemed outliers and significantly influenced the collocated precision for a given bin. These results are presented in the “Refined Weighted” columns of Table 2.1 and these values are recommended for the source apportionment modeling. To further characterize the recommended approach, **Table 2.2** shows the number of data pairs with one-or-more concentration values below the MDL and the number of data pairs deemed outliers for the refined weighted regression; and MDL values provided by HKEPD (or estimated from the collocated data) and estimated as three times the additive error value ( $a_j$ )

from the refined weighted regressions reported in Table 2.1. These are not actual MDL values but rather would be the *effective MDL* using the error structure of Polissar *et al.* (1998). The effective MDL was generally lower than the reported MDL values which demonstrates the “reported” MDL values are typically conservative for generating error structures for source apportionment modeling.

Collocated data from the TW site were used as an independent check on the modeled error structures. **Table 2.3** summarizes the number of TW data pairs with one-or-more values below the MDL; and the number of sample pairs that were deemed outliers and removed from the data set to stabilize the binned collocated precision estimates. **Figure 2.4** shows for several species the binned collocated precision for the TW data (open markers for raw data and shaded markers for data after removing outliers) and the error structures generated from the “Refined Weighted” regression coefficients for the SSP collocated data (dashed lines). In general, the concentration-dependent collocated precision for the TW data is well captured by the SSP-generated error structures. A few species had one or two concentration bins that were not explained by this error structure, for instance the highest concentration bins for As and Mg (Figure 2.4) were not matched well by the error structure. Overall, these favorable comparisons add confidence to using the “Refined Weighted” error structures of Table 2.1 for source apportionment modeling.

### 2.2.3 Methodology and Results for PM<sub>2.5</sub> Data

Record-specific uncertainties for the PM<sub>2.5</sub> dataset were provided by the sample analyst but the representativeness of these estimates must be questioned. For example, **Figure 2.5** shows the record-specific uncertainties provided for the PM<sub>2.5</sub> Pb data collected during the 2009-2010 sampling year. The error structures (uncertainty versus concentration) stratify into four groups that align with the analysis batches. It does not seem reasonable that for a given concentration the uncertainty would vary by nearly a factor of five depending on the batch. Therefore, error structures for the purpose of source apportionment modeling were generated using collocated data. Compared to the PM<sub>10</sub> collocated data set, the PM<sub>2.5</sub> collocated data set has the advantage of having concentration value estimates reported for species below the MDL and the disadvantage of being a much smaller data set (47 pairs for 2009). For each species the collocated data pairs were ordered by the pairwise mean concentration and grouped into six bins with nominally eight data pairs per bin. Collocated precision was calculated for each of the bins and then a weighted linear least squares regression was performed to determine the coefficients ( $a_j$ ,  $b_j$ ) for an assumed error structure of  $U_{ij} = a_j + b_j C_{ij}$  where:  $U_{ij}$  and  $C_{ij}$  are the uncertainty and concentration, respectively, for species  $j$  in sample  $i$ ; and  $a_j$  and  $b_j$  are additive and multiplicative error terms, respectively, for species  $j$ . This is the form proposed by Polissar *et al.* (1998) who set  $a_j$  to  $MDL_j/3$  and  $b_j$  to the relative error at concentrations much greater than the MDL. Error structure coefficients derived from regressions on the binned collocated precision are presented in the last two columns (“Weighted Regression”) of **Table 2.4**. **Figure 2.6** shows the relationships

between  $a_j$  derived from the collocated data and the MDL and limit of quantification (LOQ) values provided by DRI. In each case  $a_j$  is multiplied by a single-valued parameter that was adjusted to minimize the sum-of-squares difference between the data sets after log-transforming the data. For the MDL, the best-fit relationship is  $a_j = \text{MDL}/2.8$  which is in excellent agreement with the value of 3 assumed by Polissar *et al.* (1998). Excellent agreement is also observed between  $a_j$  and the LOQ with best-fit relationship  $a_j = \text{LOQ}/5.5$ . The error structures used for the source apportionment modeling were generated using the “Weighted Regression” coefficients presented in Table 2.4.

#### 2.2.4 Recommendations

- The PM<sub>10</sub> collocated data relationships suggest that additional significant figures could be retained in the data set for some of the species. An analysis should be conducted to determine the appropriate number of significant figures. This refinement will approve the ability to assess measurement uncertainty using the collocated data and might improve the data quality.
- For PM<sub>10</sub>, unexposed filters (laboratory blanks) are routinely analyzed by the Government Laboratory but field blanks are currently not being collected. It would be very helpful to routinely collect field blanks data so the limit of quantification (LOQ) could be assessed. The field blanks LOQs are expected to be higher than would be estimated from the laboratory blanks because there can additional contamination of the filters during transport and handling. LOQ might provide a better estimate than the laboratory MDL or effective MDL when assessing species detectability which is an important data pre-conditioning step for the source apportionment modeling.
- Collocated samples are essential to accurately estimate the uncertainties associated with the measurements. While it might be impractical to routinely collect collocated samples at each site, the frequency of collocated measurements, especially for PM<sub>2.5</sub>, could be increased compared to the prior years data to build a large data set for statistical analysis.

### 2.3 Air Mass Transport Pattern Analysis

Air mass back trajectories are commonly used to identify and interpret relationships between observed air quality parameters<sup>5</sup> and air mass history. The various approaches have been briefly reviewed by Stohl (1998). One approach to trajectory-based analysis of air quality data assigns the observed concentration at a given receptor and time to every location along the corresponding back trajectory. This is repeated for an ensemble of concentration values. The spatial domain is gridded and some measure of the receptor concentration (average value, anomaly, etc.)

---

<sup>5</sup> Termed concentrations for the purpose of this discussion, but could be receptor-modeled source contribution estimates or other parameters.

is calculated for each grid cell. Various approaches are taken to calculating the cell-specific values depending on the time the air mass spends in the cell. The resulting gridded values are typically smoothed to generate contour plots for the relationship between the concentrations at the receptor and the spatial histories of the corresponding air masses. The approaches are generally termed *residence time analysis* or the generation of *conditional probability fields*; specific techniques such as the potential source contribution function (PSCF) and quantitative transport bias analysis (QTBA) have been extensively used with substantial success. For example, in St. Louis the surface winds are rarely from the east but synoptic air masses from the east are relatively frequent due to the stalling of high pressure systems over the Great Lakes with the associated clockwise rotation bringing air from the Ohio River Valley to St. Louis. These air masses are associated with high sulfate concentrations in St. Louis which is consistent with the Ohio River Valley being a large SO<sub>2</sub> emission source region due to numerous coal-fired electric utility power plants. Similarly, we also recently demonstrated that our best estimate of the *regional component* of PM<sub>10</sub> arsenic in St. Louis follows similar behavior which is consistent with coal combustion being a major source of arsenic on the national scale.

There are challenges in the application of these tools to the Hong Kong data. Perhaps most importantly, important emission source regions cannot be broadly classified as local versus regional and a three-tiered classification is usually assumed – local, PRD (outside the Hong Kong Special Administrative Region (HKSAR) but still with a few hundred kilometers, mainly concentrated in the Pearl River Delta), and non-PRD (long-range transport). The spatial dimension for the PRD sources is on the order of the spatial resolution of the meteorology data used for the trajectory generation (nominally 100 km) and thus their impacts cannot be resolved. This could also lead to the misinterpretation of non-PRD transport impacts because the analysis would be confounded by the PRD-scale contributions. There would be value in applying such tools to the Hong Kong data if they were readily accessible (i.e. off-the-shelf applications) but this is not the case. Furthermore, to be truly meaningful it would be necessary to generate air mass trajectory fields at much finer spatial resolution than routinely available.

Given these caveats, we decided to use an alternative approach which is based on the categorization of the air mass back trajectories. In this case each trajectory is classified (“typed”) according to predominant air mass transport patterns.<sup>6</sup> This *cluster analysis* approach does not directly provide the detailed correspondence of observed concentrations with the specific upwind, prior locations of the air masses but does offer numerous advantages. First, the typing of trajectories is performed

---

<sup>6</sup> Care should be taken to distinguish this classification approach – which is based solely on air mass back trajectories and thus capturing air mass transport patterns – from “synoptic typing” which considers other properties of the air mass. Air masses with similar transport histories are presumably exposed to similar emission fields but might have different atmospheric processing capacity depending on factors such as moisture content.

independent of the air quality data. Each trajectory is assigned to an air mass transport pattern. The resulting data set can then be used to analyze and interpret any parameter of interest (e.g. Observed concentration values, source apportionment modeled source contribution estimates) without embedding that parameter in the trajectory analysis. Second, this approach informs a broader question than the residence time analysis or concentration probability fields. It simply relates the observed concentrations to air mass transport patterns and makes no inference whether the pollutant was regionally transported or the transport pattern is a proxy for local meteorological conditions that modulate local source impacts (e.g. Atmospheric ventilation). Whereas the residence time analysis and conditional probability fields approaches are operationally meaningful only for parameters dominated by regional (and in this case, only non-PRD) sources, the cluster analysis approach does not have this limitation. That said, for the case of pollutants that are truly from long-range transport the cluster analysis approach provides a much weaker link to putative emission source geographic regions.

The analysis includes two steps – the generation of an air mass trajectory data set and the clustering of these trajectories into groups with similar transport patterns. For this work we used seven day back trajectories generated by Perry Samson (University of Michigan) who, under contract from the U.S. EPA, recently compiled data sets of air mass back trajectories with coverage from 2000 to 2009 for 109 US and international sites (<http://www.world.sharedair.org/data/>). Two sites in the HKEPD's air quality monitoring network are included in the data set - the Central/Western (CW) station on Hong Kong Island and Tai Po (TP) in the New Territories – to represent the entire Hong Kong condition, as the spatial resolution of the back trajectories are  $1^{\circ} \times 1^{\circ}$  (~100km x 100km), there should not be much difference among stations in Hong Kong due to its small spatial coverage (~40 km x 30 km). Each data sets includes four trajectories per day (arrival times of 00, 06, 12, and 18 UTC) generated using Revision 4.9 of the NOAA HYSPLIT4 Model (Draxler, 1999). Meteorological data sets used to generate the trajectories were FNL (2000-2003) and GDAS1 (2004-2009;  $1^{\circ}$  resolution). The trajectory arrival height was fixed at 500 m and trajectory endpoints were generated for every two hours for up to seven days duration. The CW data set used for this study includes 14,546 trajectories out of a possible 14,608 trajectories (99.6% completeness). 70% of the trajectories include the entire seven day duration while 30% of trajectories truncated prior to seven days (e.g. Because they exited the modeling domain).

The complete set of trajectories was clustered using refinements to the methodology of Dorling *et al.* (1992) as shown in **Figure 2.7**. Briefly, a subset of N trajectories is randomly chosen as seeds. For each trajectory, the 2D Great Circle distance separating the trajectory from the same time stamp in each seed is calculated. The separation distances are summed along the trajectory history and the trajectory is assigned to the seed which minimizes the separation distance accumulated along the trajectory path. After assigning all trajectories to a cluster, new mean trajectories – which replace the seeds – are calculated to represent the clusters. The entire process

of assigning trajectories to clusters is repeated until there is no change in the trajectory-cluster matching with each update to the cluster-mean trajectories. The two closest clusters – in the spatial sense – are identified and the cluster with the fewer trajectories is removed. The trajectory assignment and cluster pruning process is repeated to obtain results for sets of  $(N-1) \rightarrow 2$  clusters. For each set of clusters, a global separation distance is calculated based on the separation of each trajectory from the mean trajectory for the respective cluster, summed over all trajectories in the data set. The final set of clusters, with trajectory assignments, is subjectively determined by choosing the cluster set for which an increase in the number of clusters does not significantly reduce the global separation distance. For this specific application, ten seeds ( $N=10$ ) were used to obtain eight sets of clusters, and the five-cluster solution was selected. As a sensitivity check, the clustering was repeated using only year 2009 trajectories and sixteen seeds; a nearly identical five-cluster solution was obtained.

The results from the cluster analysis include the assignment of each trajectory to a cluster, a measure of its separation distance from the mean trajectory for the respective cluster, and a flag indicating whether the trajectory was complete (seven days) or truncated. Time stamps were converted from UTC to Hong Kong local time.

### 3. Source Apportionment of PM<sub>10</sub> in Hong Kong

#### 3.1 Data Collection, Analysis and Pre-processing

Once every six days, 24-hour PM<sub>10</sub> sampling was carried out at ten Air Quality Monitoring Stations (AQMSs) by the HKEPD between January 1998 and December 2008 (11 years in total). Seven out of ten stations have rather complete sampling records over the eleven years, and were used in this study. These stations include Yuen Long (YL), Tung Chung (TC), Tsuen Wan (TW), Sham Shui Po (SSP), Central/Western (CW), Kwun Tong (KT), and Mong Kok (MK). The MK roadside station was relocated in late 2000 and therefore only samples collected in or after 2001 were retained for the analysis. Distribution of the seven AQMSs is shown in **Figure 3.1**, and the sampling period and environmental characteristics at each station are summarized in **Table 3.1**. Further information about these monitoring stations is available from HKEPD (2008). 22 Samples with missing values or abnormally high records for some species are removed, leading to 4167 samples retained in the analysis. Data pre-processing procedures are summarized in section 3.1 of a previous report prepared for HKEPD under project AS06-435 (Lau *et al.* 2010). Whereas Lau *et al.* (2010) used the sum of measured species mass to generate the source contribution estimates, *in this study we apportioned the quartz filter gravimetric mass* (hereafter denoted as QTM). Historically, concerns have been raised about the quality of QTM data such as bias arising from the hygroscopic nature of quartz fiber and the possibility of losing fibers during filter handling between pre-weighing and after-weighing. However, the sum of measured species approach accounts for only 62% of the QTM and, as demonstrated in section 3.2.4, QTM is a reasonably good proxy for daily-average TEOM PM<sub>10</sub> mass. A major advantage of this approach is that all PM<sub>10</sub> mass is being apportioned, not just the species that were measured. The rationale, advantages and limitations of apportioning QTM rather than the sum of measured species mass is discussed in detail in section 3.2.4.

#### 3.2 Source Apportionment by PMF

As discussed in section 2.2, uncertainty estimation is an important step for PMF application, in which each of the data values is assigned an estimated uncertainty including both measurement uncertainty and source profile variability. However, there is no simple rule in calculating uncertainty estimates. In this study, the error structures of Polissar *et al.* (1998) were applied, with the coefficients determined by weighted linear regression of collocated samples. *This approach to estimating measurement uncertainty represents a significant refinement to the approach taken by Lau et al. (2010)*. The methodology to arrive at the coefficients ( $a_j, b_j$ ) was presented in section 2.2.2 and the sample-specific uncertainties were estimated using:

$$U_{ij} = a_j + b_j \times c_{ij} \quad \text{For determined values}$$

$$U_{ij} = 5d_j/6 = 5a_j/2$$

For values below detection limits

Here,  $c_{ij}$  represents the ambient concentration of species  $j$  in sample  $i$  and  $d_j$  represents the effective MDL of the species  $j$ , calculated as three times the  $a_j$  values from the “Refined Weighted” additive error term estimates. Table 2.2 lists percentages of samples below the effective MDL for each species.

Source apportionment modeling by PMF was performed using EPA PMF Version 3.0. 10% of extra modeling uncertainty, which corresponds to  $C_3$  value of 0.1 in the PMF4.2, was imposed. 20 base runs were performed and the run with the minimum  $Q$  value was selected as the base run solution. 100 bootstrap runs were conducted with minimum correlation R-value of 0.6 to examine the stability and to estimate the uncertainty of the base run solution. All of the bootstrapped factors were uniquely matched with a factor from the base solution. For the  $Q$  values derived from the 20 base runs, their standard deviation was only ~0.0004% of their mean, indicating a very stable solution. For this modeling, an Fpeak value of zero resulted in the most meaningful results. Derived source profiles and their uncertainties are tabulated in **Table 3.2**, and the percentages of each species apportioned to each factor are plotted in **Figure 3.2**. Actual source categories are mainly determined by identifying the tracer of each source, which is typically the species that exclusively or largely resides in a particular source. **Table 3.3** lists the emission source categories and corresponding tracers used for matching PMF-resolved factors to sources in this study. Detailed information about the tracers can be found in Yuan *et al.* (2006) and Huang *et al.* (2009).  $PM_{10}$  emission source categories including vehicular exhaust, residual oil combustion, fresh sea salt, aged sea salt, crustal material/soil, secondary sulfate, secondary nitrate, trace metals and coal combustion/biomass burning, were assigned to the nine PMF-resolved factors. It is important to recognize that *the secondary sulfate and secondary nitrate factors do not include exclusively these measured species*. Ammonium and sulfate account for ~70% of the  $PM_{10}$  mass assigned to the secondary sulfate factor and ammonium and nitrate account for ~45% of the  $PM_{10}$  mass assigned the ammonium nitrate factor. As described below, in both cases most of the remaining mass can be explained by organic matter associated with the OC loaded onto these factors. This behavior is expected since factors can represent sources and/or processes and therefore secondary organic and inorganic aerosols are often loaded admixed into factors.

The change from apportioning sum of measured species mass to apportioning the QTM mass necessitates a thorough assessment of mass closure to determine whether the unidentified mass loaded onto each factor can be rationalized. Recall that, on average, the sum of measured species mass was 62% of the QTM mass; thus, 38% of the QTM mass will be unidentified even if the source apportionment modeling perfectly fits the measured species. Mass closure was assessed for each profile to evaluate whether the percentage of unidentified mass was plausible. We assumed multipliers of 1.9 for Al, 1.41 for Ca, and 2.09 for Fe to account for these elements usually being present as oxides.

- *Vehicle Exhaust.* Mass closure using the sum of measured species is 72% (i.e. unidentified mass is 28%). Mass closure is 90% using an organic matter (OM) to OC ratio of 1.6 and closure is achieved for an OM/OC ratio of 1.85. Thus, most of the missing mass in the vehicle exhaust profile can be explained using reasonable estimates for the non-carbon component of organic matter.
- *Residual Oil Combustion.* Mass closure using the sum of measured species is 62%. Mass closure is 87% using the oxides multipliers and OM/OC = 1.6, demonstrating that reasonable mass balance closure can be achieved especially given the large uncertainties assigned to both the  $\text{SO}_4^{2-}$  and OC loaded onto this source profile.
- *Fresh Sea Salt.* Mass closure using the sum of measured species is 80% and using OM/OC = 1.6 the mass closure is 83%. The source profile has a  $\text{Na}^+/\text{Cl}^-$  ratio of 0.43 which is low compared to standard seawater which has a ratio of 0.56. Given that  $\text{Na}^+$  is predominantly in the aged sea salt factor and  $\text{Cl}^-$  is almost exclusively in the fresh sea salt factor, an upward adjustment of  $\text{Na}^+$  in the fresh sea salt factor to match the standard composition of water brings the mass closure to 89%. No  $\text{SO}_4^{2-}$  loaded onto the fresh sea salt factor whereas standard seawater is about 7-8%  $\text{SO}_4^{2-}$ . If  $\text{SO}_4^{2-}$  is added proportional to the  $\text{SO}_4^{2-}/\text{Cl}^-$  ratio in standard seawater, mass closure is 95%. Thus, mass closure can be rationalized but the pulling up on measured species should be performed in EPA PMF 4, to be released in 2012, to determine whether this pulling is plausible and if so, whether the total SCE assigned to the factor is also increased (if so, there still might be a mass closure gap).
- *Aged Sea Salt.* Mass closure using the sum of measured species is 70%. A charge balance including the major cationic species in sea water ( $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{Na}^+$ ) and the anions from the aging process ( $\text{NO}_3^-$ ,  $\text{SO}_4^{2-}$ ) is balanced to within 3% so there is no evidence for such species being inadequately distributed across the resolved factors. The reason for the 30% unidentified mass is not known.
- *Crustal Soil/Dust.* Mass closure using the sum of measured species is only 22% and increases to 30% using the oxides multipliers. A major issue is that certain key species present in crustal material, especially Si, were not measured. Al, Ca, and Fe concentrations in the resolved factor were used to estimate Si and Ti from the Hong Kong  $\text{PM}_{10}$  paved road dust profile of Ho *et al.* (2003). Together with oxides multipliers of 2.41 for Si and 1.67 for Ti, the addition of these species increase the mass closure to 57-62% depending on the PMF factor to paved road profile scaling method. No OC loaded onto the resolved crustal soil/dust factor. Based on the Hong Kong soil and dust profiles of Ho *et al.* (2003) a reasonable estimate of the OC mass fraction would be 5-15% with paved road dust at the higher end of this range. The OM/OC ratio for  $\text{PM}_{10}$  crustal soil/dust is not well characterized; however, an estimated range for  $\text{PM}_{2.5}$  paved road dust is 2.8-3.4 (Bae *et al.* 2006). Thus, OM could be 15-50% of the crustal soil/dust mass and

the upper end of this range would nearly result in mass closure. It is noted that the Ho *et al.* (2006) PM<sub>10</sub> paved road dust profile mass closure using the species measured in this study is 39% when including OC and 26% when excluding OC. The latter is in good agreement with the 22% mass closure in the PMF-resolved factor. If much of the crustal soil/dust is from paved roads, there could be confounding with the OC loading onto the vehicle exhaust profile. The ability to pull OC into the PMF-resolved crustal soil/dust profile should be investigated when EPA PMF 4 is available.

- *Secondary Sulfate.* Mass closure using the sum of measured species is 75% and increases to 78% when using an OM/OC ratio of 1.6. However, the OM/OC ratio for secondary organic aerosol could be as high as 3.75 (Bae *et al.* 2006) and in this case the mass closure increases to 90%. The ammonium-to-sulfate molar ratio is 1.9 which suggests that, on average, the sulfate is nearly fully neutralized as ammonium sulfate. The reason for the missing mass in this factor is not known.
- *Secondary Nitrate.* Mass closure using the sum of measured species is 68% and increases to 81% when using an OM/OC ratio of 1.6. If the OC loaded onto this factor is secondary organic aerosol, the OM/OC ratio could be in the range 1.73-3.75 (Bae *et al.* 2006). The midpoint of this range is 2.74 and this multiplier brings the mass closure to 108%. Thus, the missing mass in the secondary nitrate factor can be explained by the non-carbon component of organic matter.
- *Trace metals.* Mass closure using the sum of measured species is 32%, and mass closure is 48% using the oxides multipliers and OM/OC = 1.6. The reason for the missing mass is not known.
- *Coal Combustion/Biomass Burning.* Mass closure using the sum of measured species is 58%, and mass closure is 79% using the oxides multipliers and OM/OC = 1.6. Biomass burning OM/OC ratios are typically in the range 1.6-1.9 and using the upper bound on this range yields 90% mass closure.

In summary, for most factors the unidentified mass can be rationalized and mass closure can be reasonably well estimated. This analysis adds confidence in apportioning QTM mass rather than the sum of measured species mass. In some cases a species that is very likely present in the emissions did not load onto the PMF-resolved factor with the most glaring case being the lack of OC in the crustal soil/dust factor. It will be constructive to use the rotational features in the next version of EPA PMF to examine whether feasible solutions can be obtained with such species pulled into the respective factors. Also, uncertainties estimated by the bootstrap method are likely conservative because this approach does not account for all sources of uncertainty. The next version of EPA PMF will include additional uncertainty estimation algorithms might better reflect the true uncertainty of the

species mass fractions in the source profiles.

### 3.2.1 Source Contributions and Their Seasonal and Spatial Variations

One objective of the HK data source apportionment modeling was to make a quantitative determination of what emission source categories are local and what are from transport, and their relative weight, since the control of local sources is fully within HK's jurisdiction while controlling non-local sources awaits regional cooperation. Therefore, the seasonal and spatial analysis are especially indicative in that local sources typically have relatively larger spatial variation and smaller seasonal variation, while non-local sources should show smaller spatial variation and larger seasonal variation, considering the clear seasonality of prevailing synoptic systems in HK. Seasons are defined as summer from 16 May to 15 September, winter from 16 November to 15 March of next year, and transition season (TS), including both spring and fall, from 16 March to 15 May and from 16 September to 15 November, based on the synoptic conditions in HK.

**Table 3.4** lists the annual-, wintertime-, summertime-, and TS- averaged source contributions at general stations during the 11-year sampling period and at the MK roadside station. Based on the general station source contribution estimates (SCE), secondary sulfate is the largest contributor and accounts for 22% of the 11-year study-average ambient PM<sub>10</sub> in HK. Contributions from vehicle exhaust, aged sea salt, crustal/soil, secondary nitrate, and coal combustion/biomass burning are comparable, with each around 12-15%. Contributions from residual oil, fresh sea salt, and trace metals are generally below 5%. Specifically, the ratio of contributions between secondary nitrate and secondary sulfate is around 0.7, much higher than the ratio between nitrate and sulfate. This is due to the loading of OC on the secondary nitrate factor.

A more detailed station-wise comparison between winter and summer is shown in **Figure 3.3**. The results in both Table 3.4 and Figure 3.3 clearly show that for sources of vehicle exhaust, residual oil, fresh sea salt, and aged sea salt, the summertime contribution is comparable, or even slightly higher, than the wintertime contribution. This indicates that they are more associated with local source characteristics since there are limited emissions in the upwind direction in summer (i.e. SE or SW) and their seasonal dependence is weak. On the other hand, for sources of crustal/soil material, secondary sulfate, secondary nitrate, trace metals and coal combustion / biomass burning, spatially averaged source contributions 3-5 times higher in winter than in summer which supports the notion that these sources were largely non-local in nature, as prevailing winds from NW or NE in winter are favorable for bringing pollutants emitted from the PRD or further inland to HK.

In addition, local sources, such as vehicle exhaust, residual oil, fresh sea salt, and aged sea salt, show relatively large contribution differences between stations, while for non-local sources, such as crustal material/soil, secondary sulfate, secondary nitrate, trace metals, and coal combustion/biomass burning, show relatively consistent

contribution between stations (Figure 3.3). This spatial pattern is another indicator for the geographic nature of each emission source category. However, when the ambient monitoring stations are ordered from the northwest to the southeast, as shown in Figure 3.3, a general northwest (higher) to southeast (lower) gradient in the source contributions of non-local pollutants is discernable which suggests that some of the non-local source contributions are still sufficiently close to the HKSAR to result in spatial gradient across the monitoring network. In contrast, such spatial gradient does not exist for the local sources.

After classifying the geographic nature of each emission source category we can calculate the relative importance of local and non-local sources to the ambient PM<sub>10</sub> levels observed in HK. For general stations, non-local sources contributed around 67% of PM<sub>10</sub> during the period from January 1998 to December 2008, and local sources contributed around 32%. Non-local contributions increase to 77% in winter and decrease to 48% in summer. At the roadside MK station, non-local sources contributed around 58% of PM<sub>10</sub> during the 8-year period, and local sources contributed around 44%, with a slight 2% over-interpretation. Non-local contributions increase to 69% in winter and decrease to 36% in summer.

### 3.2.2 Annual Trends of Source Contributions over the Eleven Years<sup>7</sup>

An advantage of source apportionment analysis by PMF is that it can provide source contribution estimates at any location in any specified period. In this section, the samples are grouped into years and the annual variation of each of the identified sources is examined. **Figure 3.4** illustrates the annual trends of the nine identified sources at both general stations and roadside station. The trends before 2005 are generally similar as those reported in Huang *et al.* (2009).

Vehicle exhaust shows a clear decreasing trend for six consecutive years from 1998 to 2003 at both types of stations. With small rebounds in 2004 and 2006, its contribution continued to decrease for three consecutive years from 2006 to 2008. This general decreasing trend clearly indicates that local emission control measures on vehicles in HK are effective. The annual average contribution from vehicles was 5.3 µg/m<sup>3</sup> at general stations in 2008, which corresponds to a 54% reduction from the contribution of 11.6 µg/m<sup>3</sup> in 1998. For the rebounds in 2004 and 2006, at this point we are unable to conclude whether they are due to the inter-annual variability of synoptic conditions that are capable of trapping more locally emitted particles in 2006, or the emission from vehicles really increased in the year of 2004 due to the economic recovery from the SARS epidemic in 2003. For the years from 1998 to 2007, there is good correlation between PMF-modeled vehicle exhaust source contribution estimate at the general stations and the annual PM emission inventory for road transport, with a correlation coefficient of 0.90 (**Figure 3.4**) and even rising up to 0.99 if data from the years of 2004 and 2006 are ignored (the emission inventory was

---

<sup>7</sup> For the sake of continuity of discussions, this section is largely excerpted from our previous report prepared for HKEPD (AS06-435), with the updated conclusions / figures / tables.

not available for 2008). This high correlation gives a strong indication of the reliability for the estimations of both source contribution and emission intensity. When the PM emissions from road transport increases 1000 tons per year, the average annual PM loading in HK would in response increase by about  $1.6 \mu\text{g}/\text{m}^3$ . The correlation coefficient at roadside MK station is not as high as that at general stations, probably because the emission variation near the single station of MK does not reflect the general situation for the entire HK. The significant intercept of the regression line may be due to the mixture of other emissions into the vehicle exhaust factor, the underestimation of road transport emission inventory, or the possible mixture of emissions outside HK territory.

On the contrary, residual oil as another local source category shows a steadily increasing trend from 1998 to 2004, slightly decreasing till 2007, and increasing again in 2008. It increased the most significantly among sources in terms of relative percentages in the eleven years, as shown in **Figure 3.6**. In HK, residual oil is mainly used as marine fuels as ocean going cargo vessels are legally able to use high sulfur fuels with sulfur content up to 4.5%, which contrasts with current requirements and availability of road vehicle fuel in HK which is 0.005% for motor diesel and petrol/gasoline (Gall and Van Rafelghem 2006). At general stations, the annual average contribution was  $0.29 \mu\text{g}/\text{m}^3$  in 2008, nearly tripled from the contribution of  $0.10 \mu\text{g}/\text{m}^3$  in 1998. Figure 3.3 shows that in summertime, TW is the station receiving the largest impact from residual oil, followed by SSP and MK. This is considered as a result of emissions from the Container Terminal to the south and southwest of the three stations. Although the absolute PM contribution is comparatively low from residual oil, some trace level species, such as nickel and vanadium, are associated with adverse health effects. Therefore, the sharp increase of residual oil emissions, while small in terms of  $\text{PM}_{10}$  mass, cannot be ignored and actions should be taken to ameliorate the fuel quality of marine vessels to and from HK to restrain its emission from increasing further.

For the two sources related to sea spray, the contribution from fresh sea salt shows a decreasing trend while the contribution from aged sea salt shows a slightly increasing trend. Therefore, in terms of percentages, more sea salt is being transformed with chloride substituted by sulfate and nitrate in the atmosphere as time evolves in the past decade. For the other natural source of crustal soil, no obvious trend characteristics can be discerned.

Both secondary sulfate and secondary nitrate show a consistent increasing trend from 1998 to 2007 with a slight decrease in 2008. This clearly indicates increasing non-local impact to HK in the past decade. Contribution from the secondary sulfate factor at general stations increased from  $9.1 \mu\text{g}/\text{m}^3$  in 1998 to  $17.1 \mu\text{g}/\text{m}^3$  in 2007, corresponding to a 87% increase. On the other hand, contribution from secondary nitrate at general stations increased from  $7.3 \mu\text{g}/\text{m}^3$  in 1998 to  $9.6 \mu\text{g}/\text{m}^3$  in 2007, corresponding to a 31% increase. Secondary sulfate generally shows a consistent spatial distribution among the general stations and also roadside station, indicating the

main source region is outside HK. In comparison, the contribution of secondary nitrate at YL and MK is consistently higher than that at other five general stations, indicating that apart from non-local contribution, local contribution cannot be neglected. Therefore, treating all secondary nitrates as non-local would overestimate to some extent the non-local relative importance. This however can be offset by the underestimation of non-local aged sea salt as in this study aged sea salt is regarded as a type of local source.

Contributions from the trace metals factor increased significantly from 1998 to 2004, and then decreased gradually to 2008. This may be the result of closing small metal-manufacturing factories in the PRD. Emissions from coal combustion/biomass burning also shared a similar trend. At the general stations, the source contribution increased from  $7.7 \mu\text{g}/\text{m}^3$  in 1998 to  $9.1 \mu\text{g}/\text{m}^3$  in 2004, and then decreased to  $6.3 \mu\text{g}/\text{m}^3$  in 2008. We speculate this pattern might be due to the increasing oil price during that period that led most companies to switch from oil to coal for energy production. In 2008, possibly due to the global economic downturn, the manufacturing activity in the PRD decreased significantly, which led to lower electric power demand and therefore lower emissions from the coal-fired power plants. Both the secondary sulfate and coal combustion/biomass burning factors exhibited a modest decrease in 2008 source contribution estimates. Another possible reason for the decreasing biomass burning contribution is the rapid urbanization process and the more stringent restriction of straw burning which used to be very common in the fall and winter seasons in the rural PRD area.

**Figure 3.7** shows the long-term trend of relative contribution from non-local sources for winter, summer, TS, and all sampling days. Due to the combined effect of local effective controls on vehicle emissions and non-local enhancement of secondary sulfate, nitrate, trace metals and coal combustion/biomass burning, an increased non-local contribution percentage over the past years could be discovered. At general stations, the non-local contribution increased from 58% in 1998 to 72% in 2004, dropped a little to 69% in 2006, and then continued to rise to 75% in 2007. The percentage variation indicates that substantial non-local influence not only offset all the local emission reduction in HK, but also raised the total ambient  $\text{PM}_{10}$  loading by  $6.8 \mu\text{g}/\text{m}^3$ , from 1998 to 2004. In wintertime, the non-local percentage is more significant, increasing from 72% in the winter of 1998-1999 to 86% in the winter of 2007-2008. This high percentages indicate that in the wintertime of 2007-2008, only one seventh of the ambient  $\text{PM}_{10}$  was contributed locally, therefore regional pollution control becomes more and more crucial to the air quality improvement in HK.

### 3.2.3 General Stations vs. Roadside Station

As shown in Table 3.1,  $\text{PM}_{10}$  data are collected at seven monitoring stations, six of which are general stations typically located on rooftops of four- to six-story buildings with unobstructed airflow from most directions. The other station, at MK, is a roadside site which is only a couple of meters above the ground and is designed to reflect the air quality experienced by pedestrians close to vehicular traffic.

As expected, contributions from vehicle exhaust show large differences between these sites with average contributions of  $8.5 \mu\text{g}/\text{m}^3$  at general stations from 1998 to 2008 and  $21.3 \mu\text{g}/\text{m}^3$  at the MK roadside station, averaged from 2001 to 2008. As the abundant emission of  $\text{NO}_x$  is ready to transform to secondary nitrate after reaction with OH and  $\text{NH}_3$ , secondary nitrate also enhanced ( $11.0 \mu\text{g}/\text{m}^3$  in roadside station as compared with  $8.3 \mu\text{g}/\text{m}^3$  at general stations). The large difference of fresh sea salt is more related to the geographical location as MK in the center of Kowloon Peninsula is relatively close to the sea as compared with those stations in the New Territories. For the other six source categories, the contribution differences between general and roadside stations were not very prominent. This demonstrates that non-local sources and local area sources generally show similar impact at both types of stations, although the stations differ substantially in terms of sampling height. At the MK roadside station, non-local sources still accounted for over half of total  $\text{PM}_{10}$  loading on an annual average. In winter this percentage would even rise to 69%, significantly dominating over the local vehicle emissions. This large percentage to some extent constrains the effectiveness of various control strategies focused on vehicular emissions to the improvement of roadside pollution exposed by pedestrians. In addition, the non-local influence at roadside station exhibits a generally increasing trend over the past years. As shown in Figure 3.7(b), the non-local percentage increased from 54% in 2001 to 61% in 2008. In winter, the non-local contribution increased from 65% in the winter of 2001-2002, again increased to 73% in the winter of 2005-2006, decreased to 67% in the winter of 2006-2007, and then rebounded to 77% in the winter of 2007-2008. Wintertime rate of increase was higher than the annual average rate of increase, suggesting the contribution from non-local sources increased significantly in the past decade.

### 3.2.4 Comparison with Results in Previous Studies

As a follow-up and refinement of our previous studies published in Yuan *et al.* (2006) and Huang *et al.* (2009) and prepared for HKEPD (Lau *et al.* 2010), it is interesting to compare source apportionment findings. Yuan *et al.* (2006) reported that from 1998 to 2002, secondary sulfate and vehicle emissions made the largest contribution (25% for each) to  $\text{PM}_{10}$  in Hong Kong, followed by secondary nitrate (12%) on an annual average. For other sources, the contribution percentages were 10% or less. Huang *et al.* (2009) did not report any average contribution from identified sources. In the report AS06-435 (Lau *et al.* 2010), secondary sulfate contribution was estimated to be 27%, followed by vehicle exhaust (17%) and secondary nitrate (15%). In this study, as shown in Table 3.4, the average contribution from secondary sulfate decreased to 22% from 1998 to 2008, while the average contribution from vehicle emissions also decreased to 15%. These comparisons are based on changes in percent contributions to sum of measured species mass (Lau *et al.* 2010) and QTM mass (this study) and it should be noted that this study is often apportioning more mass to a given source even when the percent contribution decreases between studies (e.g. The general station secondary sulfate SCE was  $10.8 \mu\text{g}/\text{m}^3$  in Lau *et al.* 2010 and  $12.3 \mu\text{g}/\text{m}^3$  in this study). In contrast to the above source that showed lower percent contributions in

this study compared to AS06-435 (Lau *et al.* 2010), the crustal soil/dust contribution increased from 5% ( $1.8 \mu\text{g}/\text{m}^3$ ) in the report AS06-435 to 12% ( $6.7 \mu\text{g}/\text{m}^3$ ) in this study.

*In comparison with previous PMF modeling studies of the HK  $\text{PM}_{10}$  data, the most significant changes for this study were the implementation of refined uncertainty estimates for the species concentrations and the use QTM as the total mass. The remainder of this section provides justification for apportioning QTM mass rather than the sum of measured species mass. As demonstrated earlier in this chapter, for most PMF-resolved factors the unidentified mass (i.e. The difference between the QTM mass and the sum of species mass apportioned to each factor) could be rationalized in terms of species that were not measured and/or possible rotational ambiguities in the PMF solution that affects the species loadings on to the profiles.*

Ambient particulate matter sampling was conducted using high volume samplers with quartz fiber filters and it is possible that quartz fibers can flake off the filters during sampling and handling. This measurement artifact could lead to the reported gravimetric mass being biased low. However, for particle mass loadings representative of conditions in Hong Kong, this artifact might be relatively small. Given that the sum-of-species approach on average accounts for only 62% of the QTM, a significant portion of the particulate matter mass is not being apportioned and this missing mass could bias the source contribution estimates. Therefore, we conducted an analysis to determine whether it would be more appropriate to apportion the sum-of-species mass or the QTM mass. There advantages and disadvantages to both approaches. Both methods do not wholly capture the  $\text{PM}_{10}$  mass present in the atmosphere because there are measurement artifacts especially for ammonium nitrate and the semivolatile organic carbon. The sum-of-species approach has the additional drawback that the source contribution estimates include only the species that are measured. In Yuan *et al.* (2006) and Lau *et al.* (2010), the non-carbonaceous component of organic matter (OM) was explicitly included in the sum-of-species approach by assuming an OM/OC ratio of 1.6. This is an assumption and different from some published OM/OC multipliers for the Hong Kong aerosol (Chen and Yu, 2007). Other potentially important species are not measured and therefore not included in the sum-of-species approach. Examples include Si and the oxygen associated with the crustal elements which are often present as oxides. Apportionment of the gravimetric mass relaxes these limitations but assumes that the gravimetric mass behaves as a conserved species. Therefore, it is susceptible to errors such as water retained by the aerosol of the filter media. Advantages of apportioning QTM include: (i) both the gravimetric mass and all of the species (elements, ions, and carbon) are measured from the same filter so there is consistency between the gravimetric mass measurement artifacts and the species-specific artifacts<sup>8</sup>;

---

<sup>8</sup> In contrast, the Chemical Speciation Network in the United States includes gravimetric mass and elements from a Teflon filter, ions from a Nylon filter, and carbon from a quartz filter. Artifacts for

and (ii) all mass is being apportioned, not just the species that were measured. Furthermore, as demonstrated earlier in this chapter mass closure for the factor loadings can be used as a quality check on the representativeness of the unidentified (unmeasured) mass that is assigned to each factor.

In Hong Kong the particulate matter Air Quality Objectives (AQOs) are assessed using data from the network of PM<sub>10</sub> TEOM monitors operated at 50°C. Thus, from an air quality planning and management perspective, the ideal case would be source apportionment of the aerosol mass as measured by the TEOM. The remainder of this section focuses on comparisons of QTM mass to TEOM mass to evaluate the appropriateness of apportioning QTM as a proxy for the TEOM mass.

For these comparisons the TEOM hourly PM<sub>10</sub> mass data were conditioned by imputing concentration values for cases with missing data for only one or two consecutive hours. These data gaps were filled by linear interpolation of the concentration values bounding the data gaps. Subsequently, only those days with 24 hours of TEOM data (after imputing) were included for the comparison to the QTM data.

For reasons noted below, data from the Yuen Long (YL) station were chosen for analysis. **Figure 3.8a** shows a scattergram of the daily-average TEOM PM<sub>10</sub> mass versus the daily-integrated QTM PM<sub>10</sub> mass measured at YL. Overall, the agreement is quite good and there is no clear evidence of systematic bias in the QTM mass from fibers flaking off the filter. Excluding the four highest data points which are truly extreme values, a reduced major axis regression of TEOM on QTM yields slope  $0.906 \pm 0.015$  (95% C.L.) and intercept  $2.8 \pm 1.0 \mu\text{g}/\text{m}^3$ . The grand mean TEOM and QTM concentrations were 59 and 62  $\mu\text{g}/\text{m}^3$ , respectively, resulting in a TEOM-to-QTM ratio of means of 0.95. These trends suggest that, at least in an average sense, QTM mass is a very good proxy for TEOM mass. **Figure 3.8b** shows the CUSUM time series for the PM<sub>10</sub> mass difference (TEOM minus QTM) at YL. The gross feature is a monotonic decay which means the QTM mass is greater than the TEOM mass, consistent with the above results. CUSUM plots for most other stations exhibited notable slope changes long the time series which suggest changes in the QTM-TEOM mass relationship such as might arise from even small changes in TEOM performance; the stability of the QTM-TEOM mass relationship, as evidenced by the CUSUM plot, was the main criterion for choosing YL for this analysis. The local maxima and minima along the CUSUM time series correspond to seasonal patterns in the TEOM-to-QTM relationship. **Figure 3.8c** is a time series of the sample day specific difference (TEOM minus QTM). In the summer the TEOM concentrations are modestly put persistently higher than the QTM concentrations. In contrast, in the winter the QTM concentrations are often significantly higher than the TEOM concentrations. The net effect is a 3  $\mu\text{g}/\text{m}^3$  difference between the two

---

nitrate and carbon might be different for the gravimetric mass measurements and the species measurements.

measurements but there are larger season-specific differences that cancel to some extent when calculating the grand averages. The reason for the summertime pattern is unknown. The wintertime patterns likely arises from semivolatile compounds (ammonium nitrate and OC) being retained by the quartz filter to varying degrees depending on environmental conditions whereas the TEOM when operated at 50°C is measuring nonvolatile mass.

Overall, the average difference between the daily TEOM-to-QTM mass is only 3%. **Figure 3.8d** shows the cumulative distribution of the TEOM/QTM ratio. The cumulative distribution is a straight line on a probability plot and thus is well represented by a normal distribution. The wintertime pattern of TEOM mass being modestly higher than QTM mass coincides with relatively low PM<sub>10</sub> mass concentrations, whereas the summertime pattern of QTM mass significantly greater than TEOM mass coincides with relatively high PM<sub>10</sub> mass concentrations. Thus, the absolute value of the *relative* difference is nearly the same for both seasons. While there is clearly some seasonal dependence in the relationships between TEOM mass and QTM mass, the QTM data are a reasonably good proxy for the TEOM mass which supports its use in the source apportionment modeling. Furthermore, the QTM mass represents the mass on the filter used for the speciation analysis and therefore is internally consistent as a total variable.

### 3.2.5 Summary

- The PMF model was used for source analysis in HK for PM<sub>10</sub> speciated data from 1998 to 2008. Nine sources, including vehicle exhaust, residual oil, fresh sea salt, aged sea salt, crustal soil/dust, secondary sulfate, secondary nitrate, trace metals, and coal combustion / biomass burning, are identified.
- A secondary sulfate factor that was ~70% ammonium and sulfate had the largest contribution, accounting for 22% of the ambient PM<sub>10</sub> mass at general stations in HK. Contributions from vehicle exhaust, aged sea salt, secondary nitrate, and, crustal soil/dust and coal combustion/biomass burning were comparable, with each around factor contributing 12-15%. Contributions from fresh sea salt and trace metals factors were 3-5% and the residual oil factor contribution was <0.5%.
- By examining the spatial and seasonal variation of each identified source, vehicle exhaust, residual oil, fresh sea salt, and aged sea salt are regarded as local sources due to their modest seasonal variation and relatively large spatial variations, while the other five categories are regarded as non-local sources due to their small spatial variation but large seasonal variations. For non-local sources, wintertime contributions were generally much larger than summertime contributions, and the source contribution estimates generally exhibited a northwest-southeast spatial gradient. In contrast, local sources exhibited wintertime contributions that were comparable to, or even slightly lower than, summertime contributions. There was no clear spatial pattern for local sources.

- Contributions from non-local sources generally greater than the contributions from local sources. Overall, the non-local sources accounted for 67% of total PM<sub>10</sub> at the general stations on an annual-averaged basis with the non-local contribution rising to 77% in winter. At the roadside station, non-local contributions were ~10% lower than at the general stations vehicle exhaust contributions were higher. This trend was observed both overall and for the wintertime and summertime seasons.
- Over the eleven year period from 1998 to 2008, vehicle exhaust contributions show a generally decreasing trend, with small upward rebounds in 2004 and 2006. Compared with the 1998 contributions, vehicle exhaust has decreased by as much as 54% in 2008. This indicates that local controls on vehicle exhaust emissions by the HK environmental authorities have been effective.
- A strong correlation between apportioned annual-average vehicle exhaust contributions and PM emission estimation from road transport was observed, with the correlation coefficient being 0.90. This provides a strong indication of the reliability for the estimations of both source contribution and emission intensity. When the PM emission from road transport decreases 1000 tons per year, the annual-average PM ambient loading in HK would in response decrease about 1.6 µg/m<sup>3</sup>.
- Residual oil shows a steadily increasing trend from 1998 to 2004, and then decreased slightly afterwards. It is the source with the most significant relative increase in contributions. Residual oil exhibits unique spatial distribution with the maximum impact on TW, MK and SSP on average. In wintertime, it poses the largest impact at the TC station. Wind directional dependence analysis (Lau *et al.* 2010) shows that residual oil in TW is mainly associated with south or southwesterly wind, blowing emissions originated from the Container Terminal. In contrast, wintertime, northerly and northwesterly winds can transport emissions from Shekou Port and Urmston Road to TC, making it the station with the largest impact of residual oil emissions in winter. While the residual oil source contributions to PM<sub>10</sub> mass are relatively small, the air toxics metals can be a concern and this emission source category warranted additional attention.
- Non-local contributions at general stations increased from 58% in 1998 to 72% in 2005, modestly decreased to 69% in 2006, and then continued to rise to 75% in 2008. In winter, non-local percentage contributions were more significant, increasing from 72% in the winter of 1998-1999 to 86% in the winter of 2007-2008. This high percentage indicates that in the wintertime of 2007-2008, only one fifth of the ambient PM<sub>10</sub> was contributed locally, therefore regional pollution control has become more and more crucial to the air quality improvement in HK.
- The Mong Kok roadside station is also impacted by increasingly significant amount of non-local pollution, 58% on average in the eight years. In winter this percentage would even rise to 69%, dominating over the local vehicle emission. This large fraction would limit the effectiveness of various control

strategies on vehicular emission to the improvement of roadside pollution exposed by pedestrians. In addition, although local control on emission of  $PM_{10}$  from vehicles is effective, there is no evidence that the emission of gaseous pollutants from vehicles is also reduced to the same degree. Further control efforts in local vehicles are still needed.

- Gravimetric mass from the quartz filter (QTM) was confirmed to be a generally reliable surrogate for total  $PM_{10}$  mass measured by the TEOM. This relationship is important because the TEOM data are used for comparison to the AQOs. In contrast to previous studies that apportioned the sum of measured species, the apportionment on QTM yielded considerably higher absolute contribution estimates because more mass is being apportioned. Crustal soil contributions increased most dramatically with this change in modeling strategy, from 5% in previous studies to 12% in the current study.

### 3.3 Source Apportionment by APCA and Unmix

#### 3.3.1 Implementation of APCA

The same eighteen species used for the PMF modeling were used in the APCA modeling. **Figure 3.9** shows the scree plot for the eighteen latent roots. Using Kaiser's criterion, latent roots with eigenvalues greater than  $\sim 1$  are retained which in this case would be the first five components ( $\lambda = 8.8$  to  $0.96$ ). The scree plot suggests that up to an additional two components might be justified if the species profiles and source contribution estimates can be physically interpreted. However, to be consistent with the PMF analysis, solutions of up to nine components were investigated. An optimal solution was selected by reviewing the regression results, G-space plots (scatter plots of the SCEs for one factor against another factor), and physical interpretation of the factor (source) profiles and explained mass profiles (the distribution of each species across the factors).

Preliminary modeling revealed that a factor with cadmium but no other trace element species was consistently resolved with as few as five factors retained in the solution. Based on G-space plots obtained for the Unmix modeling (described below), the 36 samples with cadmium concentrations greater than  $0.01 \mu\text{g}/\text{m}^3$  were eliminated from the data set. These data are identified in **Appendix A**. The elimination of the high cadmium concentrations suppressed the resolution of a cadmium-dominated factor.

**Table 3.5** shows the source contribution estimates for 4-to-9 factor solutions, mapped onto the PMF-resolved source categories where appropriate.<sup>9</sup> Results are presented for both constrained (Table 3.5a) and unconstrained (Table 3.5b). For the

---

<sup>9</sup> To facilitate the comparisons of results across receptor modeling methods, the source category names assigned to factors in the PMF modeling of section 3.2 are retained in this section. These assignments are made by matching to the tracer species identified for each PMF-resolved factor.

unconstrained regression, the constant term should tend to zero for a good APCA solution. Table 3.5b shows that a large constant was obtained for the four-factor solution and thus it is rejected; this is consistent with the scree plot demonstrating that at least five factors should be retained. For the 5-to-9 factor solutions the constant was smallest for six and nine factors. Unlike PMF it is possible to obtain large negative values for the factor loadings (source profiles) and factor scores (source contribution estimates). The presence of negative values is mathematically possible but not physically realistic, and this becomes an important criterion for accepting or rejecting certain solutions. The eight- and nine-factor solutions exhibit unacceptably large negative source contribution estimates for one or more factors and thus they are rejected. The seven factor solution includes Residual Oil, Crustal/Soil, Vehicular Exhaust, Secondary Nitrate and Secondary Sulfate distinct factors; Coal Combustion and Biomass Burning source categories lumped into a single factor; and Aged Sea Salt and Fresh Sea Salt also lumped into a single factor. Within the Secondary Sulfate factor, sulfate ion accounts for 37% of the mass assigned to that factor and chloride ion has a negative loading corresponding to 40% of the total observed chloride. The Secondary Nitrate factor contains 38% nitrate ion by mass along with other trace metals including Fe, K, Pb, Cl, etc. With six factors, nitrate ion is smeared over the resolved six factors. The Secondary Sulfate factor still has negative chloride loading corresponding to 30% of the observed chloride, but now contains ~45% sulfate. With five factor solution, most of the sulfate is apportioned to the Coal Combustion and Biomass Burning factor. Further, the Crustal/Soil factor is smeared with Biomass Burning. Note that even with higher number of factors (7-9), the Coal Combustion and Biomass Burning source categories could not be separately resolved. Similarly, the Aged Sea Salt and Fresh Sea Salt factors could not be separately resolved without allowing high negative mass for chloride in the Aged Sea Salt factor. The six- and seven-factor solutions seem most plausible and ultimately *the six factor solution was selected*. This solution provides direct comparisons to the six-factor UMNIX solution described below. Additional features and limitations of the six-factor APCA solution are described in Section 3.

### 3.3.2 Implementation of Unmix

Receptor modeling was also performed using EPA Unmix 6.0. Analysis steps and some considerations / justifications for data processing are detailed below.

1. Import input file containing data for all sites arranged with pseudo-dates and species mass concentrations (OC, EC, TC,  $\text{NO}_3^-$ ,  $\text{SO}_4^{2-}$ ,  $\text{NH}_4^+$ , Al, Ca, Cd, Cl, Fe, K, Mg, Mn, Na, Ni, Pb, S,  $\text{Cl}^-$ , V, Zn) in addition to total mass variables (quartz gravimetric mass (QTM) and sum-of-species mass (STM)) into Unmix.
2. In the *Data Processing* window, the species to be included/excluded can be selected. TC was excluded in favor of its two components; EC and OC. Sulfur was excluded to avoid redundancy with sulfate. Trial runs did not reveal any significant difference when either Cl or  $\text{Cl}^-$  was used in the model,

- thus  $\text{Cl}^-$  was excluded. STM was excluded since QTM will be used as the total mass variable.
3. In the main window, select Quartz filter mass (QTM) and mark it as the total species (mass) variable. Correspondingly, no species were set as tracer species nor were used to normalize the source contributions.
  4. To avoid Cd dominating a single factor, the same 36 samples eliminated from APCA modeling (Appendix A) were eliminated from the Unmix data set and this suppressed the resolution of a Cd-dominated factor.
  5. Using *Select Initial Species* function to select the species with the largest loadings in the varimax factor analysis of the data to identify a subset which can give a reasonable three-to-five factor solution. Al, Ca, Cd, Mg, Na,  $\text{NH}_4$ , Ni, Pb,  $\text{SO}_4$  and V were selected.
  6. The *Suggest Additional Species* function with SAFER algorithm was used (in lieu of the *Auto Unmix* option) to include the remaining species (OC, EC,  $\text{NO}_3^-$ , Cl, Fe, K, Mn,  $\text{Cl}^-$ , Zn). Unmix was unable to find convergent solutions for eight and nine factors. Solutions were obtained for five, six, and seven factors.
  7. In *Diagnostic Plots* function, source profiles/contribution and G-space plots were obtained. All three solutions (five to seven factors) had profiles with negative concentrations apportioned to certain species, with chloride being the most troublesome, especially in the Secondary Sulfate factor. With seven factors, the following sources were identified: Crustal Soil, Residual Oil, Secondary Nitrate, Secondary Sulfate, Vehicular Exhaust, Sea Salt (Aged and Fresh combined into one factor) and a single factor combining Coal Combustion and Biomass Burning. However, the Secondary Nitrate factor accounted for only 21% of the total nitrate and also included various trace metals like Fe, K, Pb etc. The six factor solution had the Secondary Nitrate smeared over all six sources. Further, the five factor solution had the Secondary Sulfate smeared over all five factors.
  8. Based on the above considerations, *the six factor solution was determined to be the best possible solution*. For this solution, an examination of G-space plots suggested no adjustments to the longitudinal spread or transverse spread were warranted.

### 3.3.3 Factor Source Profiles and Explained Mass Profiles

**Figure 3.10** presents the explained mass profiles and **Figure 3.11** presents the source profiles for the six factor APCA solution. Explained mass profiles summarize the distribution of each species across the factors. For each species, the sum of the explained mass fraction across all the factors sums to unity. Source profiles show the mass fraction of each species within each factor. The eighteen species fractions within a given factor sum to unity. **Table 3.6** presents the source profiles derived by APCA, expressed as mass percentages. The explained mass profiles are quite similar to the PMF-derived profiles presented in Figure 3.2. The key differences are: (1) nitrate ion is loaded primarily onto the Coal Combustion & Biomass Burning factor and the Vehicle Exhaust factor in APCA whereas it was resolved as a separate

factor using PMF; (2) Sea Salt was resolved as a single factor in APCA but as two factors – Fresh Sea Salt and Aged Sea Salt – using PMF; (3) Coal Combustion and Biomass Burning were resolved as separate factors using PMF but are lumped into a single factor in the APCA solution; and (4) a significant negative mass concentration of chloride ion (relative to the study-average observed chloride average concentration) was apportioned to the Secondary Sulfate factor in APCA.

Qualitatively, the source profiles of Figure 3.10 and Table 3.6 are similar the profiles presented in Table 3.2 after taking into consideration the lumping of factors in APCA that were separately resolved using PMF. A quantitative comparison of the APCA and PMF-derived source profiles is discussed in section 3.5.

Chloride ion is poorly apportioned by APCA with three of the six factors having negative chloride ion mass fractions (Table 3.6). Chloride ion is -1.5% in the Residual Oil factor and -0.5% in the Crustal/Soil factor. These loadings have relatively little impact on the chloride ion explained mass because the source contribution estimates from these two categories are relatively small. In contrast, the loading of -2% chloride ion on the Secondary Sulfate factor does significantly affect the distribution chloride ion explained mass distribution across the factors (Figure 3.10) with about a -30% contribution from this factor. This negative contribution is balanced by the explained mass for chloride ion in the Sea Salt factor being greater than 100%. The anomalous behavior for modeled chloride ion is not surprising since it is a reactive species and can be displaced from sea salt by forming hydrochloric acid. The importance of this ion displacement likely has high day-to-day variability which would add significant noise to the source profile representing truly fresh sea salt emissions. Sodium ion is loaded at -1.4% on the Combustion & Biomass Burning factor and ammonium ion is loaded at -2.2% onto the Sea Salt factor. In both cases, the resulting contribution to the explained mass is about -20%. OM loaded at -2.0% onto the Sea Salt factor but represents a negligible contribution to the OM explained mass. All remaining negative factor loadings were at mass percentages less than or equal to 0.1% with no practical consequence to explaining the distribution of species mass across the factors.

The APCA six factor solution did not resolve a Secondary Nitrate factor and nitrate ion loaded onto most factors at higher mass percentages than observed in the PMF solution. Ammonium nitrate is volatile and subject to measurement artifacts. Assuming significant nitrate ion is present as ammonium nitrate, its volatility and the artifact behavior could affect its covariance with other species and lead to the observed smearing across factors.

**Figure 3.12** presents the explained mass profiles and **Figure 3.13** presents the source profiles for the six factor Unmix solution. **Table 3.7** shows the source profiles obtained from Unmix expressed as mass percentages. The Unmix-derived explained mass profiles are quite similar to the APCA-derived explained mass profiles and are similarly mapped onto the source category names assigned to the PMF profiles.

Many of the negative loadings reported above for the APCA source profiles are also observed for the Unmix source profiles. For example, there is a negative loading of chloride ion onto the Secondary Sulfate and Residual Oil factors. The Unmix-derived Sea Salt factor has large, negative loadings of ammonium ion and sulphate ion which is different behavior from the APCA-derived Sea Salt factor.

In summary, six factor solutions resolved using APCA and Unmix show good agreement, and the explained mass profiles and source profiles are generally consistent with the results for PMF.

### 3.3.4 Comparison of APCA and Unmix Source Contribution Estimates

**Table 3.8** presents the study-average source contribution estimates (SCE) for the APCA and Unmix six factor solutions. The sample-specific SCEs for the six matched APCA- and Unmix-resolved sources are compared in **Figure 3.14**. Negative SCE values are excluded from these plots. SCE values for APCA and Unmix are highly correlated for the Sea Salt, Residual Oil and Crustal/Soil factors. For the Sea Salt and Crustal/Soil factors, however, APCA-derived contributions are ~1.5-2 times higher than the Unmix-derived contributions. In these cases, higher SCE are observed for APCA compared to Unmix because sulfate ion (both factors) and OM (Crustal/Soil factor) favorably loads onto the APCA-resolved factors compared to the corresponding Unmix-resolved factors. Vehicular Exhaust, Secondary Sulfate and the Coal Combustion/Biomass Burning SCEs for APCA and Unmix also show reasonably high correlations (Figure 3.14) and similar annual average SCEs (Table 3.8) with excellent agreement for Vehicle Exhaust. While there would be some differences in the trends extracted from the APCA and Unmix results, Figure 3.14 demonstrates that the differences are largely systematic and thus the trends would be qualitatively, and in most cases even semi-quantitatively, quite similar.

### 3.3.5 Key Trends for the Source Contributions Estimates

Section 3.2 included a detailed analysis of the PMF-modeled source contribution estimates including seasonal variations, annual trends, differences between general and roadside stations, and relationships with surface winds. Given the generally good agreement between the APCA-, Unmix-, and PMF-modeled source profiles and explained mass profiles, the detailed analyses of section 3.2 are not fully replicated here. Key trends are highlighted in this section to reinforce the findings from the PMF modeling.

Source contribution estimates for each factor were calculated with negative mass contributions zeroed out for the calculation of annual and seasonal averages. The six general stations with relatively complete data sets (CW, KT, SSP, TC, TW and YL) were used to compute the general station trends. MK site was used as the roadside station for comparison. **Table 3.9** shows the annual and seasonal SCEs averaged

over the entire data set. Table 3.9(a) presents results for APCA while Unmix results are in Table 3.9(b). Wintertime contributions from the Coal Combustion & Biomass Burning factor are about 4-6 times higher than the summer contributions. The Secondary Sulphate and Crustal/Soil factor also shows an increase by almost 2-3 times from summer to winter. These trends hold true for both the general stations and roadside station (MK). In comparison, Residual Oil and Vehicular Exhaust factors exhibit relatively little seasonal variability.

**Figure 3.15** shows the APCA and Unmix source contribution estimate differences between the general stations and roadside station. For a given factor, study-average APCA source contribution estimates (“Annual” values in Table 3.9) at the roadside station compared to the general stations agreed within 20% except for the Vehicle Exhaust factor which was 2.5 times greater at the roadside station. For Unmix, study-average source contribution estimates at the roadside station compared to the general stations agreed even better - less than 10% for all factors except Sea Salt (20%, but only  $0.4 \mu\text{g}/\text{m}^3$  difference) and Vehicular Exhaust which was 2.7 times higher at the roadside station.

**Figure 3.16** shows the trends in annual average source contribution estimates for each of the six factors resolved by Unmix. These patterns track rather well with the PMF results presented in section 3.2. Only the vehicle exhaust factor shows a dramatic difference between the general stations and roadside stations which is consistent with this source be local and the other sources being predominantly non-local. The relatively large SCEs for secondary sulfate (Figure 3.16a) arise from secondary nitrate also loading predominantly onto this factor. Contributions from these predominantly non-local sources have increased over the decade and especially in the second half of that period. Increasing contributions from these sources to a large extent offset the reductions in vehicle exhaust contributions (Figure 3.16b) over the decade. Contributions from coal/biomass combustion (Figure 3.16c) and crustal/soil (Figure 3.16d) have remained relatively constant over the decade. Sea salt contributions (Figure 3.16e) have decreased; the reason for this trend is not clear. Finally, residual oil combustion contributions (Figure 3.16e) increased rather dramatically early in the decade and then remained relatively constant thereafter.

**Figure 3.17** (APCA) and **Figure 3.18** (Unmix) show the relative changes in the source contributions over the duration of the study using the 1998 values as a baseline for the general stations and 2001 as a baseline for the roadside station. The trends derived from APCA, Unmix and PMF are quite similar.<sup>10</sup> The contribution from Residual Oil has increased by almost 250% compared to 1998 levels. Both PCA and

---

<sup>10</sup> Figure 3.6 started the roadside station time series with 2001 because the MK roadside station was moved in 2000. In this section, most annual trends for the roadside station are presented starting with 1998 to show the potential impact of the site relocation. Some roadside station trends are presented starting with 2001 because these time series focuses on percentage changes with respect to a base year and are confounded by the site being moved.

Unmix results show that Vehicle Exhaust contributions at the general stations have decreased by more than 50% of 1998 levels. At the roadside station, a decrease greater than 50% was observed over the period 2001-2008 with lower levels prior to 2001 (not shown) likely reflecting the different station environment prior to the site being moved in 2000. The contributions from Coal Combustion & Biomass Burning and Sea Salt have mainly remained nearly constant with slight decrease towards the end of the study period. The Crustal/Soil contribution has consistently stayed 50% above the 1998 levels. Secondary Sulfate however, started increasing in 2002 with a nearly 75% increase above 1998 values reached by 2008.

**Figure 3.19** (APCA) and **Figure 3.20** (Unmix) show the long term annual and seasonal trends of the relative contributions from non-local sources at the general stations (a) and roadside station (b). Again, the two models yield similar trends and are in good agreement with the trends from PMF. The factors assigned to non-local sources include Coal Combustion and Biomass Burning, Crustal/Soil and Secondary Sulfate. Perhaps sea salt can also be regarded as a non-local source but it does not show similar seasonal trends (Table 3.6). Both APCA and Unmix show higher non-local contributions at the general sites than the roadside station because the roadside station has higher Vehicle Exhaust source contribution estimates. APCA and Unmix both show ~55% non-local contribution in 1998 (general stations, annual-average). However, APCA shows a non-local contribution of ~70% in 2008 whereas Unmix estimates the 2008 non-local contribution to be 80%.

### 3.3.6 Summary and Recommendations

APCA and Unmix yielded generally consistent results. Six factors were optimally resolved in both cases, although there are several deficiencies in these modeled solutions. Negative factor loadings and negative source contribution estimates were observed for some species, especially chloride ion. Interestingly, chloride is likely the most chemically reactive of the species included in the modeling with its replacement in sea salt likely exhibiting high day-to-day variability. Nitrate ion was typically smeared over several factors. Nitrate in the form of ammonium nitrate is subject to significant volatilization and sampling artifacts which could increase substantial noise into its covariance with other species.

Recommendations specifically concerning the multivariate modeling by APCA and Unmix include: (1) Consider imposing a threshold on the maximum percentage of below detection limit values that can be tolerated; and (2) impose a non-negativity constraint to the principal components following the method presented by Rachdawong and Christensen (1997).

The deficiencies observed in this study with APCA and Unmix are consistent with the motivation for developing PMF. In addition to negative loadings and negative source contribution estimates, the exclusion of uncertainties places equal weight on all data regardless of their signal-to-noise. Despite these limitations, the key trends

extracted from the APCA and Unmix results are consistent with the PMF modeling. This provides additional confidence in these patterns being real and not merely modeling artifacts.

### 3.4 Source Apportionment by CMB

A primary focus of this project was the application of factor analytic tools (PMF, APCA, Unmix) to apportion the Hong Kong PM<sub>10</sub> aerosol. Source apportionment was also being performed using the chemical mass balance (CMB) method although it has been appreciated from the outset of this project that there would likely be significant limitations that might constrain or even compromise its applicability. First and foremost, representative source profiles are not available for the relevant sources.

#### 3.4.1 Source Profiles – Selection and Modification

The objective was to perform CMB modeling for the emission source categories significantly contributing to PM<sub>10</sub> mass concentration burdens. The relevant source categories were inferred from the receptor modeling results. Candidate source profiles were obtained from the literature and also the USEPA's SPECIATE database, focusing exclusively on profiles reported for the PM<sub>10</sub> size range. Source profiles are reported as the species mass fractions – the mass of each species divided by the total (typically gravimetric) mass. The PMF source apportionment modeling performed for this project did not use the gravimetric mass from the PM<sub>10</sub> high-volume sample quartz fiber filters but rather used the summed species mass. Therefore, to be consistent with the PMF modeling, the “base case” CMB modeling was performed using the sum-of-species mass as the total mass variable. In principle the source profiles as published could be used but this would exclude a key performance metric for CMB modeling – the sample-specific extent of mass closure between the modeled and measured total mass. Therefore, it was necessary to rescale the source profiles so the mass fractions summed over only those species used in the modeling (hereafter termed the “summed species mass”) was unity. The uncertainty assigned to each species in these rescaled profiles was recalculated using conventional uncertainty propagation rules. OC was used instead of OM to be consistent with the source profiles.

Based on the receptor modeling and existing knowledge about emission source categories that might significantly contribute to the PM<sub>10</sub> mass in Hong Kong, the following profiles were selected.

- *Vehicle exhaust:* A representative vehicle emission source profile would consider the vehicle fleet mix (e.g. gasoline versus diesel), emission control technologies, and operating conditions. These factors have likely changed over the ten year of this data set, and there is little information to support the selection of a potentially representative source profile. Light Duty Vehicle - Diesel Composite profile 3210410 in the SPECIATE database had an EC/OC

ratio (1.8) with no uncertainties reported. Another Light Duty Vehicle – Diesel Composite profile (3210210) had a significantly higher EC/OC ratio (2.7) with uncertainties reported. The profile with the lower EC/OC ratio (3210410) seemed more reasonable – EC/OC ratios for the roadside station ambient data averaged 0.7 with a ratio of 1.8 at the 98<sup>th</sup> percentile – and was chosen for the base case modeling but the profile with the higher EC/OC ratio will be used in future sensitivity studies. The uncertainties from profile 3210210 were used to estimate uncertainties for profile 3210410 which was used in the CMB modeling.

- *Soil (paved road dust):* Ho *et al.* (2003) published three source profiles for PM<sub>10</sub> fugitive dust samples collected in Hong Kong. These samples were termed Country Park Soil, Urban Soil and Paved Road Dust. These profiles could not be simultaneously used in the CMB modeling due to collinearity issues. Paved road dust was assumed to be the most appropriate soil profile for the Hong Kong data sets and was used for base case CMB modeling. However, sensitivity studies were performed by using each of these profiles in separate CMB runs.
- *Sea salt:* Pytkowicz and Kester (1971) published profiles for fresh sea salt and also sea salt with various extents of aging. PMF modeling resolved factors that were assigned to fresh and aged sea salt. However, only one sea salt factor could be included in the CMB modeling due to collinearity issues and the fresh sea salt factor was used. It is noted that the aged sea salt factor in the PMF modeling is significantly depleted in chloride (as expected) and significantly enriched in nitrate whereas the Pytkowicz and Kester aged sea salt profiles are all significantly enriched in sulfate but not nitrate.
- *Agricultural burning:* Agriculture vegetation burning profile 3625 from the SPECIATE database was used. Several SPECIATE profiles, including a profile specifically for rice straw, were tested but this generic agriculture vegetative burning profile yielded the best model performance.
- *Residual oil combustion:* Profile 1350110 from the SPECIATE database was selected. This particular profile was chosen from the various residual oil combustion profiles in the SPECIATE database because it has a Ni/V ratio closest to the ambient data. This might seem a poor justification for profile selection because it infers that residual oil combustion is the dominant source of both nickel and vanadium. However, these two species were highly correlated for the vast majority of ambient data and thus it is likely that a single source dominates their emissions.
- *Secondary sulfate:* This profile was generated using the stoichiometry of ammonium sulfate. It does not account for the Hong Kong sulfate being only partially neutralized.
- *Secondary nitrate:* This profile was generated using the stoichiometry of ammonium nitrate.

A coal combustion profile was not included in the CMB modeling. While this source category was assigned to one of the PMF-resolved profiles it is not clear

whether the source is coal combustion or metals processing operations which could have substantially different source profiles. Also, secondary organic aerosol (SOA) was not resolved as a source but there was high loading of OM (~50%) onto the PMF-resolved secondary nitrate factor. The absence of SOA in the CMB modeling is likely a significant deficiency.

**Figure 3.21** and **Table 3.10** summarize the rescaled source profiles used in the base case CMB modeling.

### 3.4.2 CMB Modeling and Diagnostics

USEPA CMB 8.2 was run for each of the 4,852 ambient samples in the data set using uncertainties generated from the error structures described in Section 4 of this report. For the base case modeling conducted using the above source profiles, SCEs were obtained for 4,652 of the samples; runs for the remaining 200 samples did not converge due to collinearity issues.

#### Mass Closure

**Figure 3.22** shows the cumulative distribution of mass balance closure expressed as the percentage explained mass, i.e. (modeled mass) / (measured sum-of-species mass)  $\times 100$ . A criterion of mass closure in the range 80-120% was adopted for the run to be used in subsequent analyses. Three samples were below the 80% threshold and 807 samples (17%) were above the 120% threshold. **Figure 3.23** shows that poor mass closure tends to occur for samples with relatively low measured sum-of-species mass. Thus, results presented later in this section are skewed towards samples with relatively higher concentrations. Rejected samples usually corresponded to the months of April to August and late October to December and occurred over the entire ten year study period. Most July and August samples were rejected. At low concentrations the modeling results split into two groups with one group of samples exceeding the 120% mass closure threshold and the other group having modeled-to-measured mass ratios less than 100%. With increasing measured concentration, the two groups converge at about 110% explained mass. Data below  $20 \mu\text{g}/\text{m}^3$  were examined to gain insights into this behavior. Source contribution estimates (SCEs) were plotted against the measured mass and it was observed that negative SCEs were relatively common for samples  $< 20 \mu\text{g}/\text{m}^3$  and exceeding the 120% mass closure threshold. This provides further evidence that the modeled solutions are poor for these samples.

To check the sensitivity of the CMB results to the soil profile used, modeling was performed with each of the three soil profiles (county park soil, urban soil and paved road dust) separately. **Figure 3.24** shows the cumulative distribution of the percentage of explained mass. The model experienced collinearity issues more often with the county park soil than the other two soil profiles, leading to nearly 10% fewer samples with converged results. Furthermore, the predicted mass concentrations were greater than 120% for nearly three-fourths of samples. Finally, with country

park soil profile, the study average soil SCE was nearly twice the SCE obtained using the urban soil or paved road dust profiles. While both urban soil and paved road dust provided similar source contribution estimates, fewer converged solutions were obtained using the urban soil profile and thus the paved road dust profile was selected for the base case. There is considerable concern about the representativeness of these profiles for the Hong Kong data set. **Figure 3.25** shows the relative distribution of Al, Ca and Fe in the ambient samples (open circles), Ho *et al.* 2003 profiles (red circles), and the PMF-resolved factors (green circles). The paved road dust profile lies at the edge of the cluster of ambient data while the county park soil and urban soil profiles are far removed from the ambient data with much higher aluminum levels. This supports the selection of the paved road dust sample for the CMB modeling. However, it is also possible that differences in analytical methods used for the Hong Kong ambient data and Ho *et al.* (2003) source profiles could cause bias, especially for light elements such as aluminum.

#### Gravimetric Mass as the Total Mass Variable

Gravimetric mass was not used in the source apportionment modeling because, consistent with prior studies on the Hong Kong PM<sub>10</sub> data, it was presumed that the gravimetric mass data is possibly biased due to artifacts such as filter fiber loss during sampling and handling. While this might indeed be a concern, the bias might be small for PM<sub>10</sub> samples given the relatively high mass loadings and thus we performed CMB modeling with the gravimetric mass (instead of the measured sum-of-species mass) as the total mass variable. For this modeling the original source profiles were used rather than the rescaled profiles. Other species present in the profile could potentially cause significant mass to be apportioned to the profile. The rescaling was most dramatic for the paved road dust and residual oil combustion profiles with the sum of the species measured in this study accounted for only 44% of the paved road dust sample mass and 76% of the residual oil combustion mass. CMB modeling for this case removed the overestimate previously observed at low concentration but the modeled mass was on average about 80% of the gravimetric mass. The species measured in this study accounted for 44% of the paved road dust profile but only 23-24% of the urban soil and country soil profiles; thus a sensitivity study was performed by doubling the road dust SCE for each sample. **Figure 3.26** shows that for this case, most of the data does fall within the mass closure thresholds of  $100 \pm 20\%$ . These analyses using the gravimetric as the total mass variable underscore that it might be more appropriate than using the sum-of-species mass. However, the subjective selection of the source profiles renders the CMB modeling results semi-quantitative at best and therefore results are presented for modeling with the scaled profiles and measured sum-of-species mass for consistency with the other source apportionments performed for this project.

CMB modeling was also conducted using a profile to account for SOA which was 62.5% OC (consistent with an OM/OC ratio of 1.6). In this case, the modeled mass significantly overestimated the gravimetric mass for virtually all samples.

## G-Space Plots

G-space plots of the sample-specific source contribution estimates were constructed for each pair of sources (**Figure 3.27**). If the data uniformly fill the plot space, this would imply that there is no correlation between source contribution estimates. A high correlation might be observed for sources that are associated in some manner, such as having a similar emission source geographic region. If the data are mostly along the two axes then significant impacts from both sources are not observed for any given sample. This might arise from the sources being in very different geographic regions, or could result from modeling error. This pattern was commonly observed for residual oil with the high residual oil SCEs all occurring during May-September. The G-space plot for secondary nitrate versus secondary sulfate shows a distinct edge; when sulfate is low the nitrate is never high whereas when sulfate is high the nitrate can be low or high. This pattern is observed for both the CMB-modeled SCEs and the measured nitrate and sulfate data. This might reflect day-specific variations in the partitioning between particulate ammonium nitrate and gas phase ammonia and nitric acid, both in the atmosphere and as a measurement artifact. Paved road dust and secondary sulfate are highly correlated. The reason for this pattern remains unknown.

### **3.4.3 CMB Results for the Base Case Modeling**

**Table 3.11** summarizes the study-average source contribution estimates for the general and roadside stations. Results are presented for both the CMB modeling and the PMF modeling. For this comparison the source category names used for the PMF modeling were adopted. PMF resolved separate fresh and aged sea salt factors; the SCEs have been summed for the entries in Table 3.11. Base case CMB modeling had 17% of samples with poor mass balance closure and thus it was necessary to remove these samples from the PMF modeling results before calculating mean SCEs. Vehicle exhaust SCEs show good agreement between the two models although this might be influenced by the method used to select the vehicle profile for the CMB modeling. Residual oil combustion SCEs are much lower for CMB compared to PMF. There are several possible explanations for these differences: the CMB profile might be unrepresentative; PMF might have artificially loaded major species onto the factor assigned to residual oil combustion due to the ubiquitous nature of the major species in the ambient data; and CMB samples failing the mass closure criterion were mainly in the summer months when residual oil combustion impacts tend to be highest. These issues point to the need for better resolution of residual oil combustion impacts on ambient PM levels in Hong Kong. Sea salt SCEs are about two times higher for the PMF modeling. This might reflect the use of a fresh sea salt profile in the CMB modeling. If only the PMF modeled fresh sea salt factor is used, the respective SCEs are 50-60% of the CMB modeled sea salt SCEs. Thus, the CMB modeled sea salt SCEs are bracketed by the PMF-modeled fresh and total sea salt SCEs. Crustal soil/dust SCEs are five times higher for the CMB results compared to the PMF results and this is perhaps the most striking

difference between the apportionments. As discussed in section 3.2 the subsequent use of QTM as a total variable significantly increased the crustal soil/dust SCEs and are likely more realistic estimates. Secondary sulfate SCEs are 10% higher and secondary nitrate SCEs are 40-50% lower for CMB compared to PMF. The difference for nitrate arises from the high loading of carbonaceous aerosol onto the PMF-modeled secondary nitrate factor. Biomass burning is 30-40% lower for CMB compared to PMF. **Figure 3.28** shows the relative contributions of each source category to the CMB modeled mass. Again, the most striking feature is the relatively high contribution from soil compared to modeling results from PMF.

The remainder of this section compares time series for the CMB-modeled annual average SCEs for the roadside and general stations (**Figure 3.29**). As discussed in section 3.2, the roadside station trends for 1998-2000 might be unrepresentative because the site was moved in 2001 and thus we focus roadside station trends for the post-2000 period. The *light duty vehicle - diesel composite exhaust* (vehicle exhaust) SCE decreases over the ten year time period, both at general stations and roadside station. At the roadside station the contribution of vehicle exhaust to the annual average sum-of-species  $PM_{10}$  decreased from 36% in 2001 to 27% in 2008. At the general stations, the SCE contribution to annual average sum-of-species  $PM_{10}$  decreased steadily over the ten years from 28% in 1998 to 15% in 2008. *Residual oil combustion* shows modestly decreasing SCE at the roadside station after 2001 and increasing SCE at the general stations from 1998-2002. The SCEs are very small ( $\sim 0.1 \mu\text{g}/\text{m}^3$ ) but the robustness of the SCEs are not clear – for reasons described above – and residual oil emissions might be important from an air toxics perspective regardless of the  $PM_{2.5}$  mass contributions. *Sea salt* SCEs are modestly higher at the roadside station compared to the general stations. This might arise from the proximity of the Mong Kok roadside station to the waterfront whereas some of the general stations are inland and might pull down the general stations average SCE for sea salt. The CMB-modeled sea salt SCE at the general stations is nearly constant over the study period; in contrast, the PMF-modeled fresh sea salt contribution decreased over the study period. *Paved road dust* SCEs are nearly identical at the roadside and general stations and show perhaps a modest increase over the study period. Generally similar trends were observed for the PMF modeling although the contributions were much lower. *Secondary sulfate* SCEs are nearly identical at the roadside and general stations and increased from 8 to  $12 \mu\text{g}/\text{m}^3$  over the ten year study period. The concentrations estimated at roadside and general stations are similar, but it represents a higher fraction of the annual source contribution at the general station due to lower  $PM_{10}$  levels at the general stations compared to the roadside station. *Secondary nitrate* SCEs are modestly higher at the roadside station compared to the general stations and have increased over the study period. *Agricultural vegetative burning* (biomass burning) SCEs were nearly identical at the roadside and general stations and steadily decreased by about 50% over the study period; this decrease was not evident in the PMF modeling results.

### 3.4.4 Summary and Recommendations

CMB modeling was performed to apportion  $PM_{10}$  measured species to seven source categories. We have significant concerns about the representativeness of the selected source profiles and the inclusion of all relevant sources, especially SOA. That said, the modeling results show generally favorable agreement with the other source apportionments performed for this project. There were significant differences between PMF and CMB for the SCEs assigned to residual oil combustion and soil. Residual oil SCEs were much lower for CMB compared to PMF while soil SCEs were much higher for CMB compared to PMF. It is not clear which of the models is providing a better estimate of the impacts from these emissions sources and future work should include a refined characterization of residual oil combustion and soil/dust impacts. Finally, while throughout this project SCEs have been presented using the sum of the measured species which loads onto each factor (or source profile), it would be instructive to use the gravimetric mass measured on the quartz fiber filter samples and include this parameter as a total variable in the modeling. The use of these gravimetric mass data have been discounted due to potential biases from the loss of fibers during handling, but given the high mass loadings onto the samples the bias might be small. This might improve the reconciliation between the CMB results and factor analytic results (PMF, APCA, and Unmix).

### 3.5 Comparison of Source Contribution Estimates by PMF, APCA, Unmix and CMB

This section presents a unified comparison of the  $PM_{10}$  SCEs obtained using PMF, APCA, Unmix, and CMB. In each case the quartz filter gravimetric mass was used as the total variable. The reported SCEs correspond to the apportionment of quartz filter gravimetric mass and not the sum of species in the resolved source profiles. PMF modeling was performed using the refined error structures reported in Section 2.2. In contrast to the PMF modeling reported in Section 3.2 an eight factor solution was deemed optimum with the difference being that the factor assigned to zinc smelting was not resolved.

#### 3.5.1 PMF, APCA, and Unmix

PMF modeling results were censored to exclude the samples listed in Appendix A which were excluded from the APCA and Unmix modeling for reasons stated in Section 3.3.1. The eight factor PMF solution was mapped onto the six factors APCA and Unmix solutions by combining the biomass combustion and coal combustion factors that were separately resolved by PMF, and combining the fresh and aged sea salt factors that were separately resolved by PMF. Furthermore, *a secondary nitrate factor was not resolved using APCA and Unmix.*

**Figure 3.30** shows the modeled SCE distributions from PMF, APCA and Unmix for samples collected at the general stations over the entire study period. All three models showed good agreement for secondary sulfate and vehicle exhaust SCEs.

For biomass/coal combustion the models yielded similar mean and median SCEs although APCA and Unmix exhibited broader SCE distributions including a higher frequency high impact days. PMF-modeled soil/crustal contributions were greater than the APCA and Unmix estimates and also had a higher frequency of high impact days. SCEs were most divergent for sea salt and residual oil. PMF-modeled sea salt SCEs were much greater than the APCA and Unmix modeled SCEs. The trend arises in part because APCA and Unmix did not resolve an aged sea salt factor. Residual oil SCEs were much higher for APCA and Unmix compared to PMF because substantially less  $\text{NH}_4^+$ ,  $\text{NO}_3^-$ ,  $\text{SO}_4^{2-}$ , EC and OC loaded onto the PMF-resolved factor.

### 3.5.2 PMF, APCA, Unmix, and CMB

PMF, APCA and CMB modeling results were censored to exclude the samples for which CMB did not converge and the samples with CMB modeling results outside the  $\pm 20\%$  criterion on mass balance closure. As described in Section 3.4.2 this censoring skews the SCEs to towards samples with relatively higher concentrations because most of the mass balance closure issues occurred for samples with relatively low concentrations. **Figure 3.31** shows the modeled SCE distributions from PMF, APCA, Unmix and CMB for samples collected at the general stations over the entire study period. For both secondary sulfate and vehicle exhaust factors the SCE distributions are very consistent across all four models. CMB predicts much lower contributions from biomass/coal burning and much higher contributions from resuspended soil/crustal material. CMB-modeled sea salt SCEs are closer to the APCA and Unmix estimates than the PMF estimates. This is consistent with the use of a fresh sea salt profile used for CMB modeling because seas salt profiles with varying degrees of aging resulted in collinearity issues when performing the CMB modeling. Finally, the residual oil SCEs for CMB are much more consistent with the PMF modeled results than the APCA and Unmix results. In the latter cases much more inorganic ions and carbonaceous mass loads onto the residual oil factor than present in the CMB source profile.

## 4. Source apportionment of PM<sub>2.5</sub> in Hong Kong

In addition to the extensive PM<sub>10</sub> speciation data set that was the focus of the previous two chapters, the EPD also periodically conducts PM<sub>2.5</sub> sampling and speciation analysis. A year of 1-in-6 day sampling was conducted at three sites in 2000/2001 – Mong Kok (MK), Tsuen Wan (TW), and Hok Tsui (HT) which represent roadside, general urban, and “regional background” conditions, respectively. Sampling conducted in 2004/2005 included a fourth site – Yuen Long (YL) – which is also an urban site. A third year of sampling was performed in 2008/2009 with these data becoming available over the performance period of this project. 24-hour integrated samples were collected using two Rupprecht & Patashnick (now Thermo) Partisol samplers deployed at each site. Samples were collected onto Teflon filters for gravimetric analysis and also elemental analysis by x-ray fluorescence (XRF). Samples were collected onto quartz filters for gravimetric analysis and also anion analysis by ion chromatography (IC), ammonium ion by automated colorimetry, sodium and potassium ions by atomic absorption spectrometry, and carbon analysis by the IMPROVE thermo-optical reflectance (TOR) method. Data sets including both concentration values and record-specific uncertainties were provided by the EPD.

### 4.1 PM<sub>2.5</sub> versus PM<sub>10</sub> Species Concentrations

Prior to performing the PM<sub>2.5</sub> source apportionment we examined the relationships between species concentrations measured in the PM<sub>2.5</sub> and PM<sub>10</sub> samples, as shown in **Table 4.1** and **Figure 4.1**. While there are differences in the field sampling hardware and analytical methods for these data sets, the comparisons do provide insights into potential data quality issues and whether certain species are likely to be in the fine fraction (PM<sub>2.5</sub>) or coarse fraction (PM<sub>10-2.5</sub>). Table 4.1 lists the PM<sub>2.5</sub>/PM<sub>10</sub> ratio of total PM mass and concentration of each species. In general, PM<sub>2.5</sub> accounts for 68% of the PM<sub>10</sub> mass in Hong Kong, indicating significant fraction of particles are within the fine sizes. The following summary refers to the plots in Figure 4.1 in order of presentation.

Total carbon (TC) shows good agreement between PM<sub>2.5</sub> and PM<sub>10</sub> samples which demonstrates the carbonaceous matter is in the fine fraction. Some of the scatter, including PM<sub>2.5</sub> TC greater than PM<sub>10</sub> TC, might arise from measurement artifacts due to the different flow rates for the samplers. For OC the PM<sub>2.5</sub> values are less than or equal to the PM<sub>10</sub> values while for EC the PM<sub>2.5</sub> values are greater than or equal to the PM<sub>10</sub> values. The EC trend is not realistic for a consistent measurement methodology but in this case both the OC and EC trends arise from differences in the thermo-optical analysis protocols – NIOSH thermo-optical transmittance (TOT) for the PM<sub>10</sub> samples and IMPROVE TOR for the PM<sub>2.5</sub> samples.

$\text{SO}_4^{2-}$  and  $\text{NH}_4^+$  are in the  $\text{PM}_{2.5}$  size range with excellent agreement between the  $\text{PM}_{2.5}$  and  $\text{PM}_{10}$  data. In contrast,  $\text{PM}_{10}$   $\text{NO}_3^-$  concentrations are often much higher than  $\text{PM}_{2.5}$   $\text{NO}_3^-$ . A caveat is that the  $\text{NO}_3^-$  measurements might suffer from differential measurement artifacts and any physical explanation for these differences should be qualified.

Al, Ca, Fe and Mn all show significant coarse particle loadings. Ca and Fe concentrations in the  $\text{PM}_{2.5}$  and  $\text{PM}_{10}$  samples are very highly correlated which suggests the same sources contribute to both fine and coarse size modes. In contrast, Al and Mg exhibit considerably more scatter and thus there might be different sources contributing to the fine and coarse modes.  $\text{Cl}^-$  is overwhelmingly in the coarse fraction. Cu is also mostly in the coarse fraction but is not considered in these studies due to contamination by copper emissions from the high-volume sampler motors. Mn, Ni, Pb, V and Zn are predominantly in the fine fraction; there appears to be some Mn and Pb in the coarse fraction. V is substantially higher in the  $\text{PM}_{2.5}$  samples than the  $\text{PM}_{10}$  samples. This is physically unreasonable and there must be a V measurement error with at least one of the methods. Moreover,  $\text{PM}_{2.5}$  elemental K – measured by XRF – shows very good agreement with  $\text{PM}_{10}$  potassium ion measured by IC.

## 4.2 PMF Modeling Strategy

Source apportionment modeling was performed using EPA PMF 3.0. From the set of species listed in Table 2.4 the following were excluded from the modeling: TC was excluded because EC and OC were used; XRF sulfur was excluded because  $\text{SO}_4^{2-}$  was used;  $\text{Cl}^-$  was excluded because XRF chlorine was used; Cu and Na were excluded due to likely contamination; and Cr, Mg, Sn, and Sr because less than 50% of samples were above the MDL. The excluded species were retained in the input file and designated as “bad”. All remaining species were assigned as “strong” and the modeling was performed with 10% extra modeling uncertainty. Record-specific uncertainty estimates were generated using the approach described in Section 2.2.3. All four sites and three sampling years were combined into a single run with sensitivity studies conducted as noted below. Each modeling run included 20 base runs and the base run with the minimum  $Q$  value was retained as the solution. Solutions for six through nine factors were examined and the eight factor solution was deemed optimal. Bootstrap runs performed on this solution were quite stable based on the narrow spread of  $Q$  values. **Figure 4.2** shows the species loadings onto the eight factor solution. The factors were assigned to the following source categories which are listed in order of decreasing source contribution estimates over the three study periods:

- Vehicular Exhaust
- Secondary Sulfate
- Biomass Burning

- Secondary Nitrate
- Trace Metals
- Residual Oil Combustion
- Soil / Dust
- Chlorine / Fresh Sea Salt

Sensitivity studies were performed by executing PMF modeling runs on subsets of the data including: one run for each site including all years (four runs); one run for each year including all sites (three runs); runs for each site and year combination (eleven runs).

PM<sub>2.5</sub> mass SCEs were estimated using four approaches: the sum of the species as modeled (using OC); the sum of the species as modeled but using OM = 1.6×OC; using the Teflon filter gravimetric mass as the Total Variable; and using the quartz fiber filter gravimetric mass as the Total Variable. **Figure 4.3** and **Table 4.2** show the sensitivity of the SCEs for these four cases. Generally consistent SCEs were obtained across the four different approaches.

### 4.3 Characteristics of the PMF-Resolved Factors

*Secondary Sulfate Factor.* 79% of the factor SCE is explained by the measured species. Factor loadings are dominated by sulfate ion (58%) and ammonium ion (18%) with the next most abundant species being OC at only 2%. It is surprising that more OC did not load onto this factor which is likely dominated by transport. About 2/3 of the measured sulfate and ammonium are loaded onto this factor. The molar ammonium-to-sulfate molar ratio in this factor is 1.6 which is between the stoichiometric ratio for ammonium bisulfate (1.0) and ammonium sulfate (2.0). Assuming these factor loadings to be representative of ambient conditions, this suggests the Hong Kong sulfate aerosol is not fully, neutralized.

*Vehicle Exhaust Factor.* 84% of the factor SCE is explained by the measured species and an OM/OC ratio of 1.6-1.7 would close the mass balance on the factor loading. The factor loadings are dominated by EC (48%) and OC (25%) with OM at 40% if OM = 1.6×OC. Sulfate, nitrate and ammonium all load in the range 1-6% of the factor SCE. The high EC/OC ratio (1.9) is expected since these samples were analyzed using the IMPROVE TOR method which generally yields higher EC than the NIOSH TOT method used for the PM<sub>10</sub> samples. The PMF-resolved vehicle exhaust factor in the PM<sub>10</sub> apportionment had an EC/OC ratio of 1.3.

*Secondary Nitrate Factor.* 69% of the factor SCE is explained by the measured species (79% using OM = 1.6×OC). Factor loadings are dominated by nitrate ion (30%) and ammonium ion (10%) with significant contribution to the factor mass also from OC (16%) and EC (10%). This factor accounts for 80% of the nitrate ion. The molar ammonium-to-nitrate ratio of 1.2 is modestly higher than the 1:1 stoichiometric ratio for ammonium nitrate.

*Biomass Burning Factor.* While potassium is a key indicator species for this factor with 63% of the measured potassium, the factor mass is dominated by OC (59%) with some nitrate (3%) and various elements associated with crustal material at 0.3-2%. 73% of the factor SCE is explained by the measured species and an OM/OC ratio of 1.4-1.5 would close the mass balance on the factor loading. Crustal material is often associated with biomass burning activities which can resuspend soil. As described below the Ca/Fe/Si ratio in this factor is similar to the country park soil and urban soil profiles in the work of Ho *et al.* (2003). 58% of the Pb also loaded onto this factor which might suggest a similar source region for biomass burning and industrial emissions.

*Trace Metals Factor.* This factor contains about 2/3 of the measured Zn and 1/3 of the measured Mn as well as significant Fe, Pb, Si, and Ti. 91% of the factor SCE is explained by the measured species (95% using  $OM = 1.6 \times OC$ ). The factor loading is nearly 2/3 sulfate with ammonium present at 1:1 molar ratio. The presence of sulfate is consistent with operations such as smelting of sulfur-laden ores and coal-field boilers (both electric utility and industrial). EC and OC loadings are in the range 5-6%.

*Residual Oil Combustion Factor.* 56% of the factor SCE is explained by the measured species (66% using  $OM = 1.6 \times OC$ ). Ni and V are key indicator species in this factor; they are present at low mass fractions (0.2 and 0.6%, respectively) and the factor mass is dominated by sulfate (28%), OC (15%) and EC (10%). The V/Ni ratio of 3.6 is high but plausible; based on comparisons to  $PM_{10}$  V concentrations the  $PM_{2.5}$  V concentrations appear to be biased high. Residual oil is often 1-5% sulfur with typical Ni mass fractions of ~10-100g/g. Thus, only a small fraction of the fuel sulfur need be emitted as primary sulfate aerosol to achieve the S/Ni ratio of ~15 observed for this factor.

*Soil/Dust Factor.* The representativeness of the soil/dust factor was examined using the ratio of major crustal elements in the resolved factor and published profiles. Ho *et al.* (2003) sampled dust sources in Hong Kong and reported  $PM_{2.5}$  and  $PM_{10}$  profiles for county park soil, urban soil, paved road dust, and industrial dusts for cement and aggregate. **Figure 4.4** shows the relative distribution of Al, Ca, Fe, and Si for each of the TW ambient samples (open circles), the county park, urban soil, and paved road profiles of Ho *et al.* (red circles), and each of the resolved factors for the TW  $PM_{2.5}$  modeling (green circles). Based on the Ca/Fe/Si distributions (Figure 4.2a) there are clearly sources of Fe in addition those captured by the PMF-resolved soil factor which is at the bottom edge of the cluster of ambient data. The soil factor is consistent with the published profiles and falls about 40% along the mixing line between the county park / urban soil profiles and the paved road profile. The Al/Ca/Si distributions (Figure 4.2b) support this interpretation with the PMF-resolved soil factor again about 60% along the mixing line between the county park / urban soil profiles and the paved road profile but in this case being slightly off the mixing line. This suggests that a composite profile of 30% county park soil, 30% urban soil, and

40% paved road dust might be a reasonable approach for CMB modeling although it is noted that the PMF-resolved factor was based on modeling data aggregated across four sites with potentially distinct soil/dust type impacts. Figure 4.4a also shows that the biomass burning factor has a Ca/Fe/Si ratio very similar to the county park and urban soils but the factor is very low in Al compared to these soil profiles (Figure 4.4b).

*Sea Salt Factor.* Sodium is a preferred indicator species for sea salt with high chloride (approaching the sodium chloride stoichiometric ratio) suggesting fresh sea salt and low chloride suggesting aged sea salt. The sodium data in this study is of sufficiently poor quality that it cannot be used in the source apportionment modeling. Thus, chloride is the sole indicator species and can capture at best the fresh sea salt contributions. 44% of the factor SCE is explained by the measured species (70% using  $OM = 1.6 \times OC$ ). The factor includes 90% of the XRF Cl but this Cl accounts for only 6% of the factor mass which is dominated by sulfate (14%), OC (11%), nitrate (8%) and EC (4%).

#### 4.4 APCA Modeling

APCA was performed on the  $PM_{2.5}$  data with all species used in the PMF analysis species (OC, EC,  $NO_3^-$ ,  $SO_4^{2-}$ ,  $NH_4^+$ , Al, Ca, Cl, Fe, K, Mg, Mn, Na, Ni, Pb, Si, Ti, V, Zn) and quartz filter mass used as the total mass variable. Detailed steps involved in APCA methodology by MYSTAT are provided in section 2.1.2 and the application of APCA to the Hong Kong  $PM_{2.5}$  data is summarized in **Appendix B**. Similar to PMF results, Na and Mg kept emerging as dominant species in separate factors until the number of factors were reduced so that these species merged into other factors. Thus, Na and Mg were removed from the analysis. **Table 4.3** shows the factor-specific mass loadings with varying number of factors retained in the modeling. Factors were linked to the PMF-resolved source categories using the tracer species. The seven factor solution was ultimately chosen with these factors assigned to following emission source categories: secondary sulfate, vehicle exhaust, soil, biomass burning, residual oil combustion, sea salt and secondary nitrate. High loading onto the soil and residual oil combustion factors was observed, with almost negligible loading onto sea salt and secondary nitrate factors.  $NO_3^-$  was also found to be smeared over various factors. To better understand these study-average SCEs, the daily SCEs were examined. **Figure 4.5** shows cumulative distributions of the daily SCEs modeled by APCA for those factors with highest number of negative daily SCE values: secondary nitrate, sea salt and biomass burning. In each case, negative values were assigned to nearly 50% of the daily SCEs. The recurring presence of negative daily source contribution estimates from the model resulted in near-zero annual average SCEs for the sea salt and secondary nitrate factors. To further examine this model performance behavior, PCA analysis was also conducted within Unmix and resulted in similar negative loadings. Thus, APCA model was deemed unsuitable to resolve factors for the HK  $PM_{2.5}$  dataset.

## 4.5 Unmix Modeling

Unmix modeling was also performed on the PM<sub>2.5</sub> data. Detailed steps involved in Unmix methodology are provided in section 3.3.2 and the application of Unmix to the Hong Kong PM<sub>2.5</sub> data is summarized in **Appendix B**. When Unmix was initially run with all species used for PMF modeling, the “Select Species” function did not pick up any species. After manually selecting all species, five- and seven-factor solutions were obtained. Both of these solutions had factors with large negative species concentrations in addition to separate factors being dominated by Na and Mg. Thus, Na and Mg were removed from the analysis. However, this modeling issue continued with Zn and thus was removed. By specifying quartz filter gravimetric mass as the total mass variable, the select species function obtained a five factor solution with OC, EC, NO<sub>3</sub><sup>-</sup>, SO<sub>4</sub><sup>2-</sup>, NH<sub>4</sub><sup>+</sup>, Ca, K, Pb and Si. With the addition of the remaining species (Al, Cl, Fe, Mn, Ni, Ti, V), five- to eight-factor solutions were obtained. Based on the negative loadings assigned to some factors in certain solutions, the seven factor solution was deemed the optimal Unmix modeling result. Secondary sulfate, vehicle exhaust, biomass burning, secondary nitrate, residual oil combustion, soil/dust and sea salt were identified as the main factors contributing to the PM mass. Species loadings onto the seven factor Unmix solution are summarized in **Table 4.4** with the source profiles shown in **Figure 4.6**. Explained mass profiles are shown in **Figure 4.7** and were used to match the Unmix-modeled factors to the PMF-modeled source category assignments.

**Figure 4.8** shows the ratio of the site-specific average SCE to the all-sites average SCE for each of the resolved factors. Vehicle exhaust exhibits the greatest variation between sites, consistent with the roadside nature of the MK station and the remote location of the HT station. Secondary sulfate, biomass burning and soil factors show virtually no spatial variation between the sites, consistent with the presumed dominant role of non-local sources for these factors. Residual oil combustion is also nearly homogeneous across the sites; this source category is generally to be local and it is possible that the HT background site is impacted by nearby marine sources given its coastal location. The spatial variability exhibited by the secondary nitrate factor is not fully understood at this time, and might reflect impacts from a combination of local and non-local sources.

Factor-specific daily source contribution estimates modeled by PMF and Unmix are presented as scattergrams in **Figure 4.9**. Unmix consistently estimates quite low SCEs for soil and sea salt compared to PMF. Vehicle exhaust and residual oil exhibit good agreement at low SCEs; in contrast, at high SCEs Unmix tends to be lower than PMF. Secondary nitrate and biomass burning SCEs are moderately higher for Unmix than PMF. For most factors the Unmix and PMF results tend to be highly correlated with varying degrees of bias. The exception is the factor sulfate which shows significant scatter.

Overall, the PMF and Unmix modeling results are generally consistent. Using quartz filter gravimetric mass as the total variable for both modeling approaches

yielded grand average SCEs for that differed by less than 5% between the models for the two dominant factors – secondary sulfate and vehicle exhaust. This performance was achieved despite the relatively high scatter in day-specific SCEs for the secondary sulfate factor. Residual oil combustion also showed relatively agreement between the models with 15% difference. For the remaining factors, both the absolute and relative differences between the models were relatively larger and the optimal solution using Unmix did not include a trace metals factor that was resolved using PMF. While both models have merit and collectively provide insights into the stability of the modeling results across different approaches, the PMF results are presumed to be the preferred source apportionment for this data set. While the basis for this decision is subjective, for this particular analysis the added value of record-specific uncertainties and the non-negativity constraint provided by PMF appears to be more critical than the edge detection capabilities of Unmix. Distinct edges might be present for sources such as motor vehicles because HT site has quite low motor vehicle impacts. However, the persistent, large contributions from non-local sources may constrain the ability to find distinct edges using Unmix.

#### 4.6 Comparison of Source Contribution Estimates for PM<sub>2.5</sub> and PM<sub>10</sub>

As discussed in previous sections, PMF provides the most reasonable source apportionment results, for both PM<sub>10</sub> and PM<sub>2.5</sub> analysis. In this section, the PMF-derived SCEs are used to compare between PM<sub>10</sub> and PM<sub>2.5</sub> results. **Figure 4.10** and **Table 4.5** show the annual-average SCEs for each of the four sites. *Secondary sulfate* and *trace metals* show consistent patterns with similar SCEs across the four sites and highest SCEs in 2005. These factors are dominated by transport and it is likely the SCEs are lower in 2009 due to lower emissions during that period. *Vehicle exhaust* SCEs is highest at MK and very low at HT which is consistent with MK being a roadside station and HT being a relatively remote site. Vehicle exhaust SCEs have monotonically decreased at the urban sites with 2009 levels at MK and TW about 50% of 2001 levels, in consistent with the PM<sub>10</sub> analysis results. *Biomass burning* SCEs are similar across the four sites and have monotonically decreased over the study period. *Residual oil combustion* and *soil/dust* SCEs are similar across the four sites and exhibit no clear temporal pattern. **Figure 4.11** shows the relative contribution of each emission source category to the modeled PM<sub>2.5</sub> quartz filter mass for each site and year.

**Figure 4.12** compares the SCEs obtained from the PM<sub>2.5</sub> and PM<sub>10</sub> modeling. The PM<sub>2.5</sub> trace metals and biomass burning SCEs were composited to be consistent with the PM<sub>10</sub> combined factor. Direct comparisons are available for TW (2001, 2005) and YL (2005) because PM<sub>2.5</sub> and PM<sub>10</sub> samplers operated on the same schedule at each of these sites. Annual averages were constructed using the paired data. The PM<sub>2.5</sub> and PM<sub>10</sub> samplers operated on a different days at MK and in this case samples during the study periods were used to construct the annual averages.

Figure 4.12 demonstrates nearly identical PM<sub>2.5</sub> and PM<sub>10</sub> SCEs for secondary sulfate

and very similar SCEs for vehicle exhaust. This is consistent with emissions from both of these source categories being dominated by fine aerosol. Biomass burning/trace metals SCEs for  $PM_{2.5}$  are higher than  $PM_{10}$ . OC loads much more onto the  $PM_{2.5}$  factor than the  $PM_{10}$  factor (59% and 28% for  $PM_{2.5}$  and  $PM_{10}$ , respectively) and more than compensates for the larger unexplained mass in the  $PM_{10}$  biomass burning factor (-9% and 35% for  $PM_{2.5}$  and  $PM_{10}$ , respectively, using an OM/OC ration of 1.6). Residual oil combustion SCEs are a factor of two higher for  $PM_{2.5}$  than  $PM_{10}$  due to the significant higher loading of EC and OC onto the  $PM_{2.5}$  source profile. Nitrate  $PM_{2.5}$  SCEs are about 50% of the  $PM_{10}$  SCEs which suggests that considerable nitrate is in the coarse particle fraction. The unexplained mass was nearly identical for both  $PM_{2.5}$  and  $PM_{10}$  nitrate factors so its contribution to the SCEs cannot explain the difference. Soil/dust SCEs are predominantly in the coarse particle size range, as expected. Fresh sea salt is scattered about the 1:1 line; the  $PM_{2.5}$  SCEs are likely unreliable because the sodium data quality was too poor to be included in the  $PM_{2.5}$  modeling and thus this factor is resolved using only chlorine.

## 4.7 Summary and Recommendations

We have the following recommendations:

- OC and EC were used in this work to be consistent with the  $PM_{10}$  receptor modeling. Modeling should be performed using the IMPROVE thermal carbon fractions rather than simply OC and EC to determine whether these parameters provide better resolving power of the emission source categories. This work should be performed with caution because there are concerns about the stability of the thermal fractions data, especially when there are abundant loadings of species such as sodium that can catalyze the carbon to evolve at lower temperatures.
- In contrast to the  $PM_{10}$  speciation sampling, the four-site  $PM_{2.5}$  speciation network is operated on the same schedule and thus there can be day-specific comparison of inter-site differences in both the concentration values and modeled source contribution estimates. Analyses should be conducted to explore inter-site patterns in the data.
- Deployment of a wider network for  $PM_{2.5}$  speciation monitoring should be implemented if resources permit. This takes on greater importance with the recent commitment to establish a  $PM_{2.5}$  AQO. A measurement platform minimizes measurement artifacts (e.g. nitrate and semivolatile OC) yet is practical to deploy and operate is encouraged. This approach would permit the examination of subtle but important patterns in spatiotemporal variability. One approach would be the measurement platform used for the U.S. Chemical Section Network with some modifications to actually improve on that design. Such implementation could reduce measurement artifacts and better represent the ambient aerosol. This potential change must be carefully considered, however, because it would introduce a discontinuity in the time

series compared to the three one-year studies conducted over the previous decade and therefore it might be more difficult to assess long-term trends in  $PM_{2.5}$ .

- As discussed in Chapter 3, the selection of source profiles is essential to the accurate estimation of SCEs by CMB, and we have significant concerns about the representativeness of the selected source profiles and the inclusion of all relevant sources, especially secondary organic aerosol (SOA), for the  $PM_{10}$  study. Therefore, we did not perform CMB on the  $PM_{2.5}$  dataset and do not suggest to do so until a complete Hong Kong-relevant source profile library for particulate matter is established.
- APCA and Unmix tend to derive negative values for certain source profiles and SCEs. In comparison, PMF limits all the elements in the factor score and factor loading matrix to be positive. Moreover, PMF is generally more powerful in disentangling minor sources with infrequent impacts and/or small contributions, due in large part to the inclusion of uncertainty estimates in the calculation. Therefore, we conclude that PMF is the most suitable receptor model for the PM analysis in Hong Kong, considering the good data quality, availability of measurement uncertainty estimate and large amount of data available. Unreliable results obtained using APCA clearly demonstrate that the features of PMF compared to APCA are important. Furthermore, while stable solutions were obtained using Unmix, for reasons discussed in Section 4.6 the weight-of-evidence tends to favor the PMF modeling results. With accurate and reliable uncertainty estimation, PMF is recommended to use for further source apportionment studies in Hong Kong. The PMF analyses should be periodically updated as more data becomes available.

## 5. Conceptual Model for Ambient Particulate Matter over Hong Kong

### 5.1 Introduction

In 1987, Hong Kong Air Quality Objectives (AQO) was established for seven widespread air pollutants including Respirable Suspended Particulates (RSP).<sup>11</sup> The PM<sub>10</sub> objectives include a 24-hour concentration of 180 µg/m<sup>3</sup> not to be exceeded more than once per year and annual arithmetic mean concentration of 55 µg/m<sup>3</sup>. In 2009 the annual average PM<sub>10</sub> AQO was achieved at all general stations – the highest value being 51 µg/m<sup>3</sup> (93% of the objective) at Yuen Long – and was not achieved at the Causeway Bay roadside station (71 µg/m<sup>3</sup>, 129% of the objective). Despite the implementation of numerous control measures within the HKSAR over the past two decades, annual average PM<sub>10</sub> concentrations have remained relatively constant at most monitoring sites throughout the territory. This trend, together with the concentration metrics being close to or in excess of the AQO, motivates the need for a refined understanding of the drivers for the observed RSP burdens. In particular, a detailed description is needed of the emissions, meteorology and atmospheric processes that modulate PM<sub>10</sub> concentrations. This “conceptual model”, with explicit emphasis on understanding long term trends and high mass concentration days over the HKSAR, could be utilized towards developing effective control strategies both within the territory and in collaboration with other regions including but not limited to Guangdong Province.

The conceptual model presented in this document is focused on the information needed to support the development and implementation of effective air quality management strategies for ambient particulate matter. The AQOs are currently being reviewed and, in addition to the AQO for PM<sub>10</sub>, there is some discussion of an adopting an AQO for PM<sub>2.5</sub>. Therefore, this conceptual model addresses both size ranges. A conceptual model for PM<sub>2.5</sub> over St. Louis was recently developed which included the following elements: annual and seasonal levels of PM<sub>2.5</sub> in relation to the mass-based standards; compositional analysis of PM<sub>2.5</sub>; meteorological influences; atmospheric processes contributing to PM<sub>2.5</sub>; sources and source regions contributing to the principle chemicals of concern; and implications to policy makers (Turner and Garlock, 2007). Content and presentation of the conceptual model for ambient particulate matter over Hong Kong loosely follows the framework used by Vickery (2004) for nine North American geographic regions and as revisited by Allen and Turner (2008), as shown in **Figure 5.1**. Specifically, the topography of Hong Kong is much more complicated than St. Louis which can bring about a series of micro-scale meteorological characteristics such as land-sea breeze circulation and

---

<sup>11</sup> RSP is defined as particulate matter smaller than 10 micrometers aerodynamic diameter and is hereafter termed PM<sub>10</sub>.

mountain-valley wind. As detailed in the next section, these characteristics can influence the pollution transport mechanism and the pollutant mass, therefore have to be considered as an important part of the conceptual model. For this study, the framework has been modified to accommodate an emphasis on those days with high PM concentrations.

This conceptual model is grounded in an analysis of 10 years of PM<sub>10</sub> hourly mass concentration data collected at 10 sites, 11 years of 24-hour integrated PM<sub>10</sub> speciation data collected at up to 10 sites, and three years of 24-hour integrated PM<sub>2.5</sub> speciation data collected at up to four sites. Both surface winds and synoptic-scale air mass transport patterns are used to interpret the observed trends. Despite this rich data set underpinning the conceptual model, it should be treated as a living document to be periodically updated as additional data become available and revised as more-detailed analysis is performed on the existing data.

## **5.2 Interaction of Meteorology and Air Quality in Hong Kong and the Pearl River Delta**

In developing the conceptual model, it is essential to understand the geographical and meteorological conditions that may lead to pollution events in the region. The PRD is located roughly at the center of the Southeastern China coastline and Hong Kong is located in the southeastern tip of the PRD. The climate of the PRD is dominated by the East Asian Monsoon, and is characterized by a long winter and summer, with a relatively shorter spring and fall in between. In summer, the southwest monsoon brings clean oceanic air to HK and PRD. This corresponds to the cleaner background air quality conditions in southern China. In winter, the prevailing northeast monsoon over the continent often brings pollutants from inland, including those from the Yangtze Delta and Fujian / Taiwan to the PRD, and further down to Western Guangdong / Hainan Island. During the transition seasons of spring and fall, the wind is generally weaker, leading sometimes enhanced accumulation and high pollutant concentrations. Furthermore, the situation could get even worse during the fall due to the higher temperature and more abundant solar radiation, both of which can increase the formation rate of secondary pollutants in the region.

In the synoptic time scale, the most important weather systems affecting the air quality over the region are the tropical cyclones in the western Pacific near the Philippines or Taiwan around 800-1000 km east to the PRD in the summer and early fall. These tropical cyclones can substantially worsen the air quality through subsidence, enhanced stability and reduced dispersion ability of the atmosphere. During these periods, the original summertime prevailing southwesterly wind can often deviate to northerly or northwesterly, bringing significant amount of accumulated pollutants to the south / southeast part of the PRD, such as Hong Kong. Some of the worst pollution events in the region are associated with such tropical cyclone conditions. The air quality pollutant variations associated with a wintertime cold front can be characterized by a short term increase in pollutant levels due to local

convergence ahead of the cold front, followed by a rapid decrease in concentration levels due to strong winds and enhanced dispersion, and finally a slow increase back to normal due to stagnation and accumulation. However, this strong wind occasionally brings sandstorms from northern China and Mongolia to southern China, especially during early spring. The sandstorms can make the Air Pollution Index surging, however, the larger particles associated with the sandstorm are generally less unhealthy and have lighter impacts on visibility degradation. In addition, the occurrence frequency of such pollution events in the PRD is rare.

Last but not the least, in the diurnal time scale, the presence of land-sea breezes over the Pearl River Estuary, which is generally resulted from high temperature, strong solar radiation and weak background wind, can also affect the pollution level through enhanced trapping and mixing of pollutants along the PRD coastline and western Hong Kong (Lo *et al.*, 2006, 2007). This land-sea breeze circulation is a unique and efficient mechanism capable of mixing pollutants emitted from various places along the PRD coastline, where industrial zones are concentrated, to other places in the delta where general prevailing wind can hardly bring about the pollutants there. Under these circumstances, a convergence zone often develops transecting Hong Kong as a result of westerly sea breeze in the western Hong Kong and the background weak easterly / northeasterly wind in the eastern Hong Kong (Fung *et al.*, 2005). The mountainous area in the center of Hong Kong also prevents the pollution brought by land-sea breeze from traveling eastward, leading to higher pollution level in the western Hong Kong than the eastern part.

The PRD has been experiencing prominent urbanization process in the past decades. The booming urban areas largely change the land-use of the delta, making the heat capacity difference between the water and the land strengthened. On the other hand, the rough underlying surface of urban land-use largely reduces the penetration efficiency of the background wind in the delta. Both factors would enhance the area controlled by the land-sea breeze circulation over the Pearl River Estuary, bringing more pollutants accumulated in the circulation, and hence elevating the pollutant levels at sites near the coastline.

The land-sea breeze effect is evident in surface winds data collected along the coastline. **Figure 5.2** shows wind roses generated using ten years of hourly surface winds data collected at the Hong Kong International Airport (HKIA) for the months of January and July. The data are stratified into three-hour intervals and calms are secluded from the plots. In January (Figure 5.2a), the nighttime and early morning hours are dominated by winds from the east. However, by late morning (10-12 hours) the winds have typically shifted to being from the northwest and this pattern persists until the late afternoon (16-18 hours). Subsequently, there is a transition back to easterly winds which persist throughout the evening and nighttime. In July (Figure 5.2b), the prevailing wind is from the southwest and daily wind direction transitions from a land-sea breeze effect are not evident. **Figure 5.3** shows a similar set of wind

roses for the Hong Kong Observatory (HKO). In January (Figure 5.3a) the winds are persistently from the east regardless of time of day. The prevailing nighttime wind directions are generally similar for HKIA and HKO, but the daytime onshore sea breeze effect does not penetrate far enough inland to affect HKO which is also strongly influenced by local topography. In July (Figure 5.3b), winds are from either the east or the west with no time of day pattern. The July wind patterns at HKO are dramatically different from HKIA.

In summary, the air quality in Hong Kong and the PRD is largely dependent on the meteorological conditions in three scales, continental-scale monsoon system, synoptic scale system such as tropical cyclone, and local scale such as land-sea breeze circulation and urban heat island effect. These factors have to be considered in the conceptual model. These three scales of meteorological conditions are also closely associated with emission sources in three different scales, which would be evaluated by this conceptual model.

### 5.3 Air Quality Conditions – Local and Non-local Impacts

The Hong Kong Environmental Protection Department (EPD) operates an extensive monitoring network for ambient PM mass and composition. Hourly PM<sub>10</sub> mass has been measured at fourteen sites for about a decade – and in some cases longer – with monitoring stations located throughout the HKSAR. Some of these stations located to capture roadside concentrations while others are located to capture neighborhood scale concentrations. PM<sub>10</sub> tapered element oscillating microbalances (TEOM) are operated at 50°C which removes particle-bound water but also other volatile material species; thus, the TEOM data represent the nonvolatile component of the ambient particulate matter. **Figure 5.4** shows the locations of the TEOM monitoring stations.

**Figure 5.5** (top) shows 5-year average PM<sub>10</sub> mass concentrations across the network. Roadside station concentrations (68-84 µg/m<sup>3</sup>) are significantly higher than general station concentrations (49-61 µg/m<sup>3</sup>). The maximum variability across the general stations (12 µg/m<sup>3</sup>) is only about 20% of the average value, demonstrating a high baseline for PM concentration levels throughout the HKSAR. Despite efforts to reduce emissions of PM and its precursors within the HKSAR, **Figure 5.5** (bottom) demonstrates that air quality conditions have not improved over the previous decade. Five-year average concentrations for the periods 1999-2003 and 2004-2008 exhibit a nearly 10 µg/m<sup>3</sup> reduction at the CB roadside station but no change at the other two roadside stations (MK and CL). Concentrations have increased 0-5 µg/m<sup>3</sup> at most of the interior sites and 7 and 9 µg/m<sup>3</sup> at the stations to the extreme southwest (TC) and northeast (TM), respectively. These stasis or even degradation of ambient PM air quality conditions has motivated detailed analyses of the PM mass and speciation data and the development a conceptual model for ambient PM over Hong Kong to provide a framework to understand the patterns of and drivers for observed conditions.

There are significant challenges to understanding the drivers for the observed

spatiotemporal patterns. Most importantly, ambient particulate matter burdens in the HKSAR are significantly impacted by local sources, transport from the greater PRD and so-called long-range transport from the entire southeast coast of China and beyond. However, in the context of data Hong Kong is an island with virtually no ground-based observations data available for the two transport regions. Some information can be obtained from satellite-based observations and modeling, but at this time these analyses only qualitatively inform our understanding of how transport from the PRD and beyond modulates ambient particulate matter levels in Hong Kong. Three analysis strategies were used to towards the development of this conceptual model. The rich data density for PM<sub>10</sub> TEOM mass was exploited by developing day-specific semi-quantitative estimates of urban- versus larger-scale contributions. Key findings are summarized in this section. Receptor modeling was performed on PM<sub>10</sub> and PM<sub>2.5</sub> speciation data to apportion PM mass to emission source categories. These results are summarized later in this report. While the assignment of local-versus- non-local contributions is subjective, these two analyses provide distinct but complementary approaches to elucidating the spatial scales of sources impacting the HKSAR. Furthermore, back trajectory analysis was performed in an attempt to separate impacts of sources with different spatial scales.

**Figure 5.6** shows a conceptualization of intraurban variability arising from PM emission sources acting on various spatial scales.<sup>12</sup> The air mass is transecting the domain with persistent surface winds from the left. The top panel shows the emission sources where the regional-scale contributions are assumed to be homogeneous across the domain (this assumption neglects interactions between regionally transported material and urban area emissions, such as urban plume processing of regionally-transported PM precursor gases). Urban-scale emissions from mobile sources, area sources, and point sources distributed across the metro core contribute to an increase in PM mass concentration (dashed line) which is skewed towards the downwind end of the domain. Neighborhood- and finer-scale emissions create local hot spots which disperse downwind and become part of the urban-scale contributions. The red line shows the concentration profile that would be observed along the transect. Monitor locations will dictate which features of the profile are captured. For example, the bottom panel shows the location of five monitors (the “network-wide baseline” will be discussed later in this report). Monitor **A** is located upwind of the urban area and measures the regional contributions. However, a wind reversal would render it downwind of the urban area and in this case it might be impacted by the urban plume (such as monitor **D**). Monitor **B** is impacted by both a local plume and upwind urban contributions. The plume would miss the site with a wind reversal. Monitor **C** is located near the core of the urban area (from an urban-scale emissions perspective). For this scenario it is not impacted by local plumes but measures the cumulative contributions from upwind sources, including point sources in the urban area. A wind reversal, however, might render monitor **C** impacted by a local plume. Monitor **D** is outside the urban area emissions zone but

<sup>12</sup> This paragraph was excerpted – with only minor modification – from Turner (2010).

nonetheless is impacted by the “urban plume” as it is advected downwind, while monitor **E** is far enough downwind that the urban plume has dispersed to the extent that the urban PM mass contributions cannot be distinguished from the underlying regional contributions. A wind reversal would render monitors **D** and **E** upwind of the urban area. This spatial concentration field represents a snapshot in time which becomes smeared for a 24-hour sample as winds shift during the day, and becomes further smeared when averaging results over many days. It is also a 1-D representation of a 2-D surface which adds complexity. Despite these limitations, this conceptualization can be useful towards analyzing and interpreting ambient particulate matter data. This conceptualization serves as a base for the development of conceptual model for Hong Kong with improved knowledge on the meteorological, geographical, physical and chemical processes in different spatial scales.

A preliminary analysis was performed by partitioning the daily-average TEOM mass measured at each station into a network-wide “base” concentration and a site-specific “excess” concentration. Valid network days were defined as those days with a valid 24-hour average concentration for at least twelve of the TEOMs. The metric used to estimate the base concentration is subjective. In this study, it was chosen as the 5<sup>th</sup>-lowest concentration for a given day. A sensitivity study was performed using other metrics for defining the base concentration (e.g. 4<sup>th</sup>- or 6<sup>th</sup>-lowest concentration) and the results exhibited similar trends albeit modestly different quantitative estimates (**Appendix C**). The time period covered 1999-2008. **Figure 5.7a** shows box plots for annual distributions of the base concentration.<sup>13</sup> The temporal pattern likely represents the influence of larger-scale sources on PM concentrations. The general pattern is an increasing trend for most metrics of the base concentration – the 95<sup>th</sup>, 75<sup>th</sup>, mean and median values are higher for the latter five years compared to the former five years. In contrast, the 5<sup>th</sup> and 25<sup>th</sup> percentiles are nearly the same over the ten year period. Thus, base concentration levels on the cleanest days (e.g. Bottom 25<sup>th</sup> percentiles) have remained unchanged but in recent years there have been an increased frequency of high concentration contributions from non-local sources. It is noted that the temporal patterns are not monotonic and could be influenced by year-to-year differences in regional scale emissions and synoptic weather. Synoptic weather influences, expressed in terms of air mass transport patterns, are discussed later in this report.

**Figure 5.7b** shows the distribution of PM<sub>10</sub> TEOM mass excess concentrations at each site over the entire time period. The sites are stratified into three groups as distinguished by the vertical lines in Figure 5.7b. The first group is a transect, across the HKSAR, of general station sites running from the northwest to the southeast. Excess concentrations are substantially higher and lower at the two extents (YL and

---

<sup>13</sup> Unless otherwise noted, all box plots in this report are formatted as follows: open circles are 5<sup>th</sup> and 95<sup>th</sup> percentile values; the bottom and top of the boxes are 25<sup>th</sup> and 75<sup>th</sup> percentile values; and the interior red and black lines are mean and median values, respectively. The number of samples is in parenthesis at the top of each plot.

EN), and high at YL perhaps due to its proximity to Shenzhen. The interior five sites of this transect have similar excess concentration distributions. The second group of sites are general stations peripheral to the transect stations. Excess concentrations are relatively low and, together with CW, the median concentrations at TP and ST are nearly equal to the base concentration and thus these sites provide a context for interpreting the base concentration. While excess concentrations at TC are much lower than at YL, they share similar characteristics with relatively broad distributions and arithmetic mean values significantly higher than the respective median values. Both of these patterns suggest relatively more dynamic behavior at these sites – either as short-term (e.g. Daily or seasonal) fluctuations or long-term trends. This will be discussed in more detail. TM also has a relatively broad distribution but in this case the arithmetic mean and median values are nearly identical which suggests different behavior from TC. The last group of sites is the roadside stations; as expected, they exhibit the highest excess mass concentrations due to near-field vehicle emissions.

Temporal patterns for the total and excess  $PM_{10}$  mass are shown in **Figure 5.8** for two sites – the TC general station and CB roadside station. Total  $PM_{10}$  mass at TC (Figure 5.8a) has increased over the ten year period which is consistent with the temporal pattern for the network-wide base concentration (Figure 5.7). Like the base concentration, the trend is not monotonic and there are year-to-year fluctuations in most concentration metrics. The  $PM_{10}$  excess mass concentration at TC (Figure 5.8b), presumed to arise from local-scale emission source impacts, has also increased over the past decade. In this case, however, the year-to-year fluctuations have been damped and the trend is more closely monotonic. The stabilized behavior arises because regional-scale impacts can be sensitive to year-to-year variations in surface wind patterns but the year-to-year variations in synoptic weather have been largely removed by subtracting out the base concentration to calculate the local excess.  $PM_{10}$  mass has dramatically decreased at the CB roadside station (Figure 5.8c). Again, the temporal pattern has been stabilized by subtracting out the base concentration to calculate the local excess concentration (Figure 5.8d).

**Figure 5.9** shows annual and monthly patterns for the base concentration and excess concentration at five sites (YL, TC, TM, SSP, and CB). Annual and especially monthly excess concentrations at YL and TC tend to follow the base concentration patterns. The increase in excess concentration at YL is small whereas the increase is more evident for TC. Despite the urbanization over the past decade around these sites it is difficult to draw conclusions about the local versus non-local contributions since the patterns so closely follow the base concentration. TM also has low excess mass (Figure 5.7b) but, as previously noted, in contrast to YL and TC the excess mass distribution at TM has nearly identical mean and median concentrations. This suggests different drivers for the excess concentration which is supported by Figure 5.9. There has been a monotonic increase in excess mass at TM and, the  $PM_{10}$  level in TM in 2008 has been exceeding the average level of the territory. In addition, in contrast to the base concentration; the excess mass is nearly constant throughout the

year. This suggests local source influences have been growing at the TM site. Annual concentrations at the CB roadside station have dramatically decreased; the excess throughout the year although modestly lower in the winter compared to the summer. This may probably arise from the relatively intense emission due to the turn-on of the air conditioner. The SSP site is one of the interior stations in the northwest-to-southeast transect previously described. The distribution of annual excess concentration has modestly decreased over the ten year period with timing consistent with the timing of the large decreases at the roadside stations such as CB. There is no discernable monthly pattern to the excess concentration at SSP.

Several key findings emerge from this analysis. First, it appears that contributions from non-local sources have increased over the past decade, affecting concentration levels across the entire HKSAR. Second, the temporal pattern for local source impacts is site specific. For example, motor vehicle emission control strategies have led to a significant decrease in the local source impacts (excess concentrations). This is clearly evident in Figures 5.7 (c,d) and is supported by the receptor modeling summarized later in this report. The area around the TC station has undergone extensive urbanization over the past decade and there has been a coincident modest increase of  $\sim 5 \mu\text{g}/\text{m}^3$  excess concentration at this site. There could be two drivers for this increase – increased local emissions due to the extensive urbanization that has occurred in this region over the past decade and/or differentially greater transport impacts at this site compared to the network-wide estimate due to its location. In contrast, the excess concentrations at TM have also increased but show a distinctly different pattern with no seasonal trend which suggests that local influences are indeed increasing. The conceptual model could be expanded and refined by examining the data from each station using this approach and developing additional analyses to exploit the hourly temporal resolution of the data.

#### **5.4 Local and Non-local Pollutant Sources**

Section 3 utilized the Hong Kong network of  $\text{PM}_{10}$  mass measurements to demonstrate spatiotemporal trends in PM mass. Evidence was provided for a long term trend of increasing impacts from non-local sources with site-specific patterns modulated by geography and urbanization. This section utilizes the PM chemical speciation network data to examine spatiotemporal patterns for impacts from specific emission source categories. HKEPD has operated a  $\text{PM}_{10}$  speciation network for more than ten years with seven stations having a relatively high level of data completeness (Table 3.1). 24-hour integrated samples are collected every sixth day (albeit not on the same days for the entire network) and analyzed for certain elements, ions and carbon. Nineteen species (eleven elements, six ion, and elemental carbon (EC) and organic carbon (OC)) were used for source apportionment modelling by Positive Matrix Factorization (PMF), Unmix, Absolute Principle Components Analysis (APCA) and the Chemical Mass Balance (CMB) methods. In this section, the PMF modelling results are used to inform the conceptual model. Generally

consistent results were obtained with the other models, although some exceptions are noted.

PMF modelling of the Hong Kong PM<sub>10</sub> aerosol resolved nine factors that were assigned to emission source categories based on their chemical characteristics (Table 3.3). Allied information, when applicable, was used to evaluate the robustness and representativeness of these assignments. For example, Figure 3.5 shows the relationship between annual motor vehicle source contributions estimates averaged across the six general stations and annual PM emissions in the road transport emission inventory prepared by HKEPD. These measures of impact and activity, respectively are highly correlated which supports the assignment of this factor to motor vehicles. However, the 1.3 µg/m<sup>3</sup> intercept suggests there are other, non-vehicle sources contributing to this factor and/or PM emissions in the road transport inventory are being underestimated. As another example, Figure 3.27 shows the relative distributions of Al, Ca and Fe in the ambient PM<sub>10</sub> samples (open circles), soil profiles presented by Ho *et al.* 2003 (red circles), and PMF-resolved factors (green circles). The paved road dust profile lies at the edge of the ambient data while the urban park and urban soil profiles are far removed from the ambient data with much higher aluminium levels. The PMF-resolved soil factor is closest in composition to the paved road dust profile with is consistent with our expectations in the highly urbanized, built environment of Hong Kong.

HKEPD has operated a three-to-four site PM<sub>2.5</sub> speciation network for three one-year periods over the past decade. PMF modelling of these data are useful not only to characterize ambient PM<sub>2.5</sub> burdens but also as a consistency check against the PM<sub>10</sub> modeling results. Figure 4.12 shows station-specific comparisons of PM<sub>2.5</sub> and PM<sub>10</sub> annual average source contribution estimates (SCE). The PM<sub>2.5</sub> trace metals and biomass burning SCEs were composited to be consistent with the PM<sub>10</sub> combined factor. Direct comparisons are available for TW (2001, 2005) and YL (2005) because PM<sub>2.5</sub> and PM<sub>10</sub> samplers operated on the same schedule at each of these sites. Annual averages were constructed using the paired data. The PM<sub>2.5</sub> and PM<sub>10</sub> samplers operated on a different days at MK and in this case samples during the study periods were used to construct the annual averages. Figure 4.12 demonstrates nearly identical PM<sub>2.5</sub> and PM<sub>10</sub> SCEs for secondary sulfate and very similar SCEs for vehicle exhaust. Biomass burning/trace metals SCEs for PM<sub>2.5</sub> are higher than PM<sub>10</sub>. Residual oil combustion SCEs are a factor of two higher for PM<sub>2.5</sub> than PM<sub>10</sub>. Nitrate PM<sub>2.5</sub> SCEs are about 50% of the PM<sub>10</sub> SCEs which suggests that considerable nitrate is in the coarse particle fraction. Soil/dust SCEs are predominantly in the coarse particle size range, as expected. Fresh sea salt is scattered about the 1:1 line; the PM<sub>2.5</sub> SCEs are likely unreliable because the sodium data quality was too poor to be included in the PM<sub>2.5</sub> modeling and thus this factor is resolved using only chlorine.

PMF modelling of the PM<sub>10</sub> data was used to examine long-term trends in the nature and scale of impacts from various emission source categories (Figure 3.4). Vehicle

exhaust shows a clear decreasing trend for six consecutive years from 1998 to 2003 at both types of stations. With small rebounds in 2004 and 2006, its contribution continued to decrease for three consecutive years from 2006 to 2008. This general decreasing trend, together with the strong correlation between PMF-modeled motor vehicle annual SCEs and the transport emission inventory shown in Figure 3.5, clearly indicates that pollution control measures on vehicles in HK are effective. The annual average contribution from vehicles was  $3.0 \mu\text{g}/\text{m}^3$  at general stations in 2008, which corresponds to a 67% reduction from the contribution of  $9.2 \mu\text{g}/\text{m}^3$  in 1998.

In contrast to the observed decreasing trend for motor vehicle emissions impacts, both secondary sulfate and secondary nitrate showed generally increasing SCEs over the period 1998-2008. This clearly indicates increasing outside impact to HK in the past decade. Contribution from secondary sulfate at general stations increased from  $6.5 \mu\text{g}/\text{m}^3$  in 1998 to  $12.6 \mu\text{g}/\text{m}^3$  in 2007, corresponding to a 93% increase. On the other hand, contribution from secondary nitrate at general stations increased from  $6.0 \mu\text{g}/\text{m}^3$  in 1998 to  $8.3 \mu\text{g}/\text{m}^3$  in 2007, corresponding to a 38% increase. Secondary sulfate impacts are generally consistent across all stations which suggest a dominant role from non-local scale sources. Secondary nitrate exhibits modestly greater spatial variability across the monitoring network with levels impacted by both local and non-local emission sources. While Figure 3.4 demonstrates that temporal patterns are evident for other emission source categories (e.g. An increasing trends for residual oil, a maximum in coal combustion/biomass burning impacts), the contributions from motor vehicles, secondary sulfate and secondary nitrate dominate the temporally varying components of the  $\text{PM}_{10}$  mass and largely explain the overall trend of little improvement in  $\text{PM}_{10}$  mass over HK in the past decade despite the implementation of emission control measured on motor vehicles.

## 5.5 Relationship between Pollutant Level, Source Contributions and Air Mass Transport Patterns

### 5.5.1 Approach

Receptor modeling results summarized in Sections 3 and 4 provide insights into the local and non-local emission source categories influencing ambient PM levels in Hong Kong. These patterns are further examined through the relationships between ambient PM and air mass transport patterns. For this analysis, air mass back trajectories were clustered into groups that capture similar synoptic scale transport behavior. Subsequently, the day-specific cluster assignments were used to interpret spatiotemporal patterns in the PMF-modeled source contribution estimates and TEOM  $\text{PM}_{10}$  mass data.

A description of the air mass back trajectory generation and clustering methodologies is presented in Section 2.3. Five air mass transport patterns were resolved with mean trajectories shown in **Figure 5.10** with the following characteristics:

- **Slow ECC** – relatively slow moving air masses transported from the north with

the centroid trajectory of the cluster located along the eastern coast of mainland China (ECC = East Coast of China);

- **Fast ECC** – relatively fast moving air masses transported from the north with the centroid trajectory of the cluster located along the eastern coast of mainland China;
- **Stagnant** – stagnant/circulating air masses that reside over the greater HKSAR area for much of the seven day period (this does not necessarily imply no air movement);
- **S/SW** – South/southwesterly flow into Hong Kong, the prevailing condition during the Monsoon season; and
- **Easterly** – air flow from the Pacific Ocean.

**Figure 5.11(top)** shows the distribution of air mass transport patterns by month over the period 2000-2009. September through March conditions are dominated by northerly air masses nominally from the eastern portion of mainland China. October is the month with the most frequent occurrence of Slow ECC (65%) while December is the month with the most frequent occurrence of Fast ECC (40%). There is only limited occurrence of Slow and Fast ECC during June (10%), July (1%) and August (7%). Since there are significant levels of pollutant emissions in the eastern portion of mainland China, Slow and Fast ECC trajectories are often associated with high pollutant levels. Starting in January the stagnant/circulating air masses pattern becomes more common with relatively high frequency of occurrence in April (32%) and May (30%). April and especially May are transition months from the wintertime north/northeasterly air mass patterns to the summertime south/southwesterly air mass patterns. South/southwesterly air mass patterns prevail during the monsoon season of June (58%), July (54%) and August (50%). Since the south/southwesterly air mass is originated from the ocean with limited emission sources, the air quality in Hong Kong is better in summer. South/southwesterly air mass patterns seldom occur during October through April, with the highest monthly frequency only 1%. September is also a transition month leading into the fall/winter season with prevailing northerly air mass patterns. This behavior demonstrates that for the case of Hong Kong, the summer and winter seasons described in Section 5.2 are good proxies for the northerly and south/southwesterly air mass transport patterns, respectively.

**Figure 5.11(bottom)** shows the distribution of air mass transport patterns by year over the period 2000-2009. There is some year-to-year variability in the relative frequency of the air mass transport patterns and this information could be considered when interpreting annual trends in air quality conditions. For example, the frequency of Slow ECC generally ranged from 30%-40% in a year. However, this percentage reached 44% in 2004. The occurrence frequency of stagnant air mass is generally below 20%. However, this percentage reached 22% in 2003 and 2004.

Both factors may contribute to the significant pollution in Hong Kong in 2004. On the contrary, the frequency of south/southwest air masses was the highest in 2008 (21%) and 2005 (19%). Correspondingly, the air pollution in these two years was less severe.

For the remaining analyses, if at least three of the four daily trajectories were assigned to the same cluster then that cluster was selected to represent the day. The remaining days were deemed to have transitional conditions and no air mass patterns were assigned to these days. This approach resulted in 82% of days over the period 2000-2008 having an air mass pattern designated.

### 5.5.2 Results for PMF-Modeled PM<sub>10</sub> Source Contribution Estimates

PMF-modeled PM<sub>10</sub> source contribution estimates (SCEs) were analyzed using the cluster analysis results. For each day over the period 2000-2008, the SCEs for each factor were averaged over the general station sites to generate a single-valued SCE to represent the HKSAR. While there are spatial gradients in the SCEs for certain factors across the HKSAR (e.g. Secondary sulfate), these variations are modest compared to the absolute SCE values. Each of the samplers is operated on a 1-in-6 day schedule but the sampling days are staggered across the network. The averaging approach exploits that sampling design to yield a daily time series with 67% completeness compared to a maximum completeness of 17% if using data from a single station. This spatial averaging approach of the SCEs is also consistent with the 1° resolution of the meteorological data set used to generate the trajectories for the cluster analysis.

**Figure 5.12** shows for each PMF-resolved factor the distribution of scaled source contribution estimates by each synoptic pattern for the period 2000-2008. SCEs were scaled such that the average SCE over all samples was unity for each factor. The interpretation for each factor is as follows.

- *Motor Vehicle*: The distribution of scaled SCEs is independent of synoptic class, with the cluster-specific average scaled SCEs ranging from 0.95 to 1.03 only. This is consistent with our expectation because the motor vehicle factor is dominated by local contributions. Different synoptic patterns could influence these local contributions through differences in atmospheric ventilation but this is likely a higher order effect. This case serves as a null hypothesis test of the approach – a local source whose impacts show no sensitivity to synoptic weather patterns.
- *Residual Oil*: Differences between the distributions are relatively small which is consistent with residual oil being a local emission source. Highest mean (1.34) and median (0.95) values are observed for stagnant weather patterns, consistent with air flow circulating over the HKSAR/PRD region.
- *Secondary Sulfate*: Highest SCEs are associated with the slow and fast moving

air mass patterns along the east coast of mainland China with modestly higher median and mean SCEs for the slow air masses (mean 1.51, median 1.25) compared to the fast air masses (mean 1.09, median 0.93). Caution must be exercised to not simply assign these contributions to long-range transport, however, because the role of regional scale transport from the PRD cannot be separated using this analysis. That stated, this behavior does motivate the need for data and analyses to better separate PRD versus non-PRD impacts. Examination of the trajectory path might be a possible way of PRD / non-PRD differentiation, as discussed in section 5.6.

- *Secondary Nitrate*: The distributions are generally similar to secondary sulfate with the transport patterns along the east coast of China corresponding to the highest SCEs. However, in contrast to secondary sulfate, mean and median SCEs are modestly higher for fast moving air masses (mean 1.75, median 1.40) compared to slow moving air masses (mean 1.20, median 0.92). This may be due to the large deposition rate of nitrate acid, which would shift the gas-particle equilibrium and be more significant for slow-moving air mass which has more time for nitrate deposition.
- *Biomass Burning*: This pattern also follows secondary sulfate and secondary nitrate behavior with high mean SCEs for Slow ECC (1.42) and Fast ECC (1.77), and very low mean SCEs for the stagnant (0.57), S/SW (0.23), and East (0.36) classes. This pattern is consistent with the timing and location of MODIS-detected fires in southeastern China and also the NRL model output.
- *Crustal/Soil*: Dramatically higher SCEs are observed for fast moving air masses (2.01) along the east coast of China compared to the other four synoptic patterns. It is possible that the soil is local in origin and the fast moving air masses at the synoptic scale correspond to high surface winds speeds at the local scale that would enhance the suspension of dust. This can be examined by comparing the day-specific synoptic class assignments to surface winds. It is also possible that there is a non-local contribution to this factor and thus needs to be examined.
- *Fresh Sea Salt*: Highest SCEs correspond to fast moving air masses (1.37) from along the east portion of mainland China and air masses from the south/southwest (1.52). High wind speeds associated with the former class and monsoon conditions associated with the latter class both might promote greater seas salt contributions.
- *Aged Sea Salt*: SCEs from aged sea salt are insensitive to the synoptic weather pattern, with the cluster-specific average scaled SCEs ranging from 0.83 to 1.14 only. It is possible this establishes a background condition for sea contributions to the Hong Kong aerosol.

The air mass pattern assignments are particularly useful to examine temporal trends in

the air quality conditions. For each PMF-resolved factor, annual median SCEs were calculated from the aforementioned composited general station time series after stratifying the SCEs by air mass cluster assignments. For each factor and air mass pattern, the annual median SCEs were regressed on year using a linear-least squares regression. The slope of this regression is a linearized estimate of the rate of SCE change per year over the period 2000-2008. As an example, **Figure 5.13** shows the annual median values and regression line for the Slow ECC air mass transport patterns and the motor vehicle and secondary sulfate factors. For air masses arriving in Hong Kong from the eastern portion of mainland China (indeed, as will later be demonstrated, for all air mass transport patterns), the general station vehicle exhaust SCE has decreased over the past decade (Figure 5.13). In contrast, for this same air mass transport pattern there has been a commensurate increase in the secondary sulfate SCE (Figure 5.13).

**Table 5.1** shows the linearized rate of change in the annual median SCE (i.e. Regression slope) by factor and air mass transport pattern. Values are presented only for those cases where the change is larger than its 95% confidence interval. Vehicle exhaust SCEs have decreased at the general stations regardless of air mass transport pattern. For example, for the Slow ECC cluster, the median scaled SCEs have been decreasing 0.12 per year, which corresponds to the absolute concentration decrease of  $0.71\mu\text{g}/\text{m}^3$  per year. This is consistent with vehicle emission control programs being implemented in the HKSAR. The rate of change does vary across the air mass transport patterns; this likely reflects the coupling between such transport patterns – which are broadly indicative of synoptic weather – and the local atmospheric ventilation conditions. The observed changes for the crustal/soil and smelting factors are small and likely inconsequential. The rate of change in the secondary sulfate factor is statistically indistinguishable from zero for the stagnant/circulating, south/southwest, and easterly air mass transport patterns. For air masses arriving from the eastern portion of China, however, there has been a significant increase in the secondary sulfate SCEs. Indeed, the increases in the secondary sulfate SCEs are nearly identical to the decreases in the motor vehicle SCEs for these two air mass transport patterns. Given that these two patterns are the prevailing conditions for most of the year, this analysis clearly demonstrates that **over the past decade, the reductions in urban-scale motor vehicle contributions to  $\text{PM}_{10}$  have been offset by transport of secondary sulfate (and associated PM) from mainland China.**

### 5.5.3 Results for $\text{PM}_{10}$ TEOM Mass Concentration Data

A similar analysis was performed on the 2000-2008  $\text{PM}_{10}$  TEOM data. While it lacks linkage to emission source categories that is embedded in the PMF-resolved factors, the daily data completeness is sufficiently high that patterns can be examined on a station-by-station basis. The hourly  $\text{PM}_{10}$  TEOM data for fourteen stations from 1998-2008 was conditioned by inputting missing values for single or two consecutive hours by linear interpolation of the adjacent reported hourly values across such data gaps. Monte Carlo simulations performed on the entire data set using a

jackknife method demonstrated that the errors from such imputing had insignificant impact on the calculated daily averages. After imputing the missing values, daily-average concentrations were calculated for those days with 24 one-hour values. Completeness for 2000-2008 was 84-95% depending on the site. For each station, annual median  $PM_{10}$  mass concentrations were calculated after stratifying the daily-average concentration values by air mass cluster assignments. For each site and air mass pattern, the annual median  $PM_{10}$  mass concentrations were regressed on year using a linear-least squares regression. The slope of this regression is a linearized estimate of the rate of  $PM_{10}$  mass concentration change per year over the period 2000-2008. As an example, **Figure 5.14** shows the annual median values and regression line for the Slow ECC air mass transport patterns and the CB roadside station (Figure 5.14) and YL general station (Figure 5.14). For air masses arriving in Hong Kong from the eastern portion of mainland China,  $PM_{10}$  mass concentration at the CB roadside station has decreased at a linearized rate of  $1.2 \mu\text{g}/\text{m}^3/\text{year}$  but this change is statistically indistinguishable from zero at the 95% confidence level ( $2.6 \mu\text{g}/\text{m}^3/\text{year}$ ). In contrast, the  $PM_{10}$  mass concentration at the YL general station has increased at a linearized rate of  $2.6 \mu\text{g}/\text{m}^3/\text{year}$  and this change is statistically indistinguishable from zero at the 95% confidence level ( $2.1 \mu\text{g}/\text{m}^3/\text{year}$ ).

**Table 5.2** shows the linearized rate of change in the annual median  $PM_{10}$  mass concentration (i.e. Regression slope) by factor and air mass transport pattern. Values are presented only for those cases where the change is larger than its 95% confidence interval. For the three roadside stations (MK, CL, and CB) and three of the general stations in Kowloon or on Hong Kong Island (SSP, KT, EN), the  $PM_{10}$  mass concentrations have decreased for one more of the air mass transport patterns of stagnant/circulating conditions, south/southwesterly flow, or easterly flow. In all six of these cases the changes have been insignificant for air mass patterns from the eastern portion of mainland China (Slow ECC and Fast ECC). This behavior is consistent with vehicle emission reduction impacts being directly observed for former case but not observed in the latter cases because the reductions from vehicle emissions (Figure 5.13b) are being offset by increased transport from mainland China (Figure 5.13a). At the northern end of the HKSAR, there has been an increase in  $PM_{10}$  of  $2\text{-}3 \mu\text{g}/\text{m}^3/\text{year}$  at YL and TM for air mass patterns from the eastern portion of mainland China (Slow ECC and Fast ECC). In most cases the changes have been statistically insignificant for the other three air mass transport patterns.

#### 5.5.4 Air Mass Transport Pattern Analysis Summary

This analysis has demonstrated that  $PM_{10}$  from vehicle exhaust has decreased over the period from 2000 to 2008. While this pattern was demonstrated in Section 4, the air mass patterns analysis confirms that statistically significant reductions have been observed for all five air mass transport clusters which reinforces this is a local (e.g. Urban-scale) effect. This analysis has also demonstrated that while motor vehicle emissions have decreased, HKSAR is being impacted by an increase in outside transport and this is specifically associated with secondary sulfate (and associated PM)

carried by air masses from the eastern portion of mainland China. The analysis lacks the spatial resolution necessary to determine the relative contributions of such impacts from emission sources in the Pearl River Delta as compared to emissions sources along the entire east coast region of mainland China.

Analyses performed on the PMF-resolved SCEs composited across all general stations and PM<sub>10</sub> TEOM mass concentrations at individual stations provide consistent patterns and complementary information. PM<sub>10</sub> mass concentrations have decreased at the more traffic-influenced sites for air mass patterns that are relatively stagnant, from the south/southwest, or from the east. The changes at these sites are insignificant for air mass patterns from the eastern portion of mainland China because the vehicle emission reductions are being offset by increased transport for these conditions. Stations located in the New Territories are strongly impacted by transport from mainland China; this suggests that at least some of the contributions are from PRD-scale sources although transport beyond the PRD might significantly contribute to – or even dominate – the observed increases. At these sites, in most cases the changes have been insignificant for the other three air mass transport patterns.

It should be noted that, back trajectory analysis for PM<sub>2.5</sub> dataset was not conducted because PM<sub>2.5</sub> samplings were only conducted at three campaigns with the sampling dates for the three sampling stations the same. Therefore, the total number of date with PM<sub>2.5</sub> speciation information was much smaller than the PM<sub>10</sub> dataset. Besides, 18% of the dates were filtered out for different clusters in different hours, while the dates with overlapping back trajectory and PM<sub>2.5</sub> speciation were further reduced. Such a small number of valid samples would definitely introduce large significant uncertainty to the results. As such, back trajectory analysis was only conducted for the PM<sub>10</sub> dataset.

## **5.6 A Preliminary Separation of PRD and Non-PRD Source Contributions using Air Mass Back Trajectories**

There has been a growing weight of evidence that PM air quality conditions in Hong Kong are significantly influenced by emission sources operating on three spatial scales – sources within the HKSAR, sources within the greater Pearl River Delta region, and more distant sources including but not limited to the East Asian Continent. In response the need to quantify PM impacts from these three scales – operationally defined as local, PRD and non-PRD sources – a preliminary analysis was conducted using PM<sub>10</sub> TEOM data and air mass back trajectories. This section summarizes the methodology and results from this preliminary analysis.

### **5.6.1 Data and Methodology**

Seven-day air mass back trajectories for Central / Western (CW) station in general AQMS during 2006 generated by Perry Samson for USEPA

(<http://www.world.sharedair.org/data/>) were used in the analysis to represent the overall situation in Hong Kong, as the spatial variation and inter-annual variation amongst Hong Kong is insignificant. For each day in the year of 2006, four trajectories were generated corresponding to arrival times in CW of 00, 06, 12 and 18 hours UTC. The trajectory arrival height was fixed at 500m and back trajectory endpoints (latitude, longitude, elevation) were generated for every two hours up to seven days duration. Hourly PM<sub>10</sub> data measured by TEOM at the CW station were used in this study. Surface winds for 2006 are not available for the CW site and thus the surface winds data measured at the Hong Kong Observatory were used to represent wind conditions at CW.

For each air mass back trajectory the following methodology was used to classify the dominant spatial scale of emissions transport.

1. The air mass back trajectories were trimmed to seven days (168 hours) prior to arrival in Hong Kong. In some cases, the air mass back trajectories truncated prior to 168 hours and trajectories were deemed valid if they had at least three days (72 hours) of endpoint data.
2. "PRD" transport days were defined as those air mass back trajectories that: (i) resided within the square area of 22.5-23.5°N and 112.0-115.0°E (Area A, roughly corresponding to the PRD) and below heights of 1500m (roughly the maximum daytime boundary layer height in the PRD) for at least 6 consecutive hours (i.e. six consecutive endpoints); and (ii) resided within the square area of 22.3-50.0°N, 100.0-140.0°E (Area B, roughly corresponding to Eastern Asian Continent where most of the anthropogenic emissions are concentrated) and below heights of 3000m for no more than 24 consecutive hours.
3. "NON-PRD" transport days were defined as those air masses back trajectories that: (i) resided in Area A and below heights of 1500m for less than 6 consecutive hours; and (ii) resided in Area B and below heights of 3000m for at least 6 consecutive hours.
4. Air mass back trajectories residing in Area A at heights of 1500m for at least 6 consecutive hours and residing in Area B at heights below 3000m for at least 6 consecutive hours were assumed to be potentially influenced by both PRD and non-PRD sources. In principle the relative contributions could be constructed from a weighted average of the time spent in both areas. For this analysis, however, this kind of back trajectories are all defined as either "PRD" or "NON-PRD" to provide the contribution bounds of the two trajectory clusters.
5. "LOCAL" transport days were defined as those air mass back trajectories that did not reside in the square areas of 22.3-50.0°N, 100.0-140.0°E or 10.0-22.3°N, 100.0-113.5°E (roughly corresponding to Indo-China Peninsula) for even one endpoint.
6. "UNDETERMINED" transport days were those air mass back trajectories that were not assigned as PRD, NON-PRD or LOCAL using the above classification criteria.

The areas designated to define PRD, NON-PRD and LOCAL trajectory classes are

illustrated in **Figure 5.15**.

As discussed in the following sections, this analysis is associated with significant level of uncertainties. The uncertainties are generated inherent in the calculation method. Specifically, it is derived from lacking the data in the PRD or larger scope to manage to perform a scale transport analysis for the small Hong Kong. Extending the analysis for the entire period would give readers wrong impression of adequate statistical significance of this analysis. Therefore, only the data in 2006 is applied in this analysis in a purpose of demonstration. We expect this analysis would be more valuable with the use of high-resolution meteorological simulations that are verified with local and regional observational data. This is beyond the scope of the current work, but should be carried out in further studies.

## 5.6.2 Results

### *Mapping of Hourly PM<sub>10</sub> Data at TW onto Classified Scales of Spatial Transport*

Hourly PM<sub>10</sub> data collected at TW in 2006 were mapped onto each air mass back trajectory that was classified as local, PRD, or non-PRD. TW samples were used for the availability for both PM<sub>10</sub> and PM<sub>2.5</sub> records. Arithmetic mean contributions were calculated for each combination of transport scale and PM<sub>10</sub> concentration. **Table 5.3** shows the averaged PM<sub>10</sub> loadings and number of occurrence for different trajectory classes. Here the definition of summer and winter is the same as Chapter 3, i.e. summer from 16 May to 15 September and winter from 1 January to 15 March and from 16 November to 31 December. It can be seen from Table 5.3 that NON-PRD trajectory class shows higher average PM<sub>10</sub> concentration in winter than in summer; PRD trajectory class shows the opposite; LOCAL trajectory shows almost equal concentration between the two seasons. The average ground-level wind speed in summer and winter are roughly the same but the direction is certainly of large difference (Figure 5.3 as an example). Another important observation is that LOCAL trajectory class is mainly concentrated in summer while PRD and NON-PRD class are mainly in the winter. Not even a single case of LOCAL trajectory class occurred in winter and therefore the associated average PM<sub>10</sub> concentration is estimated by extrapolation of the annual average and the summertime average.

It is important to note that the concentration in Table 5.3 are the average PM<sub>10</sub> concentration bounds associated with each trajectory class and not the contributions from local, PRD and non-PRD sources. That is, for a given trajectory-concentration pair all that concentration is assigned to only one class – LOCAL, PRD, or NON-PRD. Actually, the PM<sub>10</sub> concentrations in the LOCAL class, whose trajectories are mainly from the south, are the sum of local emissions within Hong Kong and PM originated from the ocean, with sea salt being the most significant fraction. A typical percentage of 30% is assigned to sea salt contribution, as determined from source apportionment analysis. The other 70% is thus assigned to local emissions within Hong Kong. Again, it should be noted that when the PRD air mass transport pattern

prevails, the source contributions are actually the sum of “local emission” and “PRD impact”. Similarly, when the NON-PRD air mass transport pattern prevails, the source contributions are actually the sum of “local emission” and “Non-PRD impact”. The local emissions contributions under different transport classes may not be the same, due much to the different wind speeds thus dispersion capacities associated. Therefore, the wind speed data were used to adjust the local contribution under different meteorological conditions. It is assumed that local source contribution is inversely proportional to the prevailing wind speed, when the emissions are constant.

With the assumptions above and the number of occurrences of each trajectory class in each season, the contributions from local, PRD and non-PRD sources can be estimated for each season. **Table 5.4** lists the absolute contributions (in unit  $\mu\text{g}/\text{m}^3$ ) and contribution percentages of local, PRD and non-PRD sources in summer and winter, and in the entire year of 2006. Table 5.4 demonstrates several interesting patterns. Local sources show equal contribution in winter (33~36%) and summer (34%), consistent with Figure 3.3. PRD sources show higher contribution in summer, while non-PRD sources show higher contribution in winter. Generally, Hong Kong’s prevailing wind from the east or northeast favors the impact of non-PRD transport. Under the general meteorological conditions, it is not common that the air mass can linger in the Pearl River Delta for more than 6 hours before reaching Hong Kong. A significant mechanism for such a transport pattern is associated with the approaching tropical cyclone that often occurs in summer. The tropical cyclone can divert the background wind to weak northwesterly, bringing significant pollutant burdens from the Pearl River Delta to Hong Kong. The two most significant non-local impact cases in the summer 2006 are associated with tropical cyclone KAEMI and BOPHA on 24 July and 9 August, respectively. Non-PRD transport shows higher impact in the winter (45-67%). It can be discovered from this analysis that most of the  $\text{PM}_{10}$  pollutants in Hong Kong are transported from outside Pearl River Delta area. The contribution percentages are over half of the total  $\text{PM}_{10}$  loading at TW.

#### Mapping of Hourly $\text{PM}_{10}$ Data at CW onto Classified Scales of Spatial Transport – High Mass Concentration Hours

Here we define the high mass concentration hours into two scenarios. Scenario 1 is defined as the hours with  $\text{PM}_{10}$  concentration higher than or equal to  $100 \mu\text{g}/\text{m}^3$ , the Interim target-2 for  $\text{PM}_{10}$  recommended by WHO, or with  $\text{PM}_{2.5}$  concentration higher than or equal to  $75 \mu\text{g}/\text{m}^3$ , the Interim target-1 for  $\text{PM}_{2.5}$  recommended by WHO. Scenario 2 is defined as the hours with top 10%  $\text{PM}_{10}$  or  $\text{PM}_{2.5}$  concentration.

**Table 5.5** shows the averaged  $\text{PM}_{10}$  loadings and number of occurrences for different trajectory classes in scenario 1. It is interesting to note that the average concentrations from PRD and NON-PRD trajectory classes in summer and winter are comparatively consistent. Again, as there is no occurrence of the LOCAL trajectory class during high mass concentration hours, the local source contributions during high mass concentration hours are estimated by a wind speed-proportional approach,

similar to the estimation of PRD and non-PRD contributions illustrated in Table 5.4.

**Table 5.6** shows the absolute contributions (in unit  $\mu\text{g}/\text{m}^3$ ) and contribution percentages of local, PRD and non-PRD sources in scenario 1. Since the absolute contribution of different trajectory classes are directly related to the number of occurrence in a particular season or year, it can be seen from Table 5.6 that on an annual average, non-PRD transport is the dominant contributor to  $\text{PM}_{10}$  at TW in 2006. From a percentage-wise point of view, PRD and non-PRD contributions present notable increase as compared with the general case as shown in Table 5.4 (PRD increased from 4~20% to 2~31% and non-PRD increased from 52~66% to 55~85%). In some pollution episodes in summer and winter, PRD contribution increased significantly (from 8~38% to 16~81% in summer and from 0~19% to 0~37% in winter), while the contribution percentages of local sources decreased (from 35% to 10% in summer and from 33~36% to 24~29% in winter). There is no discernible change of Non-PRD impacts during high PM hours.

**Table 5.7** shows the averaged  $\text{PM}_{10}$  loadings and number of occurrences for different trajectory classes in scenario 2. **Table 5.8** shows the absolute contributions (in unit  $\mu\text{g}/\text{m}^3$ ) and contribution percentages of local, PRD and non-PRD sources in scenario 2. It can be seen that non-PRD transport is still the dominant contributor to  $\text{PM}_{10}$  at TW in 2006. From a percentage-wise point of view, PRD contribution increased significantly (from 8~38% to 13~63% in summer and from 0~19% to 0~38% in winter), while the contribution percentages of local sources decreased (from 35% to 14~15% in summer and from 33~36% to 23~28% in winter). There is no discernible change of Non-PRD impacts during high PM hours.

In summary, although the contributions from PRD sources on normal days are relatively low, they increased significantly during some high mass concentration episodes. Non-PRD sources are always important contributors to the  $\text{PM}_{10}$  levels at TW, regardless of on normal days or during high mass concentration episodes. Local sources, on the other hand, are comparatively more significant during high mass concentration periods in wintertime.

#### *PM<sub>2.5</sub> Analysis*

**Table 5.9-5.14** show the corresponding information of  $\text{PM}_{2.5}$ . It can be discovered that the contribution percentages of local, PRD and non-PRD sources are very similar to those for  $\text{PM}_{10}$ , as listed in Table 5.3-5.8. This shows that in Hong Kong,  $\text{PM}_{2.5}$  and  $\text{PM}_{10}$  almost share the same conceptual model. In other words, pollutants associated with local, PRD and non-PRD sources are almost with the same  $\text{PM}_{2.5}/\text{PM}_{10}$  ratio. From previous discussions, sea salt aerosols are likely the dominant species for coarse particles when LOCAL trajectory class is dominated, while crustal soil / dust aerosols are likely the dominant species for coarse particles when PRD or NON-PRD trajectory class is dominated. Although coarse particles may have larger deposition rate than fine particles, such a difference is negligible for

the same size of HK territory.

### Comparison with 3-D Air Quality Modeling

As a way of quality assurance / quality check on the trajectory-based analysis, we also conducted some preliminary source apportionment analysis with the Particulate Source Apportionment Technology (PSAT) module in the CAMx modeling system, using the dataset from CW AQMS. The analysis period was the entire month of November 2006. The absolute contributions of sources from the PRD and non-PRD transport are illustrated in **Figure 5.16**. The conclusion is that in the entire month of November 2006, local sources contribute 10-20% of total PM<sub>10</sub> in Hong Kong, PRD sources 25-35%, and non-PRD sources 45-65%. During pollution episodes, however, local sources contribute 15-25%, PRD sources 35-55%, and non-PRD sources 20-35%. The back trajectory analysis results are generally consistent with the 3-D air quality modeling results.

### **5.6.3 Discussion and Recommendations**

Accurate separation of contributions from local, PRD and non-PRD sources to the air pollutants in Hong Kong is an important and challenging topic, as they are associated with distinct control strategies. In this project, back trajectory analysis was applied as a simple but rough tool to separate PM<sub>10</sub> level at CW in the year of 2006. Although reasonable results are reached from the back trajectory analysis, the back trajectory approach is associated with significant uncertainties, as detailed below.

- The designated area may not be representative of the source impact areas. In this study, two square areas are prescribed to represent the PRD and Eastern Asian Continent, where most of PRD and non-PRD sources are located. However, the area with source impact may be different from the emission area. For example, in the wintertime, there is very often a high pressure system dominating the Eastern Asian Continent. As a result, pollutant emitted over the continent can be readily transported hundreds or even thousands of kilometers into the ocean. Therefore, air mass originated from the ocean may also have chance to bring about pollutants that were originally emitted from the Continent. This would lead to an overestimate of local sources in the wintertime.
- The wind field of back trajectory analysis is with the resolution of 1°x1°, which means probably the entire PRD are only covered by a small number of wind field grids. Such a coarse resolution would miss some important micro-meteorology characteristics inside the PRD, such as land-sea breeze circulation, mountain valley wind circulation and urban heat island effect, which are generally more prominent during pollution episode. As a result, the time of air mass lingering within the PRD will be underestimated and impact of PRD sources will be correspondingly underestimated. The impact of long-range transport, on the contrary, will be overestimated.
- Each back trajectory is subjectively classified as one of LOCAL, PRD and NON-PRD classes. After the classification, the source classes impacting the

station are fixed. In other words, there would not be PRD sources contributing to the samples with LOCAL or NON-PRD trajectories. Similarly, there would not be non-PRD sources contributing to the analysis time with LOCAL or PRD trajectories. Such an assumption may not be reasonable in the real situation. Large number of analysis samples is needed to minimize the associated discrepancies. It is therefore of relatively larger uncertainty for some short-period analysis, e.g. seasonal analysis, annual trend analysis or episodic analysis.

- The number of samples available is too few. For example, we cannot find any high mass hours in 2006 associated with pure “PRD” trajectory, i.e. Trajectory passing only Area A but not Area B. This would introduce large uncertainty in estimating PRD contributions during high mass hours. In future work, this back trajectory analysis should be extended to cover the entire sampling decade to get more robust results and temporal trends in association with each trajectory clusters. In addition, wind field with higher spatial resolution (down to kms) should be generated by meteorological model and applied in the back trajectory analysis.

Apart from the present analysis, we also propose the “SO<sub>2</sub>-sulfate conversion” and “remote sensing” methods based on back trajectory analysis that may serve as alternative approaches for the source separation. In the “SO<sub>2</sub>-conversion” method, sample-specific 'retention time' within the PRD can be derived for each back trajectory, and a typical threshold rate of SO<sub>4</sub><sup>2-</sup>/SO<sub>2</sub> can be estimated, in consideration of the general oxidation capacity in the regional atmosphere. The observed SO<sub>4</sub><sup>2-</sup>/SO<sub>2</sub> ratio higher than the threshold ratio can be deemed as with non-PRD contribution, and its portion can be estimated. In the “remote sensing” method, we may overlay the back trajectories on the satellite images and estimate the annual variation of PRD (AOD integration along the trajectories within PRD) and non-PRD (AOD integration along the trajectories outside PRD) contributions, in consideration of frequencies of back trajectories in each year and probably a dry deposition-associated removal factor. Then, this relative portion can be applied on the estimation of non-local sources by PMF analysis (which actually is the sum of PRD and non-PRD) to estimate the actual contribution of PRD and non-PRD sources.

Certainly, each analysis approach has its own advantages and limitations, and the explicit analysis details need to be refined. Therefore, it is suggested these approaches are conducted in parallel, and the results can be mutually validated. This is certainly well beyond the scope of the present study, but it is an interesting and meaningful study to improve our understanding of source contributions from different scales to the pollution level in Hong Kong. Furthermore, when resources are available, our source apportionment study based on 3-D air quality models can be extended with larger temporal coverage to provide more piece of information for results validation.

## 5.7 A Compact Summary of the Conceptual Model for Ambient Particulate Matter over Hong Kong

With the information summarized above, the conceptual model for Hong Kong can be visualized as **Figure 5.17** under general conditions and **Figure 5.18** for Scenario 1 and **Figure 5.19** for Scenario 2 during high PM hours. Again, it is noted that this description is based on analyses that could be performed with the available data and must be refined as new knowledge becomes available.

Under general conditions, the prevailing wind is from the northeast in winter and southeast in summer. If a transect is made from NE to SW across HK, the contributions of PRD and non-PRD sources should be consistent. Local sources, on the other hand, are mostly concentrated in the urban areas. The relative contribution fractions of sources in three different scales are determined by the back trajectory analysis discussed in section 5.6.

During high PM hours, the background wind during summer and winter very often shifts to northwesterly with relatively weak wind speed. In view of this, the transect of interest is shifted from NE-SW to NW-SE. As expected, non-PRD contribution is still spatially consistent for its larger scale. However, PRD contribution increases significantly with an obvious NW-SE spatial gradient. In particular, land-sea breeze circulation develops significantly during weak background NW wind which acts as an efficient mechanism to bring the PRD pollutants into HK. As this land-sea breeze circulation can only penetrate HK around 10-20 km and with the blockage effect of the mountain in the centre of HK, the PRD sources pose greater impact to the western HK than the eastern HK. Therefore, we may expect the PRD impact is spatially stratified. Moreover, with the stabilized mixing layer, the contributions of local sources during high PM days in winter are significantly higher than that in summer. Again, the relative contribution fractions of sources from the three different scales are determined by the back trajectory analysis discussed in previous section, which may change when the uncertainty of back trajectory analysis reduced. Moreover, as discussed previously, this conceptual model applies for both  $PM_{10}$  and  $PM_{2.5}$ .

## 6. Conclusions and Recommendations

This project has identified the emission source categories exerting influence on Hong Kong PM burdens and quantified their impacts including how these patterns have changed over the previous decade. Emphasis was placed on a weight-of-evidence approach using independent data sets and analyses to assess the reliability of key findings. This work has culminated in the development of a conceptual model for ambient PM over Hong Kong. A battery of analyses has been used to inform the development of this conceptual model for ambient PM over Hong Kong. While the major findings are not necessarily new revelations, these analyses significantly contribute to the weight-of-evidence underlying such findings and place them on a firmer scientific foundation. Major conclusions, policy-relevant implications and recommendations are as follows.

### 6.1 Key Findings

Ambient PM attributed to vehicular exhaust has significantly decreased over the period of this study. As discussed in section 3.2.2, 54% of vehicle exhaust was reduced in 2008 as compared to the level in 1998. The weight-of-evidence for this conclusion is strong and includes source apportionment results for PM<sub>10</sub> and PM<sub>2.5</sub> speciation network data using multiple source apportionment methods (e.g. PMF, Unmix, APCA, CMB), and an analysis of the PM<sub>10</sub> TEOM data that isolates site-specific excess PM mass from an area wide PM baseline. A nearly consistent rate of decrease in vehicle exhaust impacts, with the median absolute contribution decreasing at a rate of 0.6 µg/m<sup>3</sup> per year, was observed for all air mass transport patterns. Year-to-year changes in the vehicle exhaust PM<sub>10</sub> ambient burdens are linearly proportional to changes in the HKSAR transport sector PM emission inventory. Vehicle-related emission control strategies implemented over the previous fifteen years has resulted in a substantial positive impact by decreasing ambient burdens of – and thus exposures to – vehicle exhaust. In contrast to the late 1990s, vehicle exhaust is no longer the dominant emission source category contributing to PM<sub>10</sub> burdens at the general stations, with its contribution percentage decreased from 22% in 1998 to 10% in 2008. However, it remains among the most dominant source categories at the roadside stations and remains a significant contributor at the general stations. In addition, the advancement in the performance of internal combustion engines often yields more fine particles in greater numbers. This may well be bad news from a public health perspective. The European Union is now looking into the prospect of regulating ultrafine particles in 2013. Therefore, it is recommended that Hong Kong should endeavor to convert the existing vehicle fleet to low- or no-emission models to comprehensively address the vehicle emissions problem.

In this study, “local sources” were defined as those exhibiting spatiotemporal characteristics similar to those expected for sources emitted from within the Hong

Kong territory, while “non-local sources” referred to those with spatiotemporal characteristics across the HK network that are consistent with patterns expected from various scales of regional transport. We found that ambient PM attributed to transport has significantly increased over the period of this study. The weight-of-evidence for this conclusion is strong and includes source apportionment results for  $PM_{10}$  and  $PM_{2.5}$  speciation network data using multiple source apportionment methods (e.g. PMF, Unmix, APCA, CMB), and an analysis of the  $PM_{10}$  TEOM data that isolates an areawide PM baseline. In the source apportionment modeling the relevant emission source categories include secondary sulfate and secondary nitrate. The case is compelling for secondary sulfate being overwhelmingly from PRD and/or non-PRD sources. Spatiotemporal patterns in the source apportionment modeled secondary sulfate factor were nearly identical across the HKSAR with no discernible difference between the general stations and the roadside station.  $PM_{10}$  source contributions from secondary sulfate increased over the ten year study period only for air masses from the north with negligible change over the decade for air mass transport patterns from other directions. Similarly,  $PM_{10}$  TEOM data stratified by air mass transport patterns increased for air masses from the north and remained unchanged or decreased for virtually all sites when the air mass transport patterns were from other directions. Secondary nitrate appears to have both local and non-local contributions; the precise split is currently unknown. Overall, the increases in secondary sulfate and nitrate emission source categories have virtually offset the decreases in vehicle exhaust impacts. The relative role of the emissions from Guangdong Province versus the emissions from more distant regions of mainland China – and especially the entire east coast – is not known and this is an important area for future research. However, it is clear that future air quality conditions in Hong Kong will depend upon cooperation from provinces in mainland China; such cooperation is already underway with Guangdong Province and perhaps a coordinated, regional-scale air quality management plan will best serve both the HKSAR and neighboring provinces. There exist models for such coordination such as the Ozone Transport Commission (OTC) in the Northeastern United States.

Source apportionment modeling suggests that PM from residual oil combustion has increased over the study period. These emissions are attributed to marine vessels and this increase is generally consistent with increased marine activity in the greater PRD region over the past decade. While the PM burdens from this source category appear to be relatively small, they are not negligible and, furthermore, emissions from residual oil combustion tend to be rich in air toxics metals. From a public health standpoint there is substantial motivation to better characterize marine-related activities and emissions and consider the implementation of air quality management plans targeting this sector. This will require cooperation with neighboring provinces to be most effective.

Given the importance of outside transport, a preliminary analysis was conducted to estimate relative contributions from local sources (i.e. within the HKSAR), PRD sources, and non-PRD sources including but not limited to the East Asian Continent.

Specifically, within the “non-local” source sector, we also classified contributions into “PRD” and “non-PRD” according to the projected geographic locations of the emissions. Air mass back trajectories were classified into these three spatial categories based on their presence and residence time in these geographic zones. PM<sub>10</sub> mass data from 2006 was used for this study, with most analyses performed on the Central Western site data. During the summertime, which is dominated by monsoon conditions, PM<sub>10</sub> contributions are greatest from non-PRD sources (27~56%), followed by local (35%) and PRD sources (8~38%). For the scenario of high PM hours in the summertime, however, PRD contributions (16~81%) increase with a corresponding drop in local contributions (10%). In the wintertime, contributions are greatest from transport beyond the PRD (45~67%), followed by local contributions (33~36%) and finally PRD contributions (0~19%). On high PM days in the wintertime, however, the contributions are quite different with PRD (0~37%) and non-PRD (33~76%) transport dominating in some cases with slightly reduced local sources (24~29%). Similar percentages were derived for PM<sub>2.5</sub> contributions, suggesting PM<sub>10</sub> and PM<sub>2.5</sub> behave in a conceptually similar way in Hong Kong.

Estimated relative source contributions across the three different spatial scales are generally consistent with the case study in November 2006 by CAMx air quality model yet it is believed the results have considerable uncertainty, in large part because there were a limited number of samples available and the coarse resolution of wind fields to generate the trajectories. In future work, this back trajectory analysis should be extended to cover the entire sampling decade to get more robust results and temporal trends in association with each trajectory clusters. In addition, wind fields with higher spatial resolution (on the scale of a few kilometers) should be generated by meteorological modeling and applied in the back trajectory analysis.

*Policy Implications.* The results of this study demonstrate the need for an air quality management strategy that targets emission reductions on the local, PRD, and non-PRD scales to comprehensively address the ambient PM air quality challenge in Hong Kong. While there is strong evidence that prior local emission control programs have been effective, local emission sources – especially vehicle emissions – are still significant and further vehicle emission reductions within the Hong Kong SAR would be beneficial. There is also evidence for significant contributions from regional- and larger-scale pollution transport. Our current state-of-knowledge suggests emission reductions from within the PRD would benefit Hong Kong especially on high PM days. Emission reductions from more distant sources such as the eastern portion of mainland China would also benefit Hong Kong in particular by reducing concentrations for persistent wintertime conditions that lead to relatively high annual-average concentrations.

## 6.2 Recommendations

Several recommendations have already been made in this report concerning the data quality assessment (section 2.2.4), source apportionment modeling (sections 3.3.6,

3.4.4, and 4.7), and transport scales analysis (section 5.6.3). Additional general recommendations are provided in this section.

Effective air quality management and planning requires appropriate data collection and analyses. The HKSAR benefits from a relatively dense network of PM<sub>10</sub> TEOM monitors and speciation samplers that have been operated for more than a decade. Together with high quality analytical approaches, the resulting data set is unprecedented in its spatial and temporal coverage for a region the size of the HKSAR. That said, there are some monitoring considerations that have policy-relevant implications.

- First, there is growing attention worldwide on PM<sub>2.5</sub>. HKEPD has shown foresight by conducting a series of one-year PM<sub>2.5</sub> speciation studies at three-to-four sites over the past decade. We are pleased to recently learn that the PM<sub>2.5</sub> network in Hong Kong has been expanded to match the PM<sub>10</sub> network. We envision that the parallel measurement of PM<sub>10</sub> and PM<sub>2.5</sub> will provide the opportunity to gain significant insights in terms of pollutant emission, formation, transport and depletion processes in Hong Kong and the PRD.
- Second, even if there is greater emphasis on PM<sub>2.5</sub> measurements in the future, a robust network of PM<sub>10</sub> speciation sampling and continuous mass monitoring should be sustained to continue the time series for assessing long term trends. While there might be tension to change the PM<sub>10</sub> measurement protocols to reduce artifacts, the effect of such changes on the continuity of the data time series should be carefully considered. If an approach is taken that expands PM<sub>2.5</sub> measurements at the expense of PM<sub>10</sub> measurements, several of the analyses underlying this conceptual model could be used as part of a network assessment to determine optimal configurations for a given level of resources.
- Third, the analyses underlying this conceptual model have identified several key data gaps. The current Supersite-inspired studies should substantially contribute to many of these key science questions. Additional issues include a need for vertical profiling of PM concentrations, a more-detailed characterization of carbonaceous aerosol and a better understanding of sea salt dynamics.
  - Many areas in the HKSAR are strongly influenced by land/sea breezes which can confound the interpretation of ground-levels measurements, especially when integrated over periods such as 24 hours. There is evidence of land/sea breeze effects in the PM<sub>10</sub> TEOM data diurnal profiles. Vertical information would help to elucidate whether elevated concentrations during onshore winds conditions result from material being previously blown offshore and then subsequently recirculating back onshore, versus onshore winds bringing in material from other regions.

- A shortcoming of the source apportionment modeling was the inability to clearly resolve secondary organic aerosol (SOA). This is typically the case when all of using bulk measures of carbon such as EC and OC. Source apportionment of speciated organics holds much promise to quantify contributions from secondary organic aerosol. While such data has been collected in Hong Kong for special studies, at this time it is prohibitively expensive to be integrated into a routine monitoring network. Consideration should be given on ways to best integrate such information, including periodic special studies conducted at existing network stations, to better interpret the routine speciation monitoring data including constraints on the source contribution estimates obtained from source apportionment modeling. Perhaps a conceptual model should be developed specifically for SOA over Hong Kong to inform what measurements would be most useful in the future.
- Sea salt is a significant contributor to PM<sub>10</sub> burdens over Hong Kong. Source profiles and temporal patterns for sea salt source contribution estimates strongly suggest the sea salt is being aged to varying degrees. In this case, it cannot be considered a purely natural emission source. More work is needed to refine the understanding of this aging process and its impact on sea salt composition. The Supersite continuous measurements of inorganic ions might provide insights.

A concerted effort was made to critically examine the source contributions estimates generated by the receptor modeling. This weight-of-evidence approach can further be strengthened through additional data collection (e.g., as noted above) and analyses. Chemical transport modeling (CTM) will be a very important addition towards informing the conceptual model because it draws quantitative relationships between emissions and observed impacts. Source apportionment modeling and other data analyses can be useful diagnostic tools for evaluating CTM performance and filling in knowledge gaps when the CTM models are deemed inadequate. The robustness and interpretation of the receptor modeling results can be further evaluated. For example, the PM<sub>10</sub> and PM<sub>2.5</sub> TEOM data can be used to estimate coarse PM (PM<sub>10-2.5</sub>, or PMc) concentrations which we would expect to be highly correlated with the crustal soil/dust factor. Relationships between PMc and source contribution estimates for the resolved factors should be examined. The loading of OC onto the secondary nitrate factor is consistent with the source of this carbon being SOA. However, it is also possible that the volatile ammonium nitrate co-varies with that portion of the OC that is also semivolatile, regardless of whether that carbon is of primary or secondary origin. More work is needed to understand how SOA is distributed across the resolved factors and whether the OC associated with the secondary nitrate factor is truly SOA or merely semivolatile OC.

### 6.3 The Path Forward

As Hong Kong embarks on implementing new Air Quality Objectives (AQOs) including its first AQOs for  $PM_{2.5}$ , considerable effort need be invested in the development of a robust framework to both monitor compliance with the AQOs and provide the data needed to strengthen the linkages of air quality planning and management to the underlying science. Numerous lessons can be learned from HKEPD's historical approach to PM monitoring as well as approaches taken overseas. This section predominantly focuses on the U.S. experience. In 1997 the USEPA promulgated revisions to the U.S. National Ambient Air Quality Standards (NAAQS) to include standards for  $PM_{2.5}$ . Thus, there is nearly fifteen years of experience implementing this standard.

*PM<sub>2.5</sub> NAAQS Level and Form.* Current daily- and annual-average  $PM_{2.5}$  NAAQS are 35 and 15  $\mu\text{g}/\text{m}^3$ , respectively. An important nuance is the precise calculations that lead to the "design values" used to determine compliance. In addition to minimum criteria for data completeness, the following calculations are applied at the monitoring station level.

- Daily-average  $PM_{2.5}$  NAAQS design values are determined by first calculating the annual 98<sup>th</sup> concentration for each year. The design value is the three-year average of these annual 98<sup>th</sup> percentile concentrations.
- Annual-average  $PM_{2.5}$  NAAQS design values are determined by first calculating the quarterly average of the daily concentrations and then averaging the four quarterly values to obtain an annual average concentration for each year. This approach equally weights the data collected in each calendar quarter and thus avoids bias from any differences in data completeness across the year; it is particularly important for sites that have significant seasonal variations in ambient concentrations where an unbalanced number of samples in each season could introduce bias in the annual average. The design value is the three-year average of these annual average concentrations.

These calculation methods are intended to suppress bias from temporal variations in data completeness and also suppress, to some extent, year-to-year concentration variations that are modulated by changes in weather patterns.

*PM<sub>2.5</sub> NAAQS Compliance Monitoring Methods.* The  $PM_{2.5}$  Federal Reference Method (FRM) is filter-based gravimetric method. It is a combination of design- and performance based specifications for sampler design, sampler operation, and sample handling and conditioning. The A Federal Equivalent Method (FEM) must meet certain design and performance specifications including demonstrated agreement with the FRM at a mandated level of accuracy. In recent years several semi-continuous monitoring methods have received FRM designations and thus can be used for compliance monitoring.

There are sound reasons to designate a reference method and require any alternative

methods to demonstrate an adequate level of equivalency to the reference method. However, the selection of a reference method requires careful consideration. The  $PM_{2.5}$  FRM has often been criticized but it must be appreciated that the intent was to tether the FRM to methods used for the health studies underlying the NAAQS. The FRM measurement does not represent the ambient  $PM_{2.5}$  mass concentration – wet or dry – but rather represents the  $PM_{2.5}$  mass concentration after the sample has been collected and conditioned according to a prescribed method. The method does control for aerosol water content which can otherwise dramatically affect the mass concentration for hygroscopic particles. Measurement artifacts include but are not limited to the partial loss of semi-volatile compounds such as ammonium nitrate and certain organic compounds. Thus, it is generally accepted that the FRM mass is often biased low compared to an ambient aerosol for which the water is removed without other losses. This bias depends on environmental conditions but the FRM does provide a standardized approach.

A conundrum arises in the design of semicontinuous monitors. Over the past decade there have been several instrument advances that lead to a more quantitative determination of the ambient PM mass concentration. However, to meet the requirements for FEM designation these instruments must be operated in a manner that mimics the bias in the FRM measurement. While this need for consistency is understandable, from some perspectives it can be viewed as not capitalizing on technology advances (at best) to being regressive (at worst).

To place the Hong Kong monitoring program in context, it first must be recognized that the 50°C TEOM methods currently used for  $PM_{10}$  and subsequently  $PM_{2.5}$  were truly state of the art at the time of network inception. It effectively removes aerosol water, albeit at the expense of also removing other volatile compounds. Indeed, it is recognized to be a very stable and robust measurement of nonvolatile PM mass. Measurement methods were subsequently developed that can more closely match the FRM and/or the ambient PM mass concentrations. However, adopting these newer methods could lead to significant discontinuities in the historical record. This is a fundamental tension in managing all long term networks – the desire to adopt the best available measurement methods yet not introduce discontinuities into the time series. *Perhaps the ideal monitoring strategy would include measurements for both total and nonvolatile PM with the former measurements used for determining compliance with the AQOs and supporting scientific investigations, and the latter used for preserving measurement method consistency within the long term record.* Assuming the goal is to meet both objectives, it must be determined whether this can be accomplished using a single monitor or would require multiple instruments. For example, the Thermo FDMS TEOM provides semi-continuous measurement of both the nonvolatile and volatile components of ambient PM mass. However, the split between volatile and nonvolatile mass is operationally defined and, at least for the units with FEM designation, the nonvolatile measurement is tuned to match the FRM rather than a measurement consistent with a 50°C TEOM.

*PM<sub>2.5</sub> Speciation Monitoring Methods.* The U.S. has two major PM<sub>2.5</sub> speciation monitoring networks – the Chemical Speciation Network (CSN) and the IMPROVE network. CSN is primarily focused on measurements in urban areas whereas IMPROVE monitors are primarily deployed in Class I areas such as national parks in support of the U.S. regulations on visibility and haze. The CSN network ideally meets several objectives such as providing data for developing air quality control programs and supporting health effects studies. A subset of the CSN – the Speciation Trends Network – meets the objective for long term monitoring with the remaining CSN sites operated at the discretion of state and local air quality agencies. It is beyond the scope of this summary to comprehensively articulate all of the considerations underlying the selection of sampling methods, chemical analysis methods, and network design. However, a few points deserve mention. First, the objectives of the speciation monitoring must be established. In the U.S., the speciation measurements actually play a formal and important role in the development of control strategies for areas violating the PM<sub>2.5</sub> NAAQS. For example, control strategy development for the annual standard requires chemical transport modeling using tools such as CMAQ or CAMx. The limitations of these models are recognized and thus the models are used to develop relative reduction factors for classes of particle components and the speciation data are used to convert these component-specific relative reductions to absolute PM mass concentration reductions. In this case, a goal is to have the speciation data match as closely as possible the mass collected on the FRM filter, including all of the sampling and measurement biases. For example, the CSN method uses a denuder and Nylon filter to provide a reasonably robust measurement of ambient PM nitrate and calculations are performed on these data to estimate the portion of the ambient PM nitrate that would be retained by the FRM Teflon filter. The CSN organic carbon measurements are prone to measurement artifacts although recent changes have been adopted to reduce the magnitude of the artifacts. Therefore, the organic matter retained by the Teflon filter is estimated by using the speciation data to estimate the mass on the filter from all components other than organic matter, and then assign to organic matter the remaining mass needed to close the mass balance on the gravimetric measurement. Clearly there are compromises between strategies that provide the best estimate of ambient PM components and those components actually present on the FRM filter. In this spirit, Hong Kong must decide the objectives for speciation monitoring and use these objectives to guide the selection of sampling hardware and analytical methods.

*QA/QC Programs.* Quality assurance and quality control are important elements of any monitoring program with a particularly high bar set for measurements with regulatory implications. Similar to the discussion of speciation monitoring, it is beyond the scope of this summary to comprehensively address the many QA/QC aspects and thus merely a few perspectives are provided. The adoption of more advanced sampling and monitoring methods will place an added burden on QA/QC. For example, many PM samplers have hardware that is much more complex than the high volume sampler and many semi-continuous PM monitors are more complex than

the 50°C TEOM. Operators must be well trained in preventive maintenance and troubleshooting, and a QA/QC program must be in place to quickly identify measurement and resolve issues. Audit samplers and monitors should be periodically deployed at each station to provide short term collocated measurements (this is meant to complement and not replace collocated measurements conducted to ascertain measurement precision). In addition to standard QA/QC practices for the field and laboratory operations, a complementary and particularly effective approach is to establish a program to routinely review the data as soon as practicable after data collection. Comparisons between monitors at a given station and across the network can be made with emphasis on detecting anomalies. For example, CUSUM plots can be a powerful method to detect anomalies in times series data. It is generally accepted that networks taking advantage of contemporary measurement methods require added levels of not only data validation but also data quality assessment using various types of data analyses; these assessments must be provided in an ongoing and time manner to inform the network operations.

In summary, recent developments to establish PM<sub>2.5</sub> AQOs for Hong Kong provide an unprecedented opportunity to rethink the PM mass and speciation monitoring networks. There will necessarily be multiple objectives, and numerous factors must be considered to arrive at the measurement methods and network design that achieves an appropriate balance in light of the objectives and operating costs and other constraints. A formal process is encouraged that clearly articulates the various goals and challenges so the decision-making can be transparent and effectively articulated.

## 7. References

- Bae MS, Schauer JJ, Turner JR (2006) Estimation of the monthly average ratios of organic matter to organic carbon for fine particulate matter at an urban site. *Aerosol Science and Technology*, 40: 1123-1139.
- Chen X, Yu, JZ (2007) Measurement of organic mass to organic carbon ratio in ambient aerosol samples using a gravimetric technique in combination with chemical analysis. *Atmospheric Environment*, 41: 8857-8864.
- Dorling SR, Davies TD, Pierce CE (1992) Cluster analysis: a technique for estimating the synoptic meteorological controls on air and pollution chemistry. *Atmospheric Environment*, 26A: 2575-2581.
- Draxler, RR (1999) HYSPLIT4 user's guide. NOAA Tech. Memo. ERL ARL-230, NOAA Air Resources Laboratory, Silver Spring, MD.
- Fung JCH, Lau AKH, Lam JSL, Yuan ZB (2003) Observational and modeling analysis of a severe air pollution episode in western Hong Kong, *Journal of Geophysical Research*, 110, doi: 10.1029/2004JD005105.
- Gall C, Van Rafelghem M (2006) Marine emission reduction options for Hong Kong and the Pearl River Delta. Available at: <http://www.civic-exchange.org/publications/2006/marineemission-e.pdf>
- Guo H, Ding AJ, So KL, Ayoko G, Li YS, Hung WT (2009) Receptor modeling of source apportionment of Hong Kong aerosols and the implication of urban and regional contribution. *Atmospheric Environment*, 43:1150-1169.
- Halliwell B, Gutteridge JMC (1999) Free radicals in Biology and Medicine. Oxford University Press: Oxford, United Kingdom.
- HKEPD (Hong Kong Environmental Protection Department) (2008) Air quality in Hong Kong 2007. Available at: <http://www.epd-asg.gov.hk/english/report/files/aqr07e.pdf>
- Ho KF, Lee SC, Chan CK, Yu JC, Chow JC, Yao XH (2003) Characterization of chemical species in PM<sub>2.5</sub> and PM<sub>10</sub> aerosols in Hong Kong. *Atmospheric Environment*, 37: 31-39.
- Ho KF, Lee SC, Chow JC and Watson JG (2003) Characterization of PM<sub>10</sub> and PM<sub>2.5</sub> source profiles for fugitive dust in Hong Kong. *Atmospheric Environment*, 37:1023-1032.
- Huang XF, Yu JZ, Yuan Z, Lau AKH, Louie PKK (2009) Source analysis of high particulate days in Hong Kong. *Atmospheric Environment* 43:1196-1203.
- Hyslop H and White W (2008a) An empirical approach to estimating detection limits

- using collocated data, *Environmental Science and Technology*, 42:5235-5240.
- Hyslop H and White W (2008b) An evaluation of interagency monitoring of protected visual environments (IMPROVE) collocated precision and uncertainty estimates, *Atmospheric Environment*, 42:2691-2705.
- Kaiser HF (1960) the application of electronic computers to factor analysis, *Educational and Psychological Measurement*, 20:141-151.
- Lau, AKH, Yuan Z, Bian EQ (2010) Final report of *Trend and Apportionment Analysis of Air Pollution Sources over Hong Kong* (AS06-435), prepared for the Hong Kong Environmental Protection Department, November 2010.
- Lo JCF, Lau AKH, Fung JCH, Chen F (2006) Investigation of enhanced cross-city transport and trapping of air pollutants by coastal and urban land-sea breeze circulations, *Journal of Geophysical Research*, 111, doi:10.1029/2005JD006837.
- Lo, JCF, Lau, AKH, Chen F, Fung JCH, Leung KKM (2007) Urban modification in a mesoscale model and the effects on the local circulation in the Pearl River Delta region, *Journal of Applied Meteorology and Climatology*, 46, 457-476.
- Louie PKK, Watson JG, Chow JC, Chen A, Sin DWM, Lau AKH (2005) Seasonal characteristics and regional transport of PM<sub>2.5</sub> in Hong Kong. *Atmospheric Environment* 39:1695-1710.
- Moody JL, Munger JW, Goldstein AH, Jacob DJ, Wofsy SC (1998) Harvard Forest regional-scale air mass composition by Patterns in Atmospheric Transport History (PATH). *Journal of Geophysical Research*, 103: 13181-13194.
- Polissar AV, Hopke PK, Paatero P, Malm WC, Sisler, JF (1998) Atmospheric aerosol over Alaska 2. Elemental composition and sources. *Journal of Geophysical Research* 103:19045-19057.
- Pun BK, Seigneur C (1999) Understanding particulate matter formation in the California San Joaquin Valley: conceptual model and data needs. *Atmospheric Environment* 33:4856-4875.
- Pytkowicz RM and Kester (1971) The physical chemistry of sea water, *Oceanogr. Mar. Biol. Ann. Rev.*, 9: 11-60.
- Qin Y, Oduyemi K (2003) Chemical composition of atmospheric aerosol in Dundee, UK. *Atmospheric Environment* 37:93-104.
- Rachdawong P and Christensen ER (1997) Determination of PCB sources by a principal components method with nonnegative constraints, *Environmental Science and Technology*, 31: 2691.
- Stohl A (1998) Computation, accuracy and applications of trajectories – a review and

bibliography. *Atmospheric Environment*, 32:947-966.

Thurston GD and Spengler JD (1985) A quantitative assessment of source contributions to inhalable particulate matter pollution in metropolitan Boston, *Atmospheric Environment*, 19: 9-25.

Turner JR, Garlock JG (2007) A conceptual model for ambient fine particulate matter over the St. Louis Area; Version 3.0, prepared by Washington University in St. Louis for the Missouri Department of Natural Resources, October 2007.

Turner J (2010) A Conceptual Model for PM<sub>2.5</sub> over Southeast Michigan: Intraurban Variability and High Concentration Days. Paper No. 2009-A-410-AWMA. *Proceedings of the 102<sup>nd</sup> Annual Meeting of the Air & Waste Management Association*, Detroit, MI, June 16-19, 2009.

Turpin BJ, Lim HJ (2001) Species contributions to PM<sub>2.5</sub> mass concentrations: revisiting common assumptions for estimating organic mass. *Aerosol Science and Technology* 35:602-610.

Wang T (2003) Study of visibility reduction and its cause in Hong Kong – Final report to the Hong Kong Environmental Protection Department. The Hong Kong Polytechnic University, 236pp.

Watson JG, Zhu T, Chow JC, Engelbrecht J, Fujita EM, Wilson WE (2002) Receptor modeling application framework for particle source apportionment. *Chemosphere* 49:1093-1136.

WHO (World Health Organization) (2006) WHO air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide.

Yuan ZB, Lau AKH, Zhang HY, Yu Z, Louie PKK and Fung JCH (2006) Identification and spatiotemporal variations of dominant PM<sub>10</sub> sources over Hong Kong, *Atmospheric Environment*, 40:1803-1815.

**APPENDIX A – Samples Eliminated from the Modeled Data Set  
(Cd > 0.01 µg/m<sup>3</sup>)**

Pseudo-Date	Cd, µg/m <sup>3</sup>	Site	Date
04/11/2101	0.014	CW	08/09/2006
11/24/2101	0.011	KC	11/17/1999
01/12/2102	0.013	KC	09/12/2000
09/09/2112	0.012	KC	04/14/1998
08/19/2102	0.011	KT	01/07/2002
11/03/2102	0.012	KT	02/14/2004
11/11/2104	0.012	MK	04/12/2006
04/11/2105	0.021	MK	10/04/2008
02/19/2107	0.013	ST	05/19/2000
03/10/2107	0.019	ST	09/10/2000
06/19/2107	0.016	TC	08/31/2000
06/21/2107	0.021	TC	09/12/2000
06/30/2107	0.013	TC	11/05/2000
07/21/2107	0.019	TC	03/11/2001
01/04/2108	0.012	TC	12/26/2003
08/09/2108	0.012	TC	10/06/2007
03/04/2109	0.017	TP	09/11/2000
03/13/2109	0.015	TP	11/04/2000
07/05/2109	0.023	TW	04/04/2000
01/16/2110	0.016	TW	01/02/2004
01/23/2110	0.014	TW	02/13/2004
01/28/2110	0.011	TW	03/14/2004
05/31/2110	0.016	TW	03/16/2006
03/20/2111	0.02	YL	08/31/2000
03/22/2111	0.011	YL	09/15/2000
03/31/2111	0.017	YL	11/05/2000
04/19/2111	0.08	YL	03/10/2001
06/07/2111	0.012	YL	12/29/2001
08/06/2111	0.014	YL	01/11/2003
08/08/2111	0.013	YL	01/23/2003
09/29/2111	0.011	YL	12/31/2003
10/04/2111	0.015	YL	01/30/2004
10/20/2111	0.026	YL	05/05/2004
02/07/2112	0.014	YL	03/16/2006
05/12/2112	0.011	YL	10/07/2007
04/02/2113	0.011	YL	04/14/1998

## APPENDIX B: Steps of Performing PMF, APCA and Unmix

### PMF Steps

1. Import the concentration data file containing data for all sites arranged with pseudo-dates and species mass concentrations (OC, EC, TC, NO<sub>3</sub><sup>-</sup>, SO<sub>4</sub><sup>2-</sup>, NH<sub>4</sub><sup>+</sup>, Al, Ca, Cl, Fe, K, Mg, Mn, Na, Ni, Pb, Si, S, Cl<sup>-</sup>, Ti, V, Zn) in addition to total mass variables (quartz gravimetric mass (QTM)) into the *Concentration Data File* box. Import the uncertainty data file into the *Uncertainty Data File* box. The dimension of the two datasets should be identical.
2. Assign a folder for output files in the *Output Folder* box. Comma-delimited text (.csv) is suggested for output files. After all configurations (input files, output file location, qualifier, file type, species categorization, and all run specifications from the *Model Execution* screen) are set, the configuration can be saved for loading in the future runs.
3. In the *Concentration/Uncertainty* tab of *Analyze Input Data* screen, the category of all the species can be set. Apart from setting the QTM as Total Variable in this study, all species are kept as Strong. Extra Modeling Uncertainty is set as 10%.
4. In the *Model Execution* screen, set the *Number of Runs* as 20. A fixed *seed* is preferably for further reproduction of the analysis results, set as 7 in this study. *Number of Factors* should be set with a random small number and increase once at a time until the results are not interpretable. We started with 5 and ended up with 9 in this study. The *Output File Prefix* can be set with any characters. After all these sets are made, click *Run* button to execute the model.
5. After the base runs are finished, the Q values are listed in the *Base Run Model Summary* window with the lowest Q (robust) value automatically boldfaced. It should be noted that only converged solutions should be investigated further using other tools. Base Model Results can be seen in the *Base Model Results* tab in different forms, and are output into the output folder specified.
6. In the bootstrap run, the *Selected Base Run* is set automatically and should not be modified if there is no strong justification. *Number of Runs* is set as 100 and *Minimum Correlation R-value* is 0.6 in this study. *Block Size* is calculated by the model and should not be modified if there is no strong justification. After all these sets are made, click *Run* button to execute the model.
7. After the bootstrap runs are finished, the variability in bootstrap runs is shown graphically in the *Box Plot* tab of the *Bootstrap Model Results* screen. The text results are listed in the *Summary* tab and are output into the output folder specified.

### APCA – MYSTAT Steps

1. Import the input file containing standardized concentration matrix (**Z**) data for all sites arranged with pseudo-dates and species standardized mass concentrations (OC, EC, TC,  $\text{NO}_3^-$ ,  $\text{SO}_4^{2-}$ ,  $\text{NH}_4^+$ , Al, Ca, Cl, Fe, K, Mg, Mn, Na, Ni, Pb, Si, S,  $\text{Cl}^-$ , Ti, V, Zn).
2. Go to *Analyze>Factor Analysis*.
3. Select the species for factor analysis; (OC, EC,  $\text{NO}_3^-$ ,  $\text{SO}_4^{2-}$ ,  $\text{NH}_4^+$ , Al, Ca, Cl, Fe, K, Mg, Mn, Na, Ni, Pb, Si, Ti, V, Zn). TC is excluded in favor of its two components; EC and OC. Sulfur is excluded to avoid redundancy with sulfate.  $\text{Cl}^-$  was excluded in favor of Cl to remain consistent with  $\text{PM}_{10}$  modeling.
4. Specify the number of factors (between 5 and 9 for this analysis), select the varimax rotation and save the *Factor Coefficient* and *Factor Scores*. The analysis needs to be performed twice to save these results separately. The *Factor Coefficient* output contains ‘Coefficient of Standardized Factor Scores’ and ‘Rotated Loading Matrix’, while the *Factor Scores* contains the rotated principal component scores (**P\***).
5. Multiply the species-specific ‘Coefficient of Standardized Factor Score’ by the ratio of mean to standard deviation of the corresponding standardized species concentration to obtain the Z-scores at zero concentration **Z<sub>0</sub>**. Within each factor, add the **Z<sub>0</sub>** values for all species to obtain the rotated absolute scores (**P<sub>0</sub>\***).
6. A matrix of Absolute Principal Component Scores (**APCS**) is calculated by subtraction **P\* - P<sub>0</sub>\***.
7. The vector of total particulate mass (quartz filter gravimetric mass in this analysis) is regressed on **APCS** and the regression coefficients together with **APCS** are used to reconstruct the SCEs. The multivariable regression is also performed using MYSTAT. The constant term in the regression, when not forced to be zero, provides a measure of bias in the modeled concentration and can reinforce the selection of the optimal number of factors.
8. The vectors of the reconstructed SCEs are regressed on each species mass concentration to obtain the factor profiles. These multivariable regressions are also performed in MYSTAT.
9. The source profiles are examined to check for the physical interpretability of the solution and to match them to possible source categories.
10. Na and Mg appeared as dominant species in separate factors and were excluded before repeating the analysis.
11. Based on the suite of results (SCEs, G-space plots, physical interpretability of the factor profiles and explained mass profiles), and in particular the frequency of large negative species loadings and SCEs, APCA was deemed unable to resolve interpretable factors for  $\text{PM}_{2.5}$  dataset.

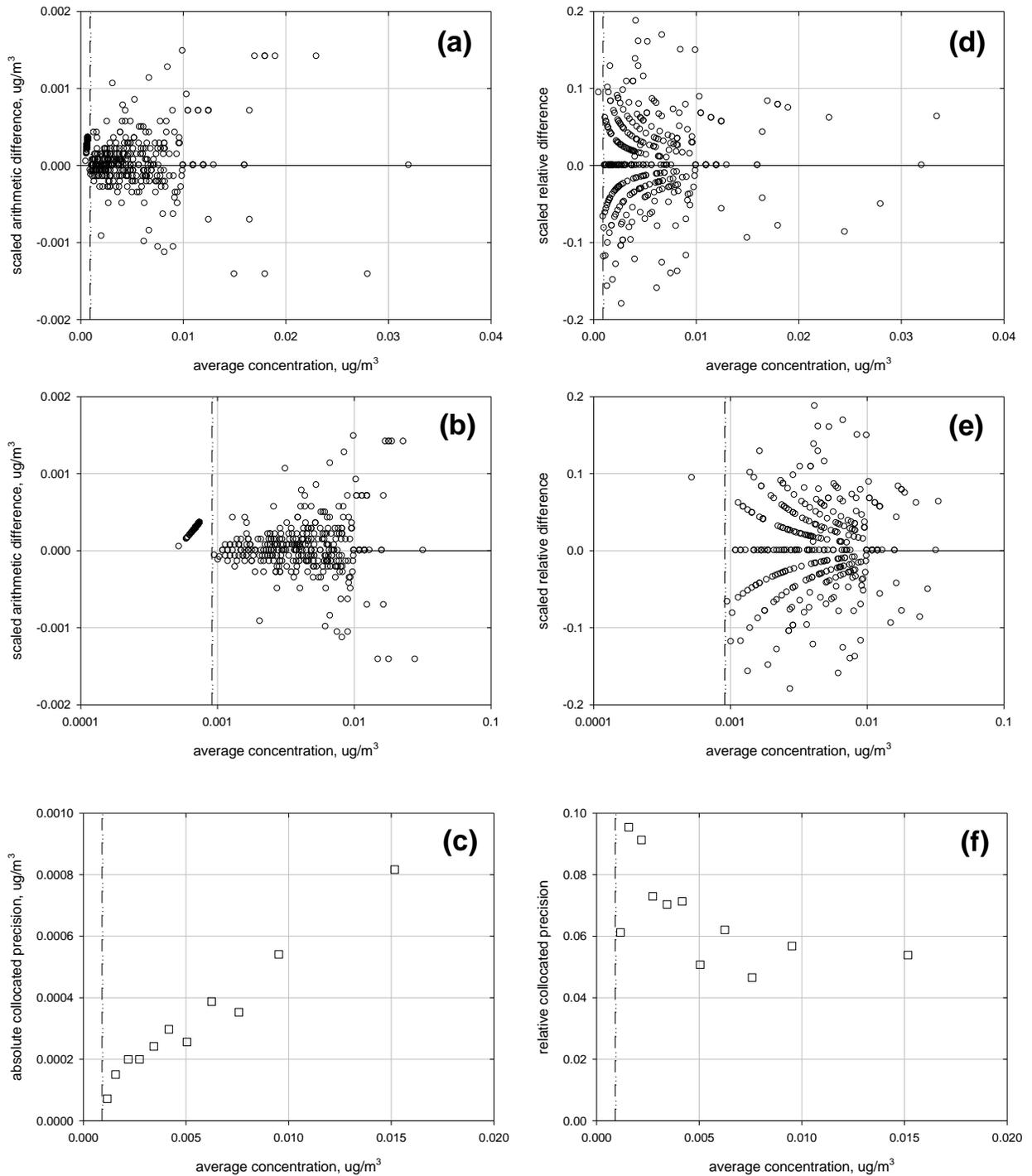
### Unmix Steps

1. Import the input file containing data for all sites arranged with pseudo-dates and species mass concentrations (OC, EC, TC,  $\text{NO}_3^-$ ,  $\text{SO}_4^{2-}$ ,  $\text{NH}_4^+$ , Al, Ca, Cl, Fe, K, Mg, Mn, Na, Ni, Pb, Si, S,  $\text{Cl}^-$ , Ti, V, Zn) in addition to total mass variables (quartz gravimetric mass (QTM), Teflon gravimetric mass (TTM) and sum-of-species mass (STM)) into Unmix.
2. In the *Data Processing* window, the species to be included / excluded can be selected. TC was excluded in favor of its two components; EC and OC. Sulfur was excluded to avoid redundancy with sulfate. Trial runs did not reveal any significant difference when either Cl or  $\text{Cl}^-$  was used in the model, thus  $\text{Cl}^-$  was excluded to remain consistent with  $\text{PM}_{10}$  modeling. STM and TTM were excluded since QTM will be used as the total mass variable.
3. In the main window, select Quartz filter mass (QTM) and mark it as the TOTAL (mass) variable.
4. Using *Select Initial Species* function a five factor solution was obtained with OC, EC,  $\text{NO}_3^-$ ,  $\text{SO}_4^{2-}$ ,  $\text{NH}_4^+$ , Ca, K, Pb and Si.
5. The *Suggest Additional Species* function with SAFER algorithm suggested Na, Ni, V, Zn and Mg, while the remaining species represented trace metals and sea salt (Al, Fe, Mn, Ti and Cl). Hence, all the species were manually selected to obtain 5-to-8 factor solutions.
6. In *Diagnostic Plots* function, source profiles/contribution and G-space plots were obtained. The factor profiles obtained had large negative concentrations apportioned to certain species ( $\text{SO}_4^{2-}$ , OC and other species). Two of the factors were found to be separately dominated by Na and Mg, respectively, and were excluded.
7. Subsequently Zn now dominated a single factor and was also removed.
8. Based on the negative loadings assigned to certain resolved factors and the G-space plots, the seven factor solution was deemed the most reasonable solution.
9. Secondary sulfate, vehicle exhaust, biomass burning, secondary nitrate, residual oil combustion, soil/dust and sea salt were identified by their tracer species as the main factors contributing to the PM mass.

## APPENDIX C: PM<sub>10</sub> TEOM Mass Partitioning into Base and Excess Contributions

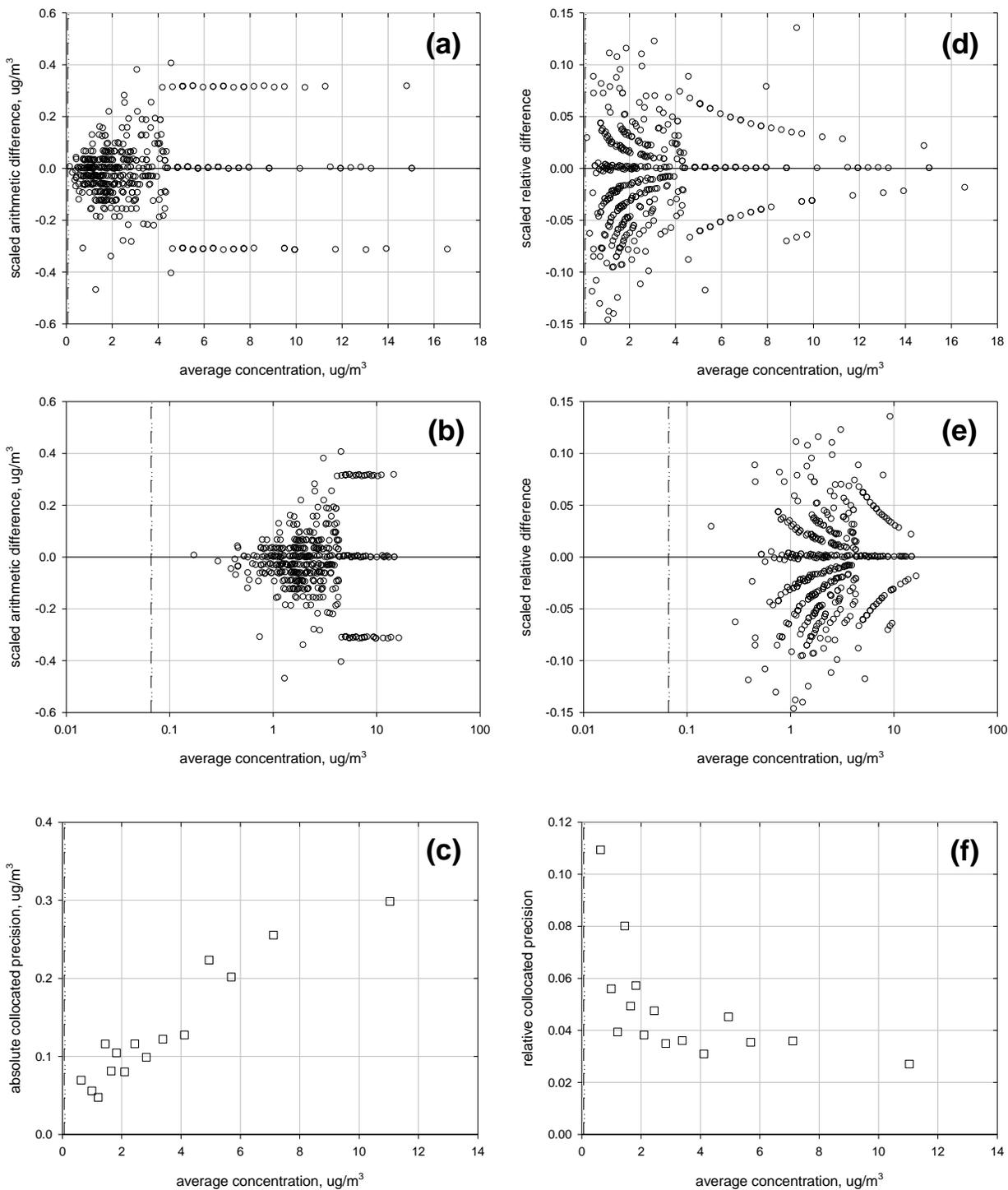
As described in the main body of the report, a preliminary analysis was performed by partitioning the daily-average TEOM mass measured at each station over the period 1999-2008 into a network-wide “base” concentration and a site-specific “excess” concentration. Valid network days were defined as those days with a valid 24-hour average concentration for at least twelve of the TEOMs. In this study, the 5<sup>th</sup>-lowest concentration for a given day was chosen as the base concentration. **Figure C-1** shows the relationships between the daily 5<sup>th</sup>-minimum value and the daily 4<sup>th</sup>- and 6<sup>th</sup> minimum values. In both cases the concentration differences are clustered tightly just above the 1:1 line. This demonstrates that on a daily basis this portion of portion of the concentration distribution across all sites is well behaved. **Figure C-2** shows how the grand average base concentration changes for a change in the metric from one rank to the next rank (e.g. From the 4<sup>th</sup>-lowest to the 5<sup>th</sup>-lowest). As metric is changed from the 4<sup>th</sup>-lowest to the 8<sup>th</sup>-lowest value, the base concentration increases by about 2 µg/m<sup>3</sup> with each step. In general, any of these metrics could have been chosen to represent the base concentration with no impact on the qualitative patterns but up to a 10 µg/m<sup>3</sup> difference in the actual value assigned to the base. This behavior reflects on the general uniformity in PM<sub>10</sub> concentrations across the general stations in the network. **Figure C-3** shows the 24-hour average concentrations at a general station (CW) and a roadside station (CL) versus the daily base concentration. At the general station the data are tightly clustered around the 1:1 line; this site closely tracks the base concentration. In contrast, at the roadside station the values are consistently above the base concentration with higher day-to-day variability due to daily changes in local source impacts.

# Arsenic



**Figure 2.1** SSP collocated data for arsenic: (a) scaled arithmetic difference, linear concentration scale; (b) scaled arithmetic difference, logarithmic concentration scale; (c) binned absolute collocated precision; (d) scaled relative difference, linear concentration scale; (e) scaled relative difference, logarithmic concentration scale; and (f) binned relative collocated precision. The dashed vertical line is the MDL.

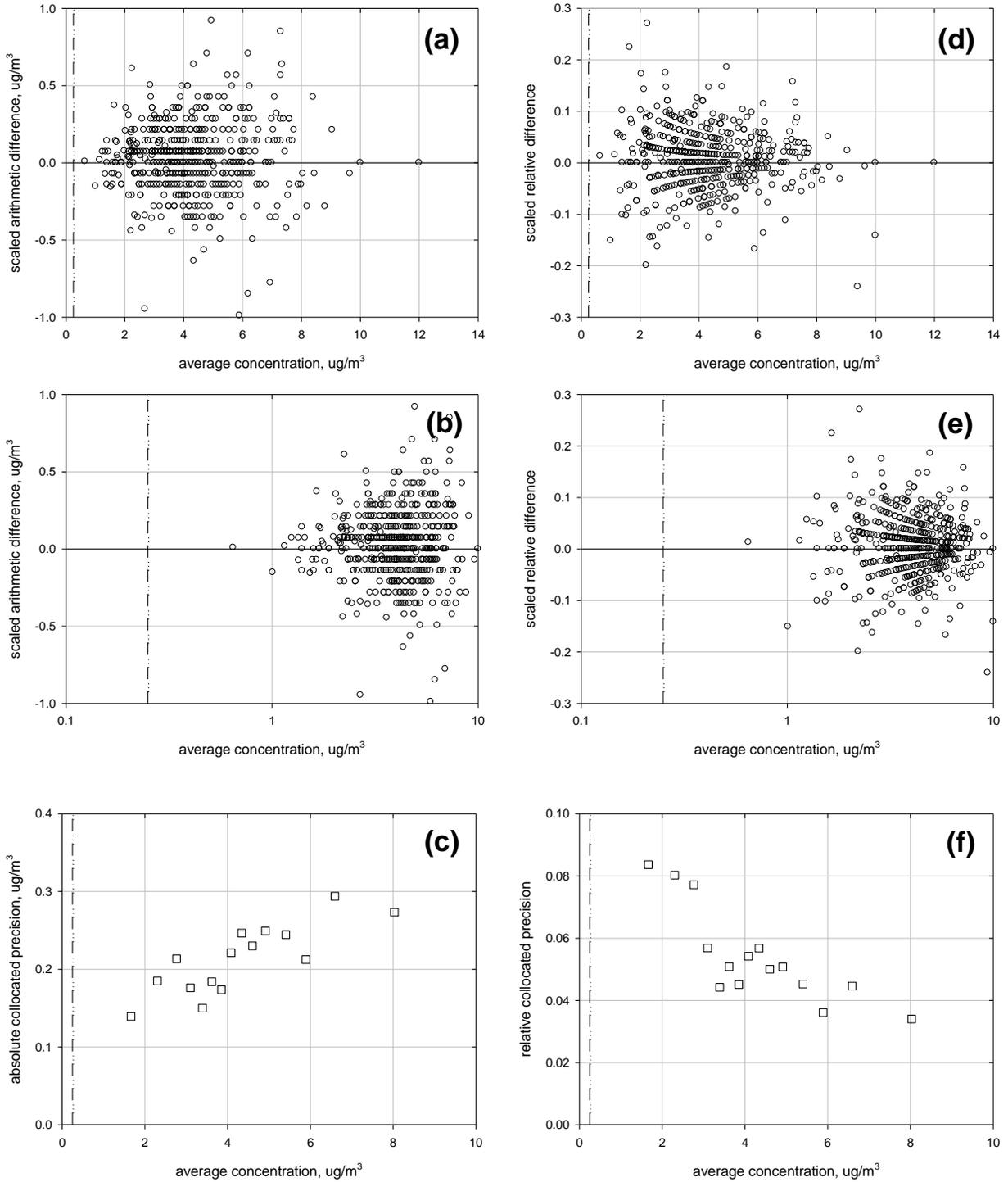
# Nitrate



**Figure 2.2** SSP collocated data for nitrate: (a) scaled arithmetic difference, linear concentration scale; (b) scaled arithmetic difference, logarithmic concentration scale; (c) binned absolute collocated precision; (d) scaled relative difference, linear concentration

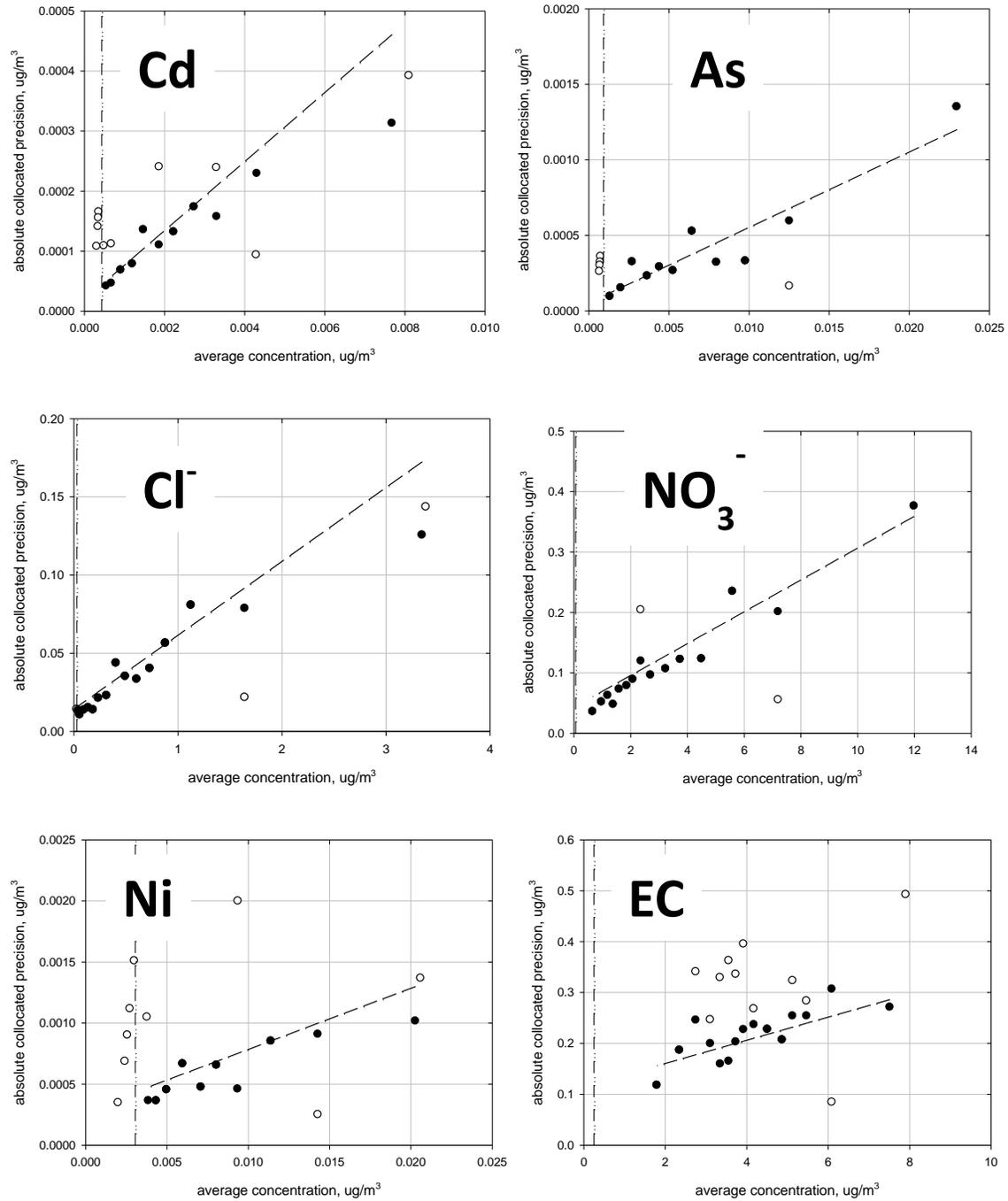
scale; (e) scaled relative difference, logarithmic concentration scale; and (f) binned relative collocated precision. The dashed vertical line is the MDL.

## Elemental Carbon



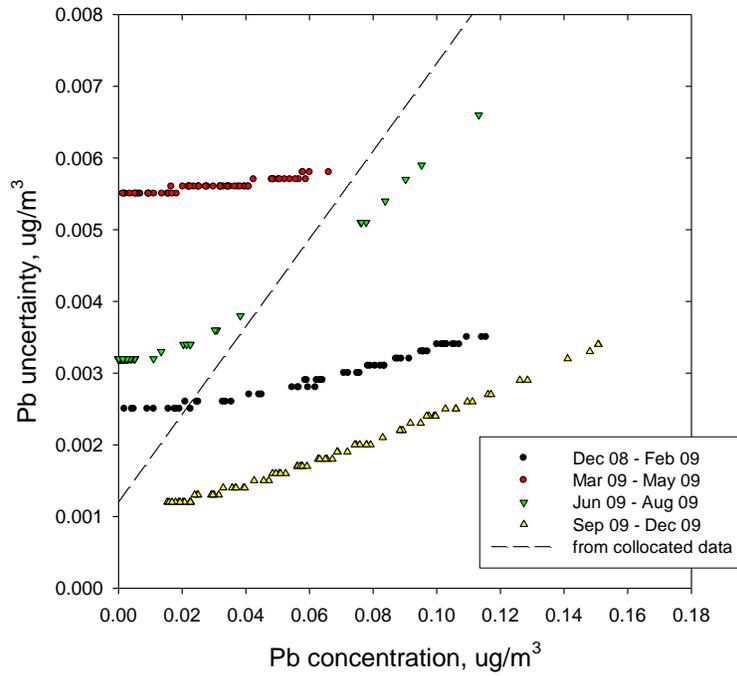
**Figure 2.3** SSP collocated data for elemental carbon: (a) scaled arithmetic difference, linear concentration scale; (b) scaled arithmetic difference, logarithmic concentration

scale; (c) binned absolute collocated precision; (d) scaled relative difference, linear concentration scale; (e) scaled relative difference, logarithmic concentration scale; and (f) binned relative collocated precision. The dashed vertical line is the MDL.

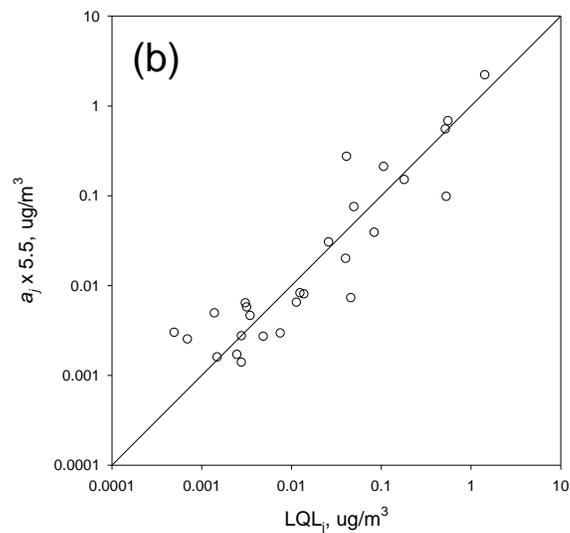
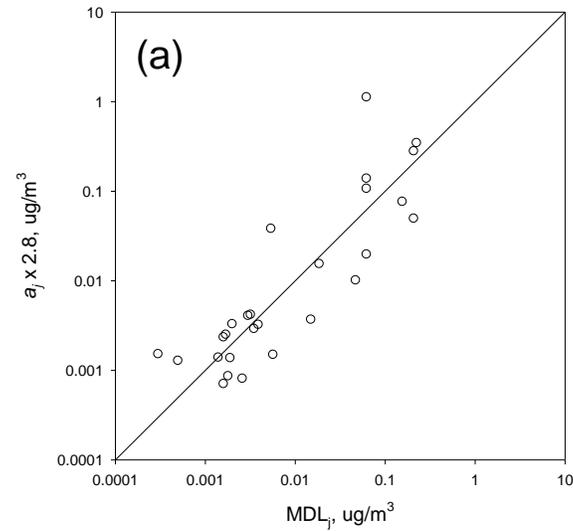


**Figure 2.4** Six examples of the binned absolute collocated precision for the TW data set using raw data (open circles) and outliers removed (shaded circles); and the error

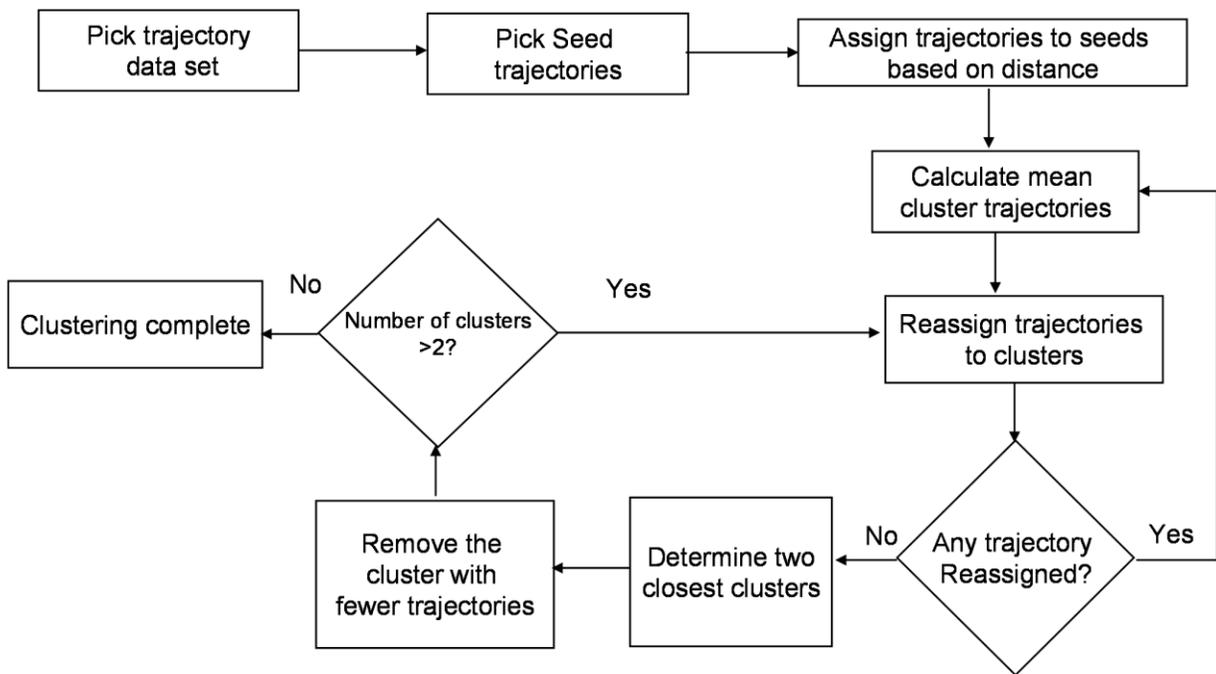
structures estimated regressions of the SSP collocated data (dashed lines). The dot-do-dash vertical lines are the MDL values.



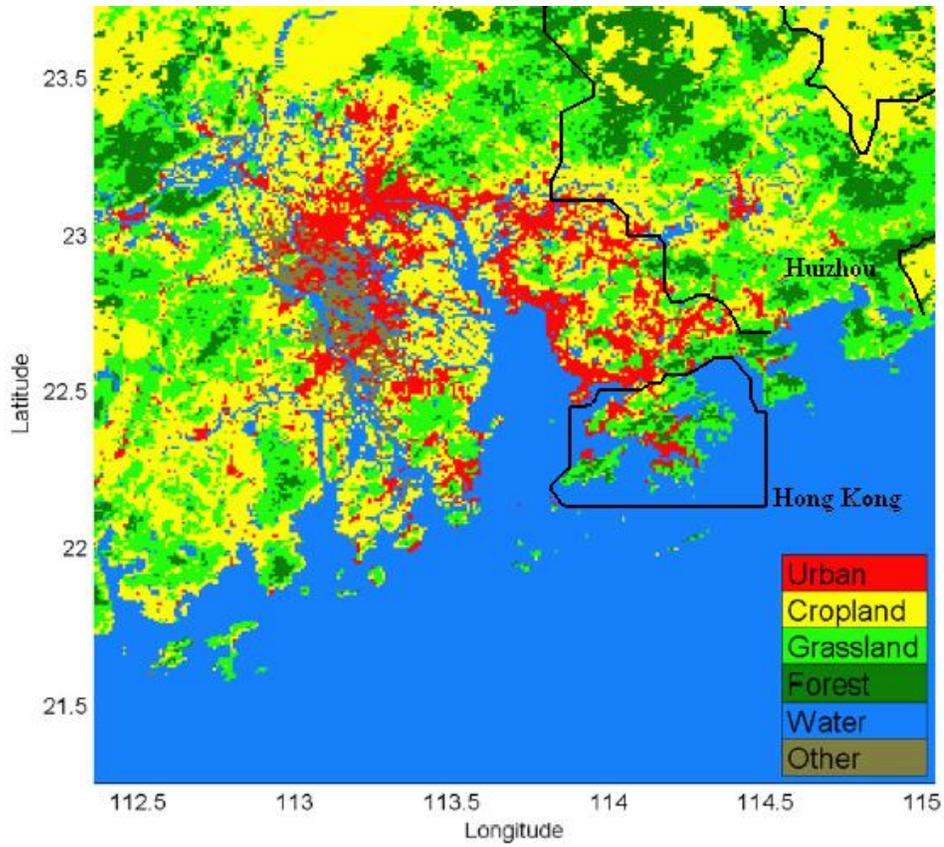
**Figure 2.5** Error structures for Pb reported by DRI for the four analysis batches of 2009 PM<sub>2.5</sub> samples (markers) and the error structure derived from the analysis of collocated precision data for 2009 (dashed line).



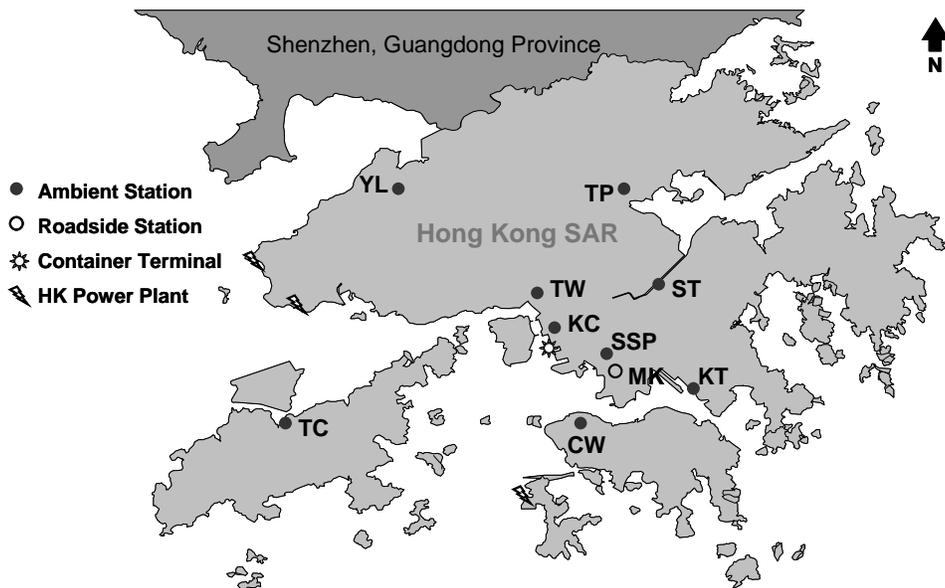
**Figure 2.6** Additive term ( $a_j$ ) calculated from the weighted regression of binned collocated precision on concentration versus DRI-reported MDL (figure a) and LOQ (figure b) values. For each comparison a multiplicative term was calculated by minimizing the sum-of-squares difference between  $a_j$  and either  $MDL_j$  or  $LOQ_j$  after log-transforming the data.



**Figure 2.7** Flow chart for the cluster analysis methodology used in this study.

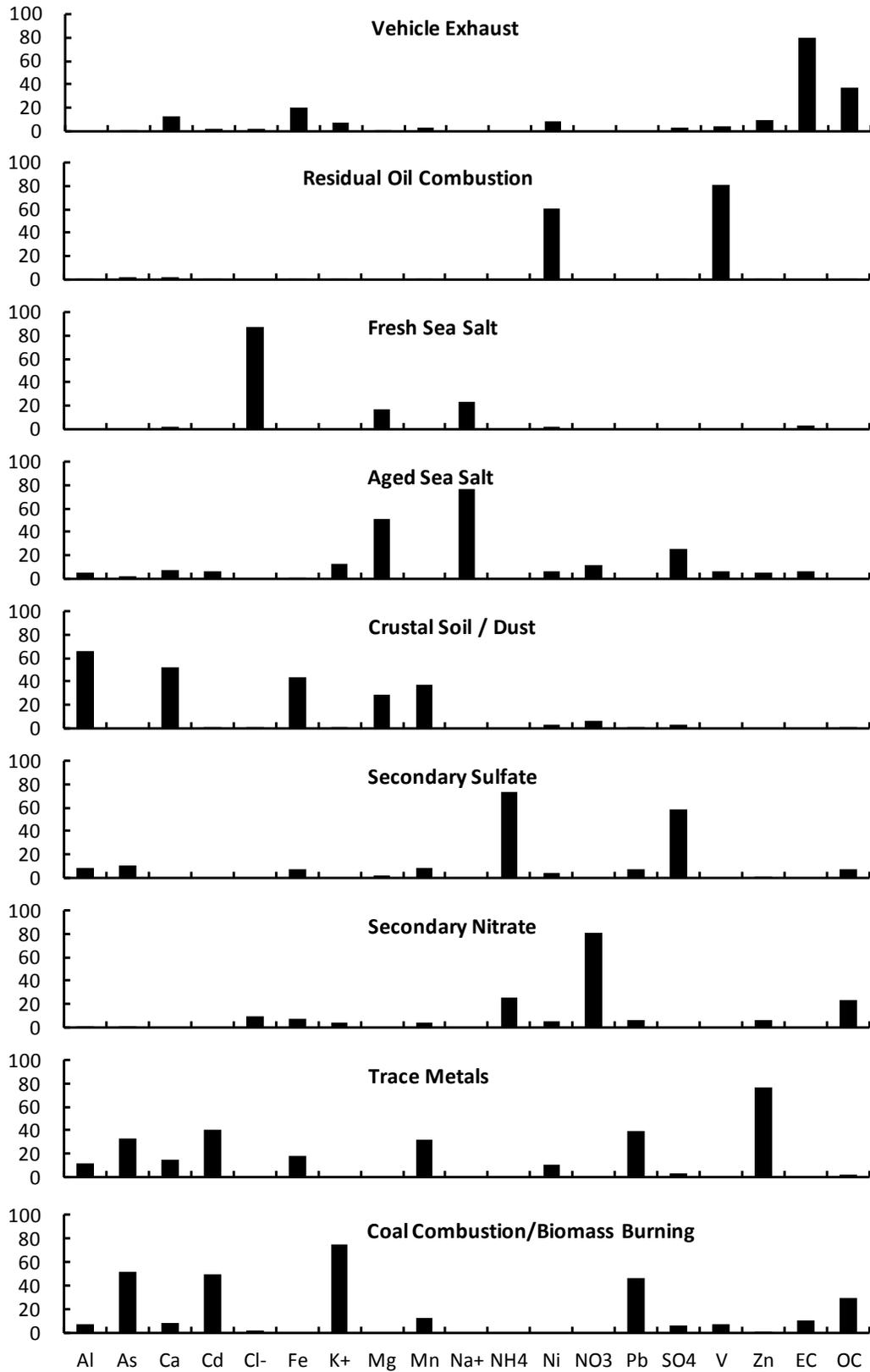


(a)

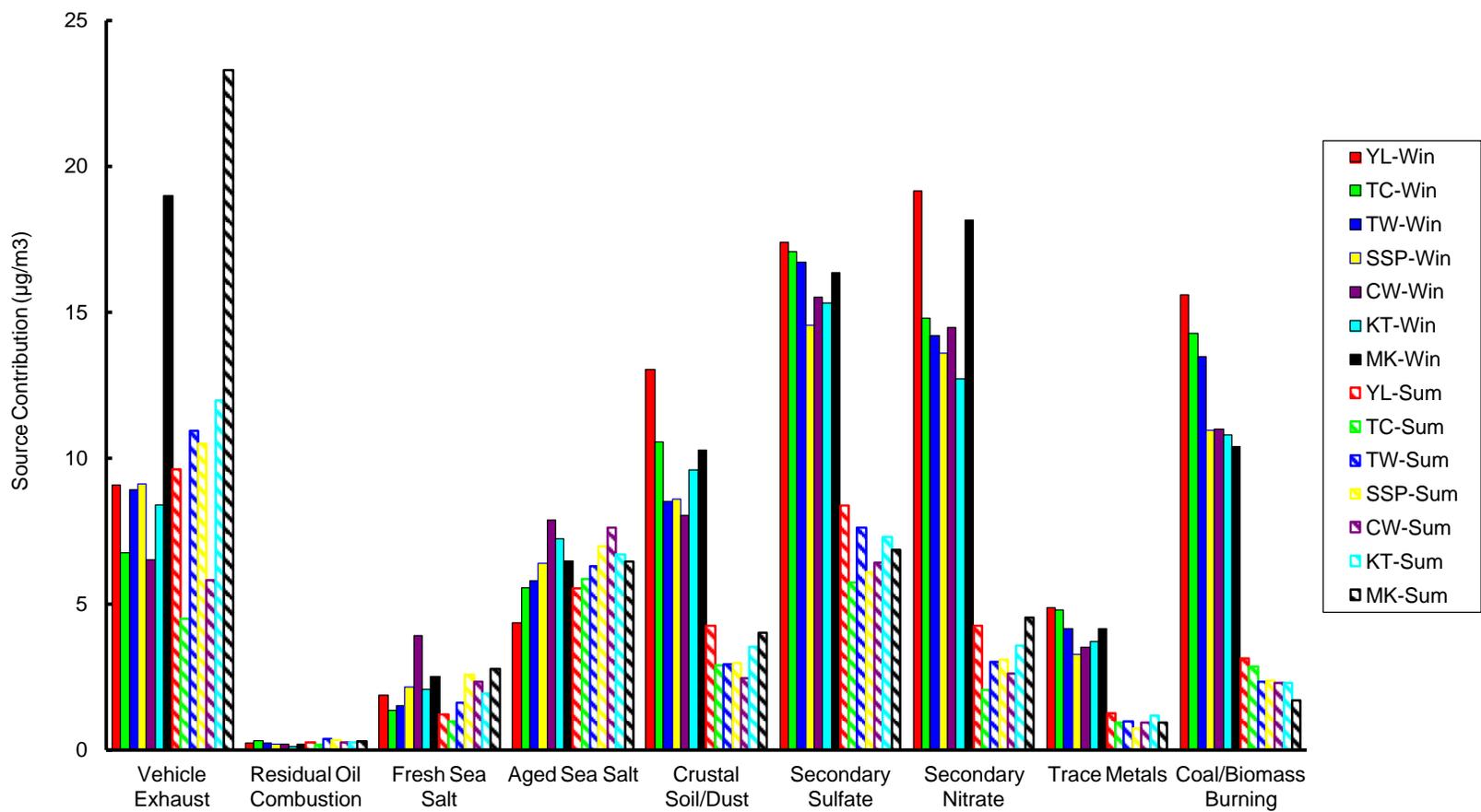


(b)

**Figure 3.1** (a) Land-use Map of the Pearl River Delta, and (b) Geographical distribution of the AQMSs in Hong Kong.



**Figure 3.2** Percentage Explained Variation for the source profiles derived by PMF



(The contribution in winter is significantly higher than in summer for the last five sources, while for the first four sources, no significant seasonal variation could be observed.)

**Figure 3.3** Comparison of source contributions in summer and winter, derived by PMF (unit:  $\mu\text{g}/\text{m}^3$ ).

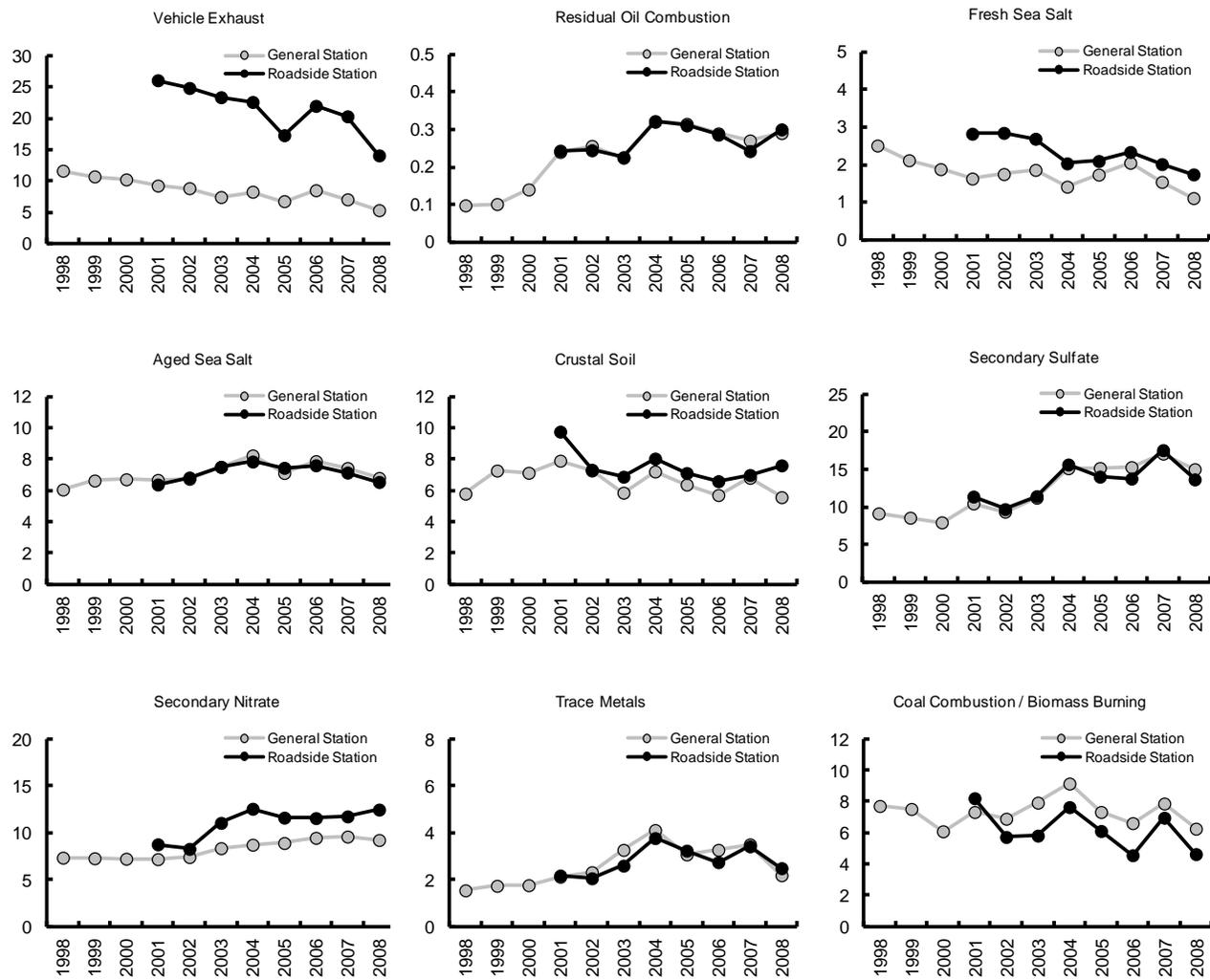
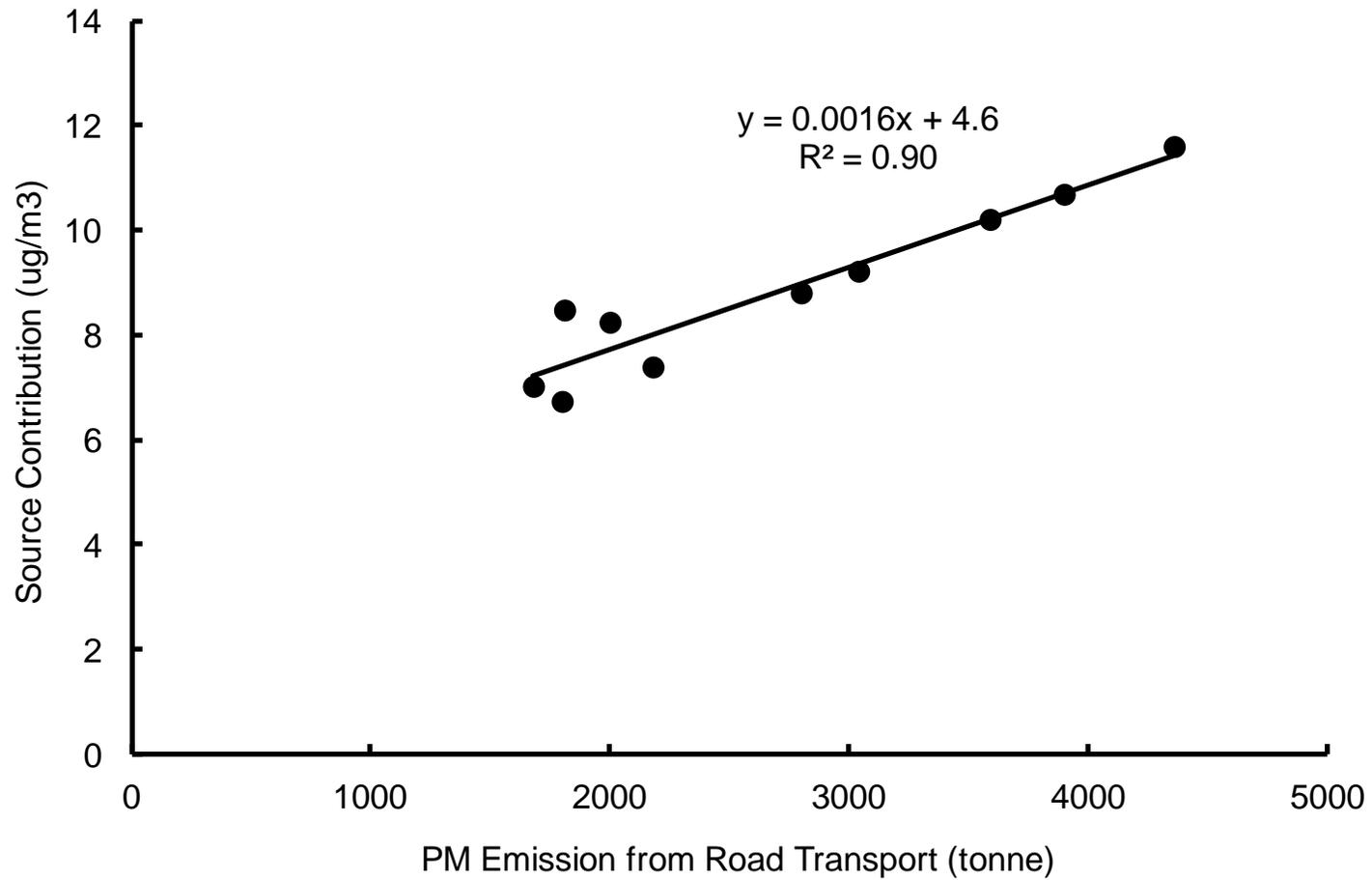
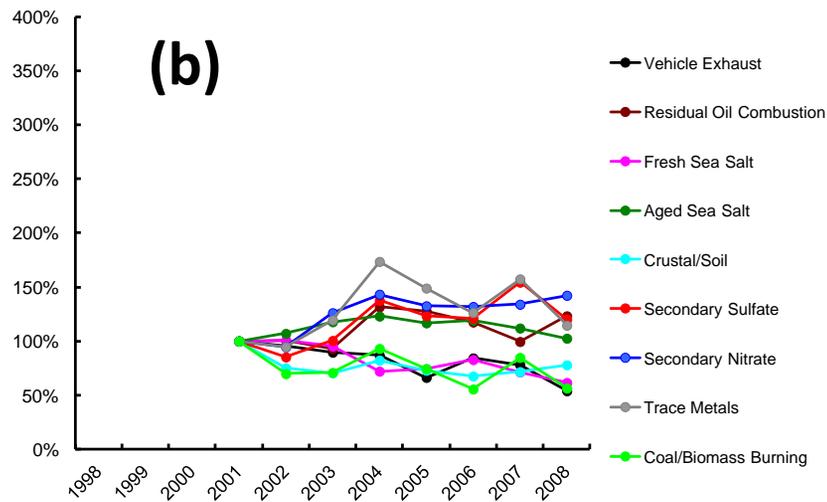
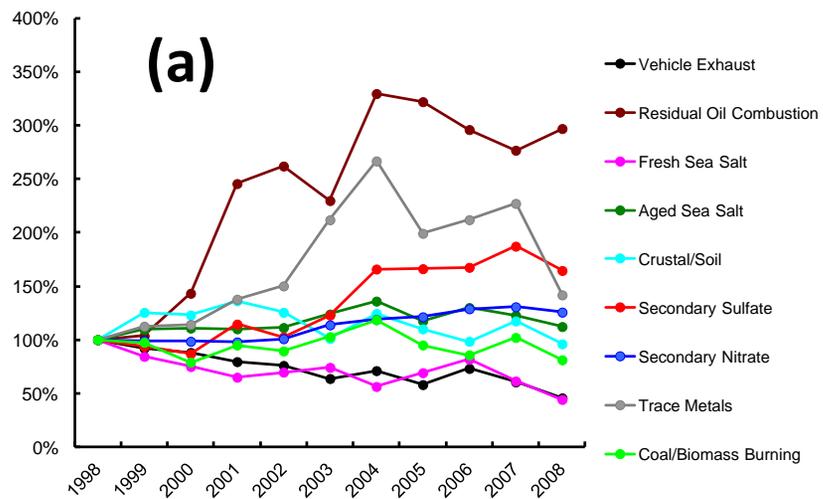


Figure 3.4 Annual trends of identified PM<sub>10</sub> sources impacting Hong Kong at general stations and roadside station Mong Kok.

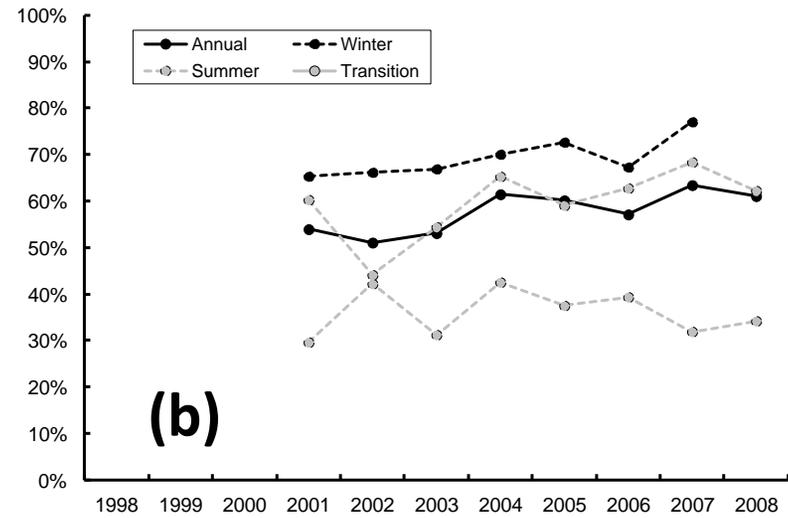
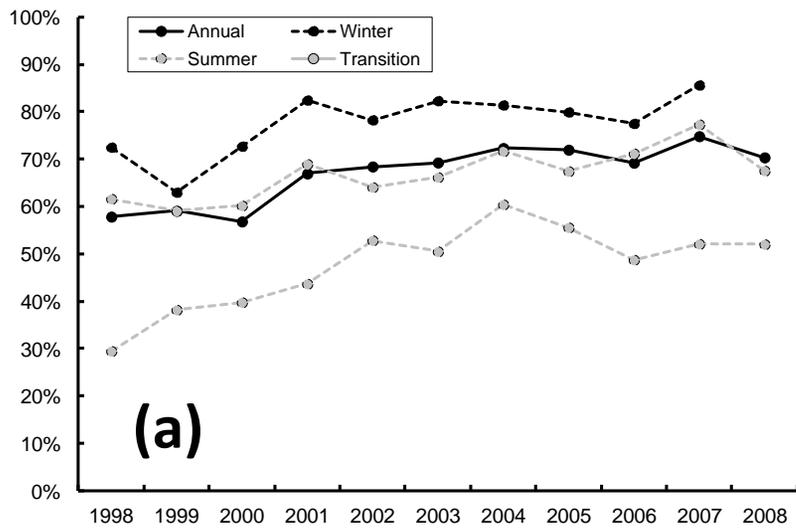


Source: Hong Kong air pollutant emission inventory, available at: [http://www.epd.gov.hk/epd/english/environmentinhk/air/data/files/table\\_pm.pdf](http://www.epd.gov.hk/epd/english/environmentinhk/air/data/files/table_pm.pdf)

**Figure 3.5** Strong correlation between apportioned vehicle contribution and estimation of PM emission from road transport.

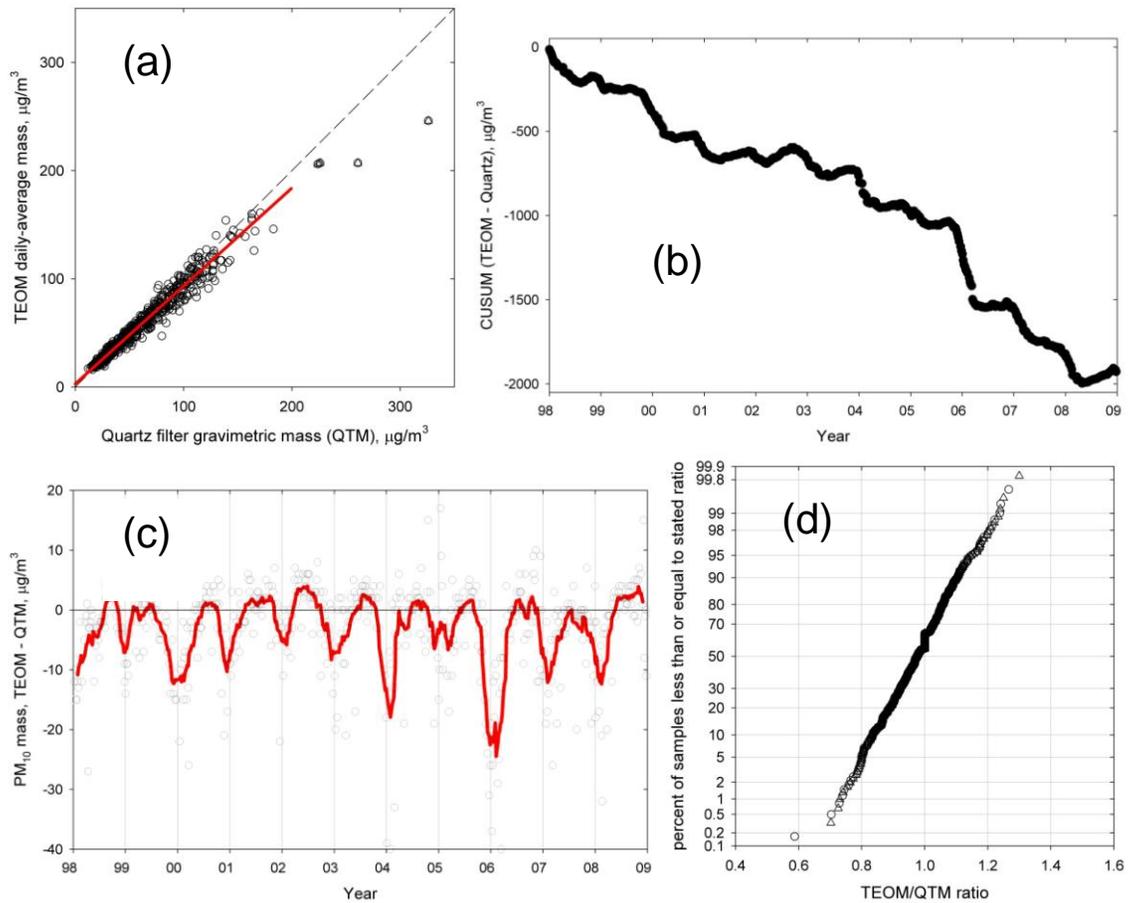


**Figure 3.6** Relative annual contribution variation of identified sources with respect to their contribution (a) in 1998 at general stations, and (b) in 2001 at roadside station Mong Kok.

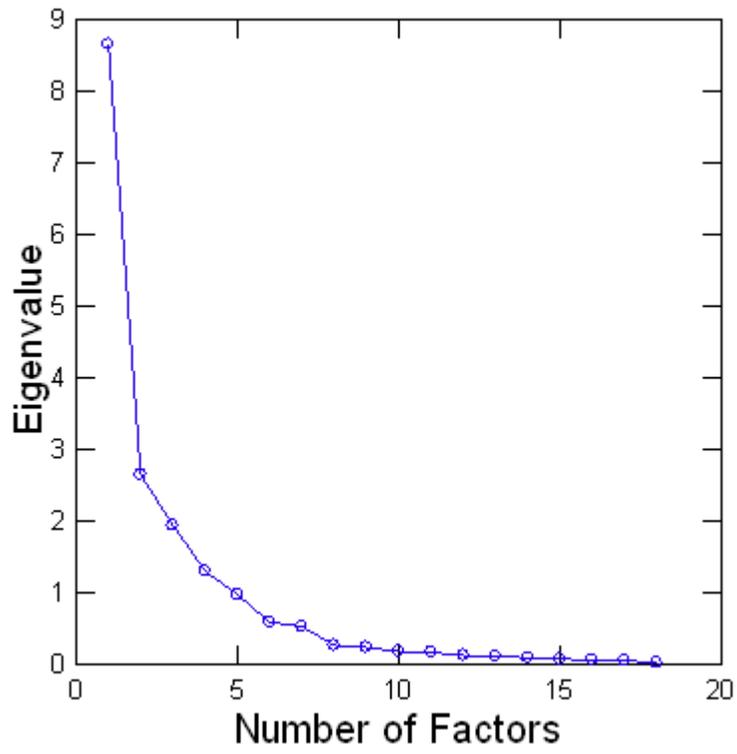


**Figure 3.7** Long-term trend of relative contribution from non-local sources in winter, summer, transition season and all sampling days (a) at general stations, and (b) at roadside station Mong Kok.

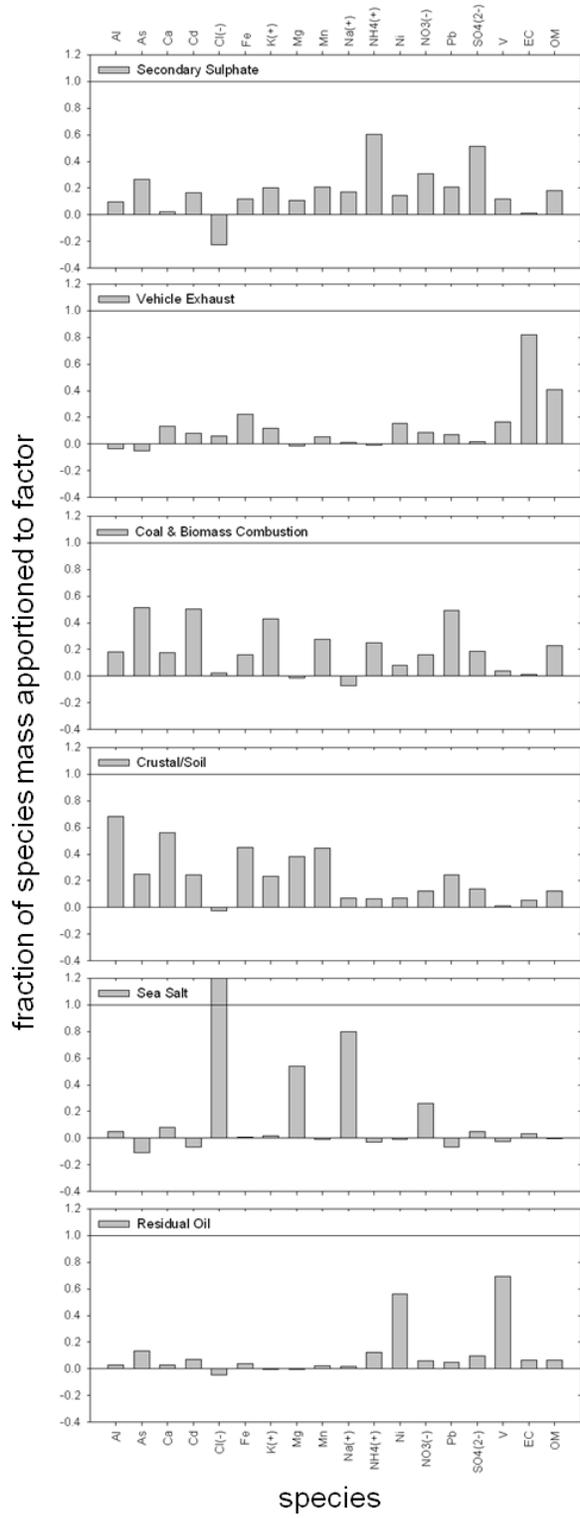




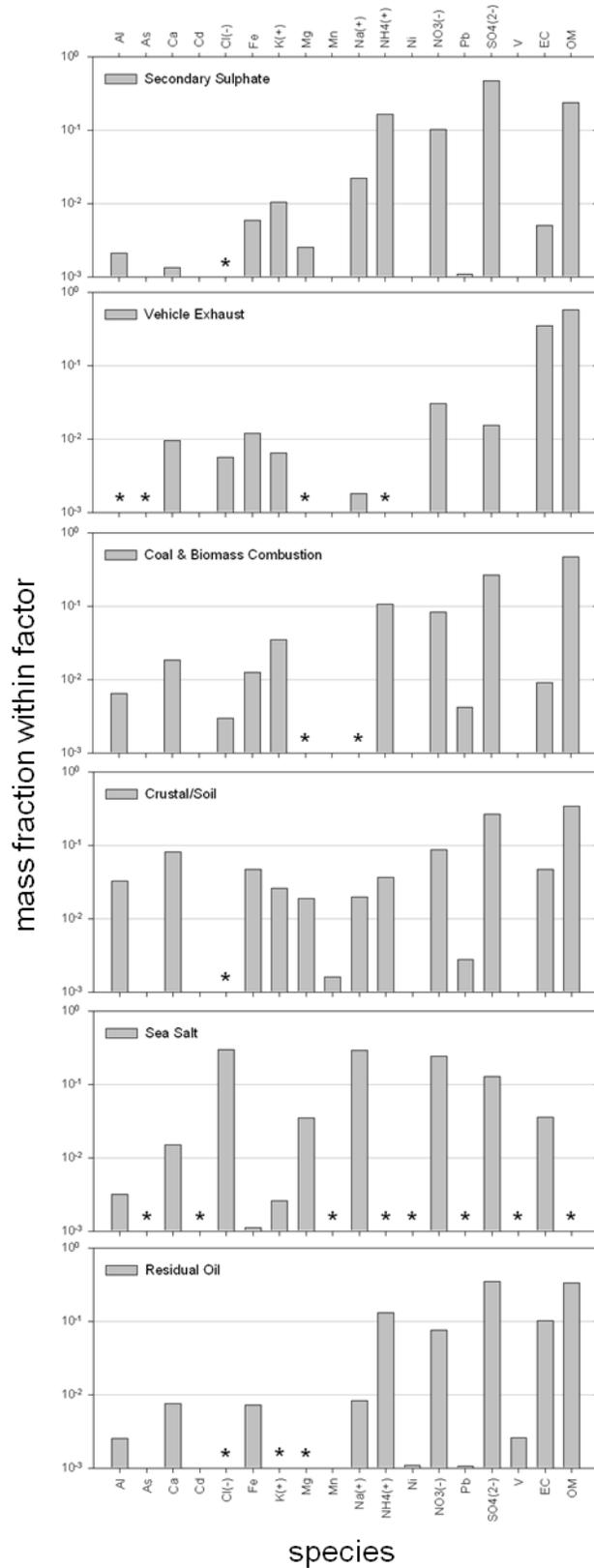
**Figure 3.8** Relationships between PM<sub>10</sub> 24-average TEOM mass and 24-hour integrated QTM mass at Yuen Long (YL): (a) scatter plot of TEOM on QTM; (b) cumulative summation (CUSUM) time series plot for TEOM minus QTM; (c) time series of TEOM minus QTM (markers) including a centered 11-sample arithmetic mean smoother (line); and (d) cumulative distribution of the TEOM/QTM mass ratio.



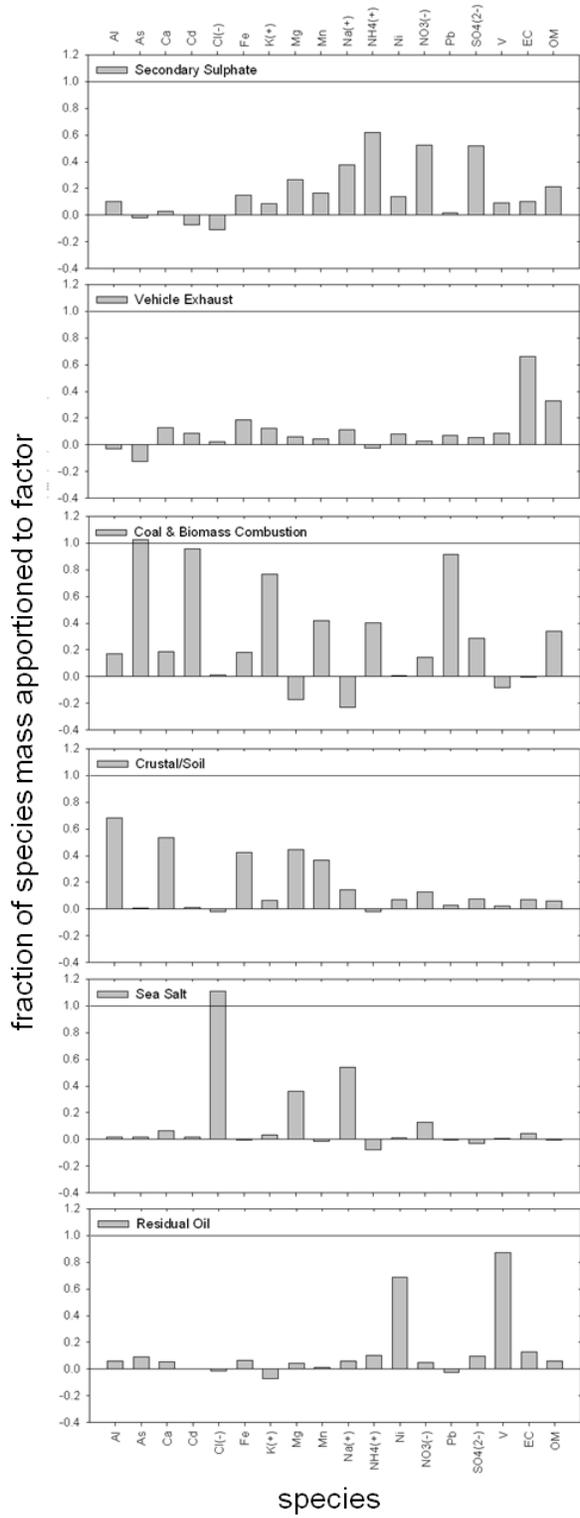
**Figure 3.9** Scree plot for the standardized, conditioned data.



**Figure 3.10** Explained mass profiles (fraction of species mass) for the six factor APCA solution.



**Figure 3.11** Source profiles (species mass fractions) for the six factor APCA solution.



**Figure 3.12** Explained mass profiles (fraction of species mass) for the six factor UNMIX solution.

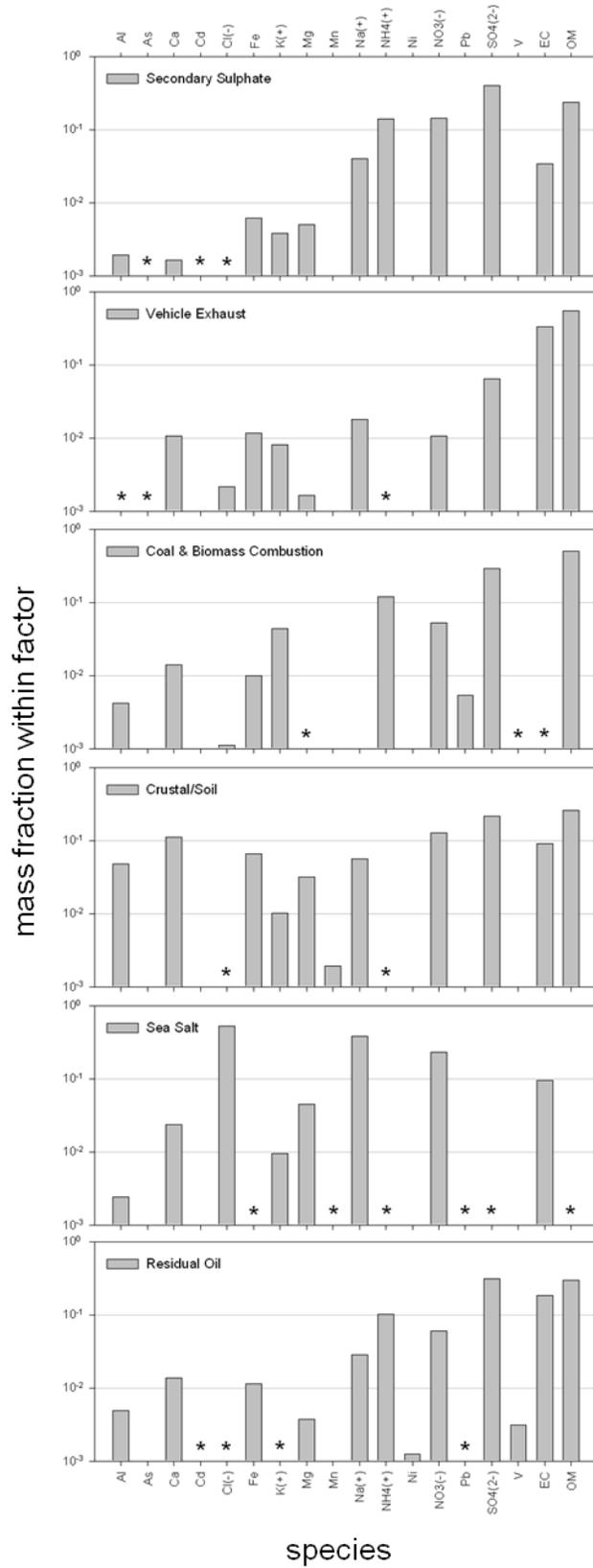
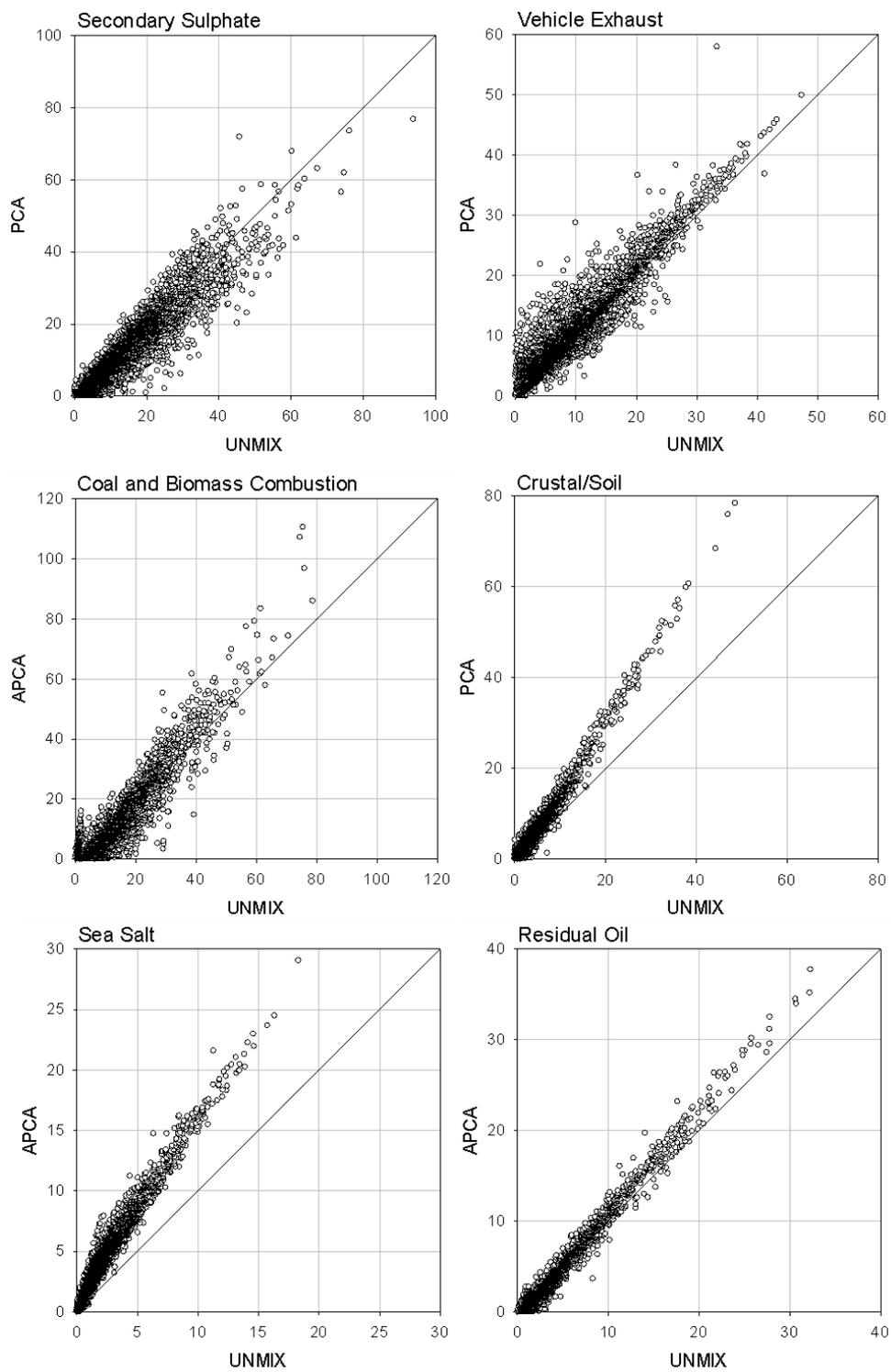
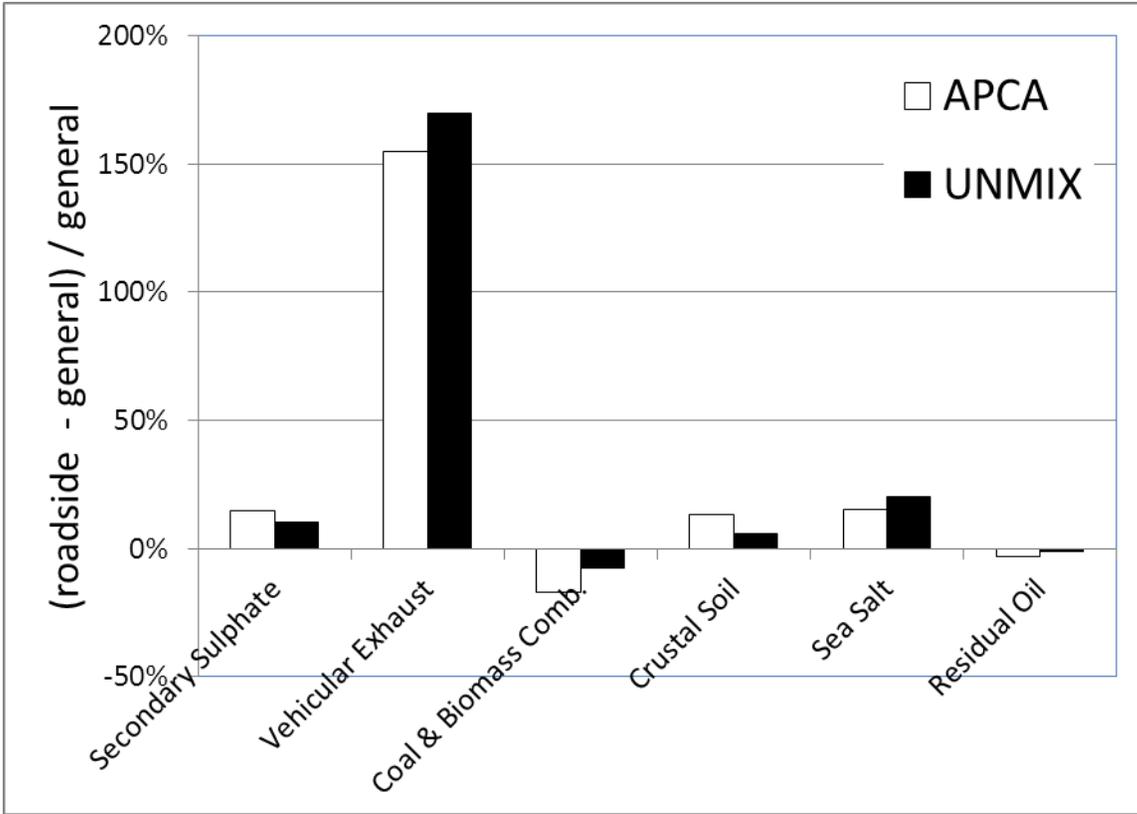


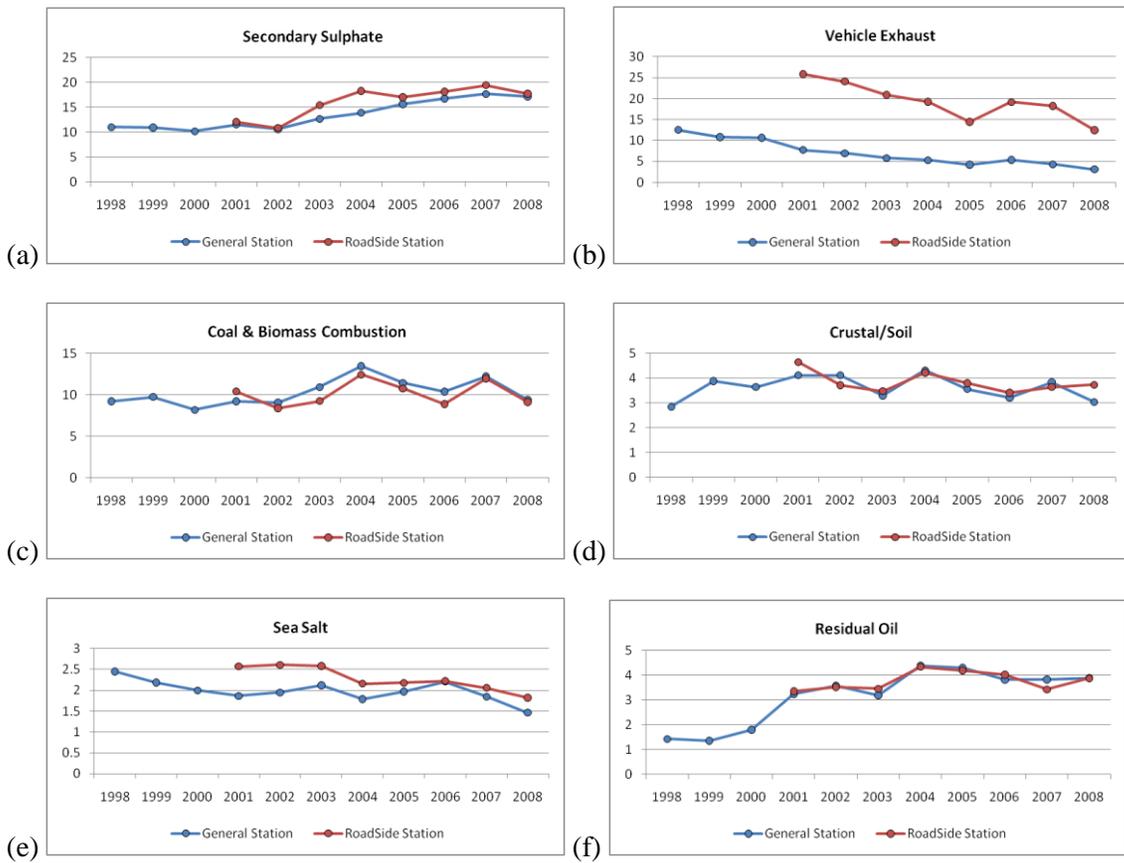
Figure 3.13 Source profiles (species mass fractions) for the six factor UNMIX solution.



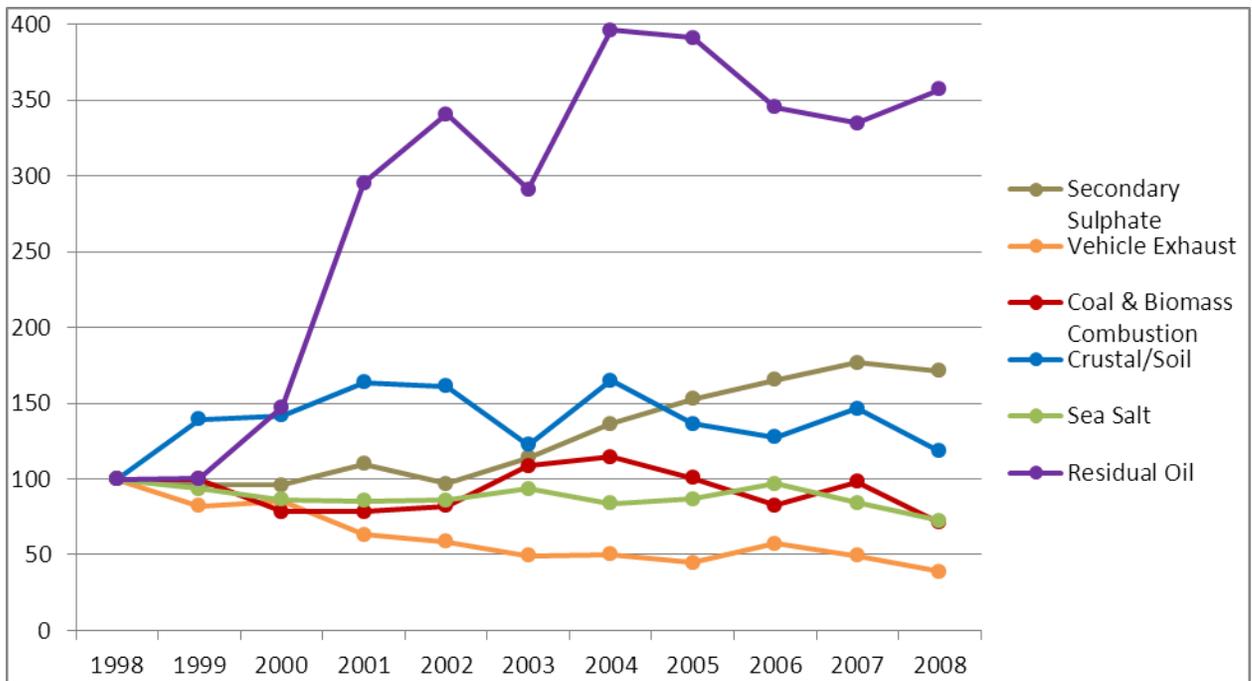
**Figure 3.14** Scatter plots for the sample-specific source contribution estimates derived from APCA and UNMIX analyses. All concentration units are  $\mu\text{g}/\text{m}^3$ .



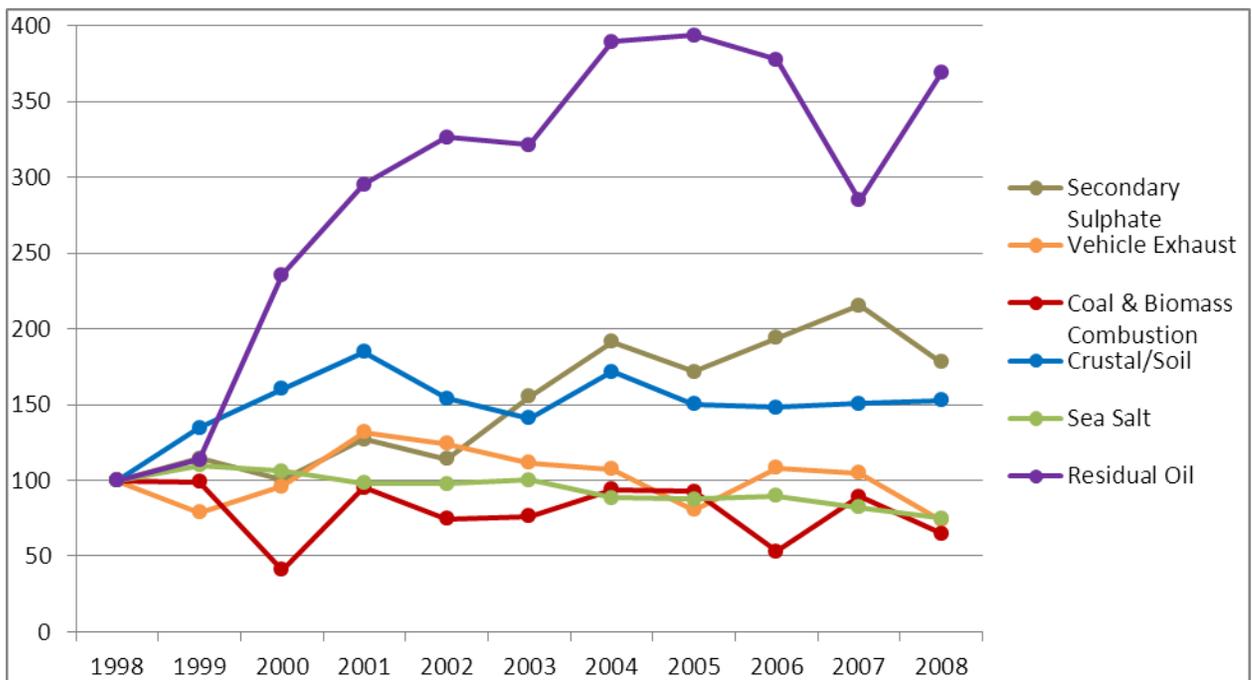
**Figure 3.15** Relative differences in study-average source contribution differences between the roadside station and general stations for APCA and UNMIX.



**Figure 3.16** UNMIX-modeled annual- average source contribution estimates at the roadside stations (red) and general stations (blue) for each emission source category. Source contribution estimates are reported in  $\mu\text{g}/\text{m}^3$ .

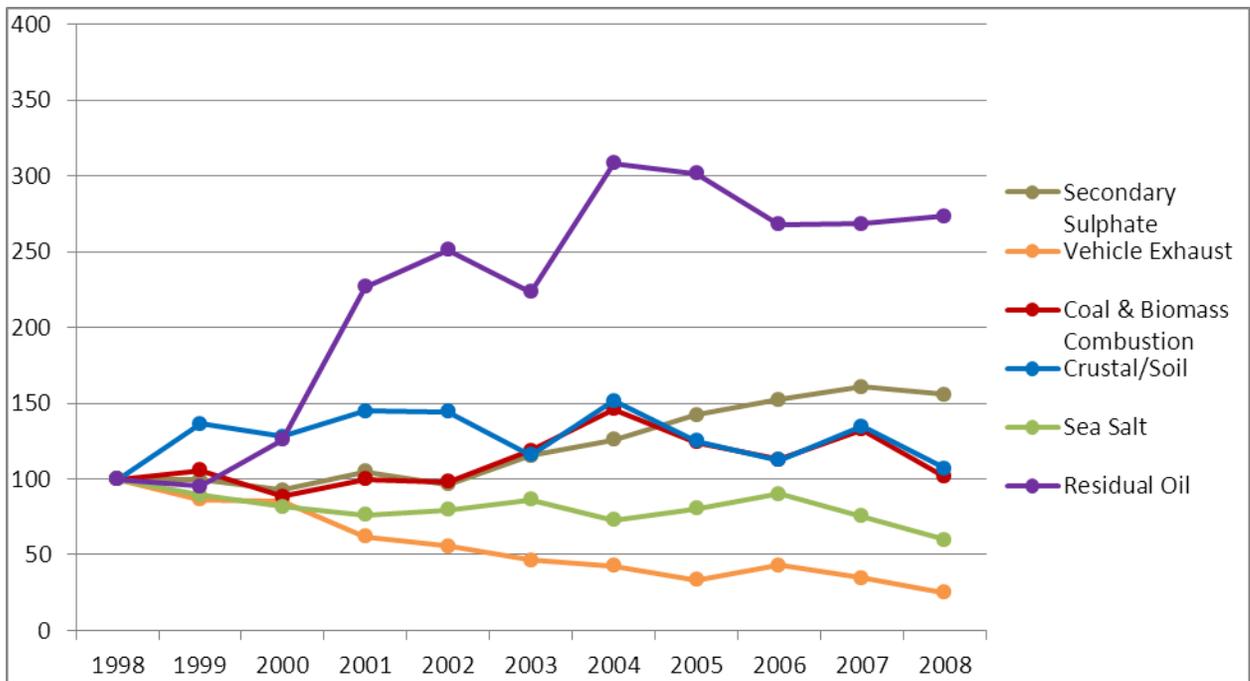


(a) General Station

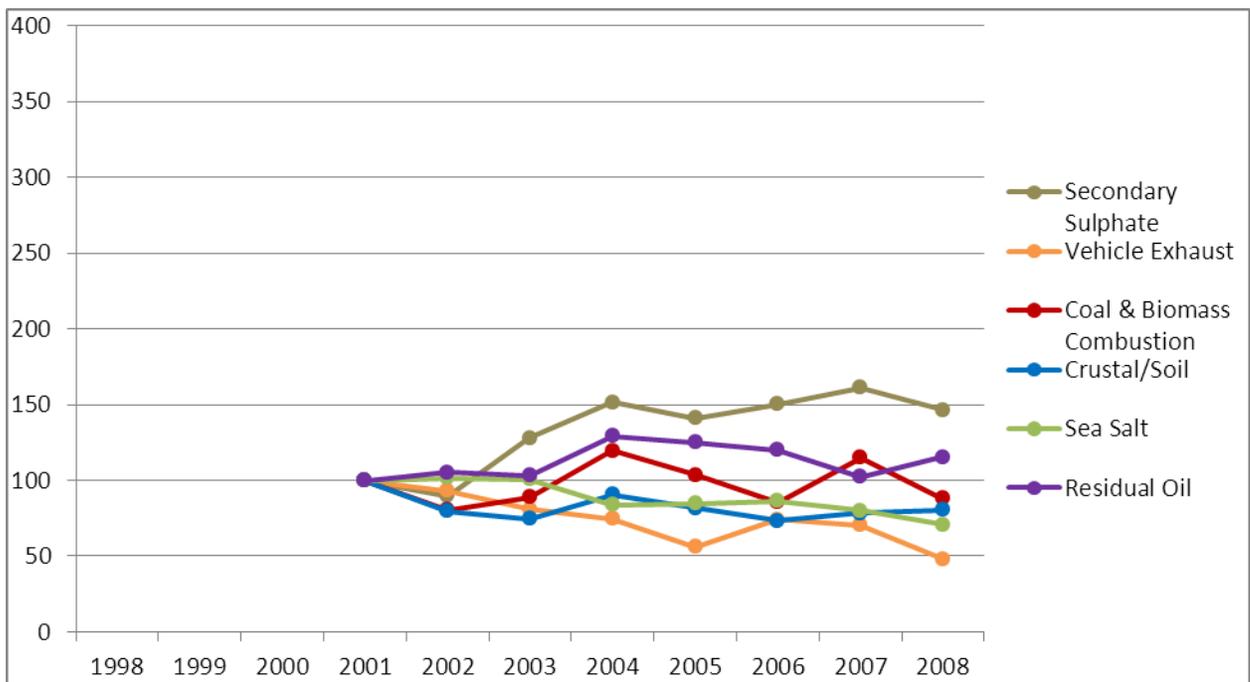


(b) Roadside Station

**Figure 3.17** APCA: Relative annual contribution variation of identified sources with respect to their contribution in 1998 for the general stations (a) and roadside station (b).

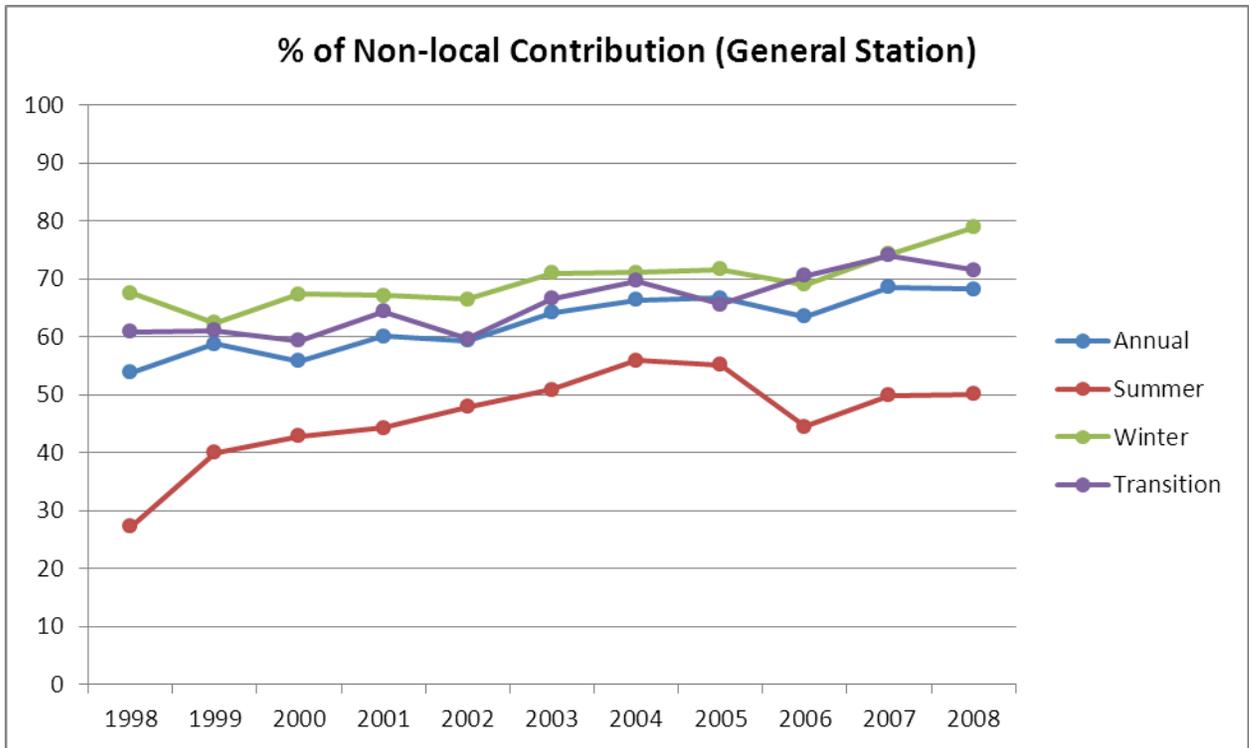


(a) General Station

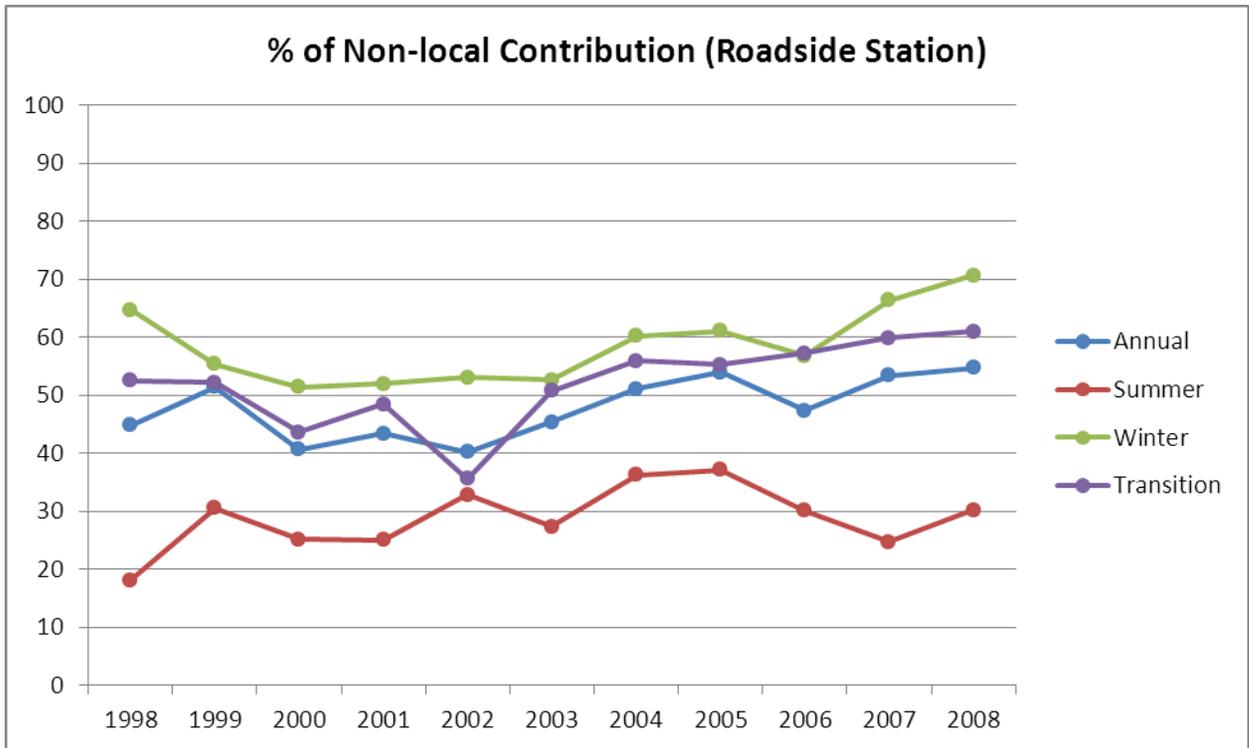


(b) Roadside Station

**Figure 3.18** UNMIX: Relative annual contribution variation of identified sources with respect to their contribution in 1998 for the general stations (a) and with respect to their contribution in 2001 at the roadside station (b).

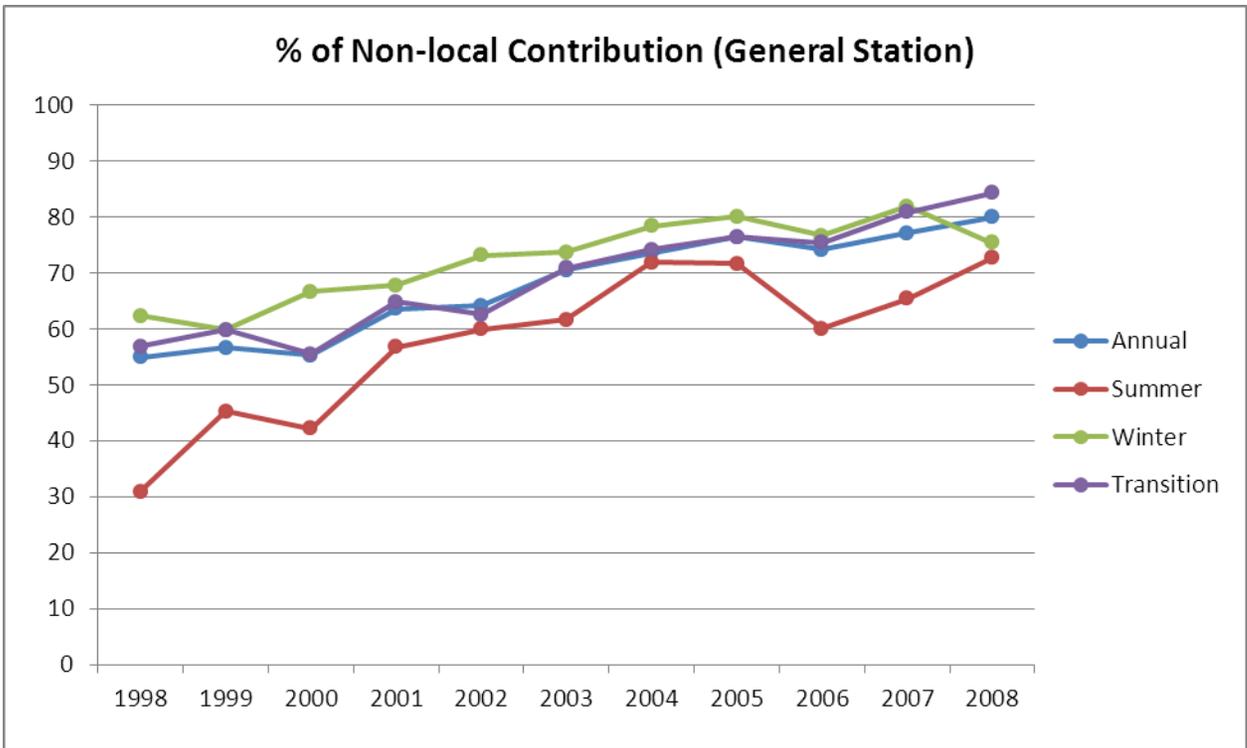


(a) General Station

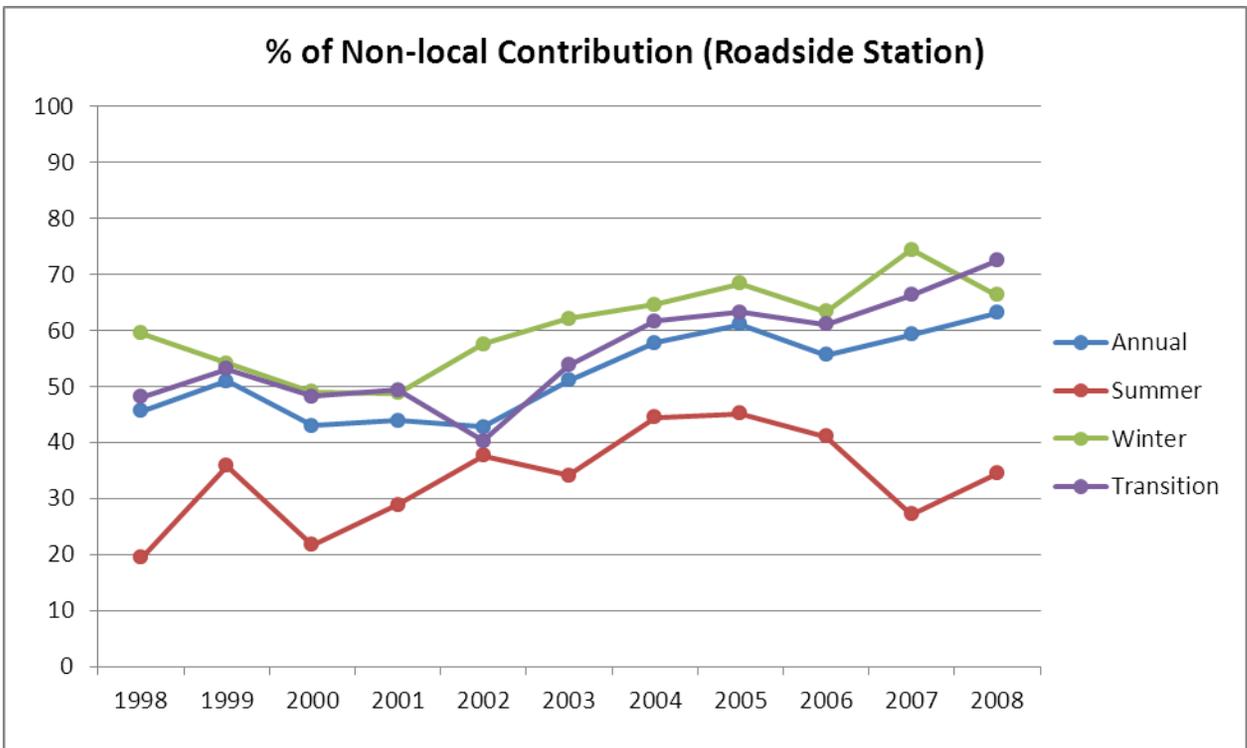


(b) Roadside station

**Figure 3.19** APCA - Long term annual and seasonal trends of relative contribution from non-local sources at the general stations (a) and roadside station (b).

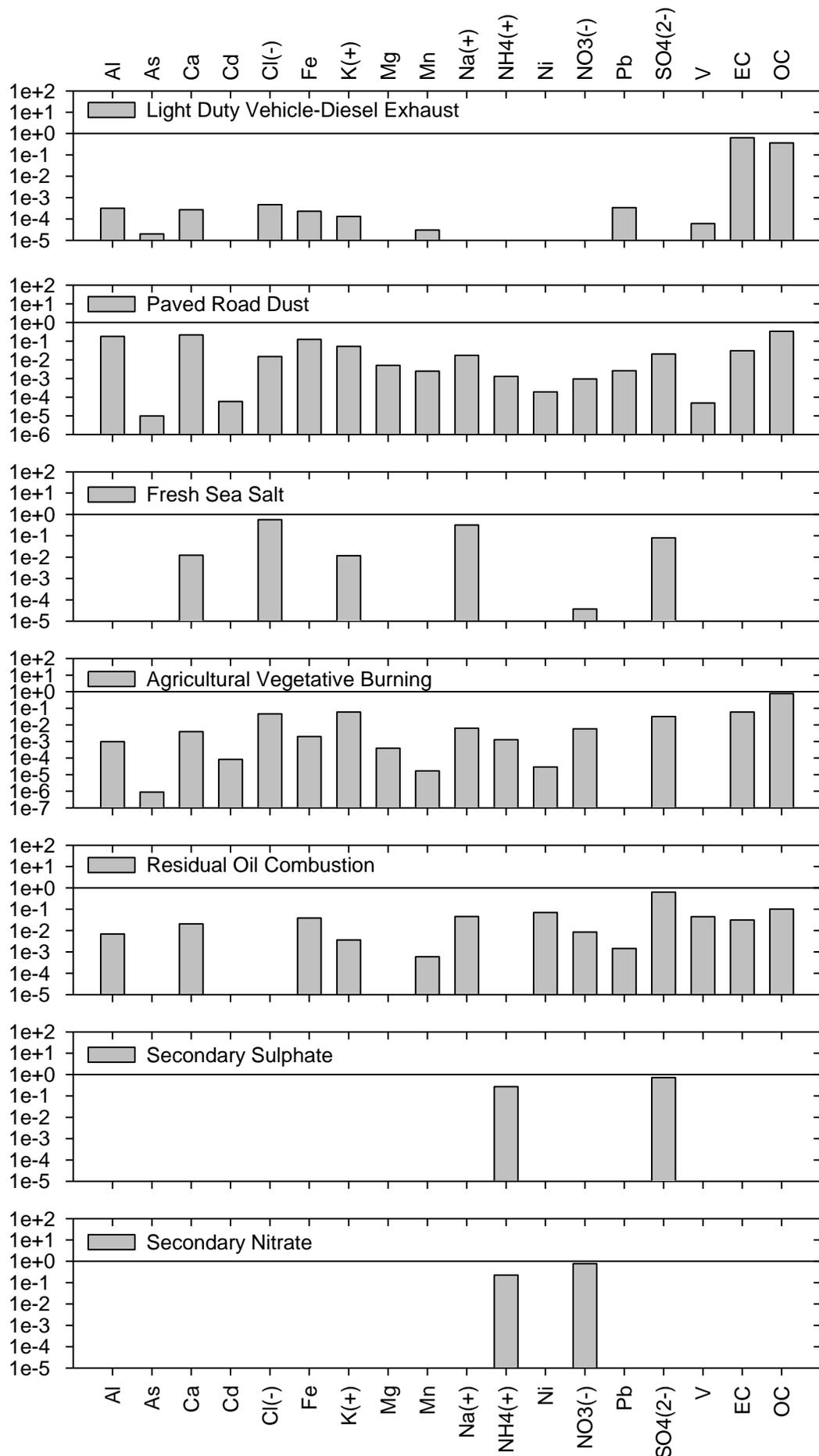


(a) General Station

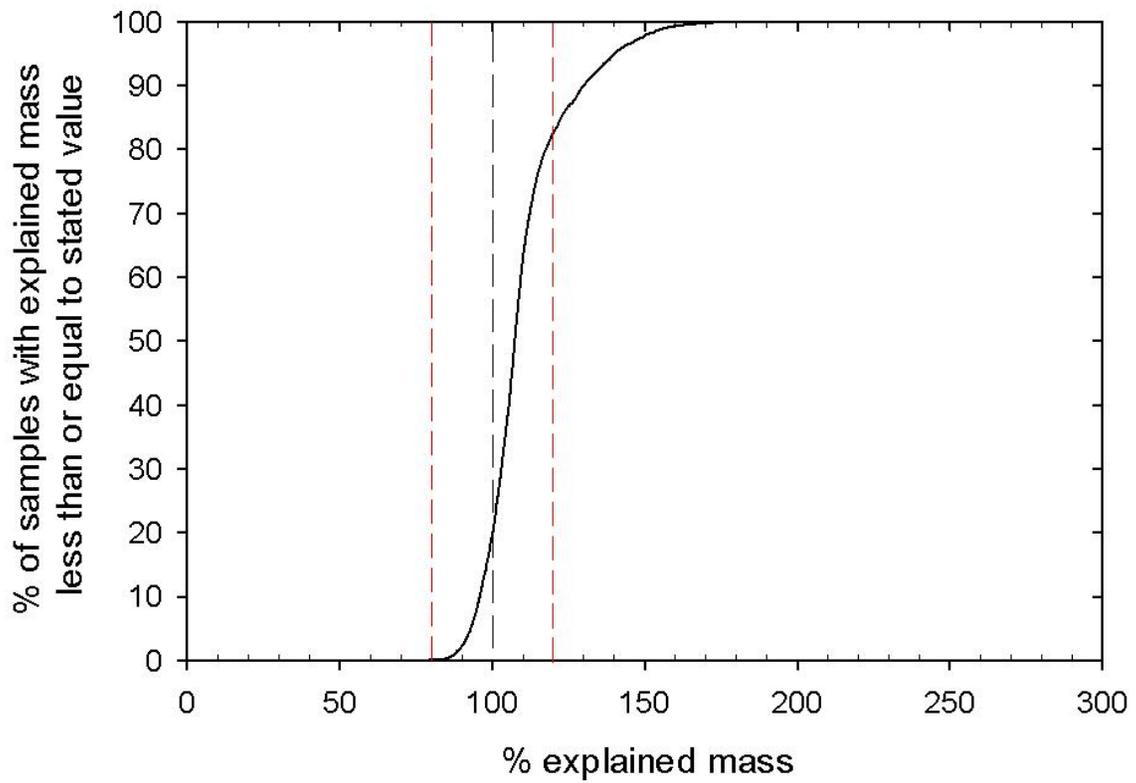


(b) Roadside station

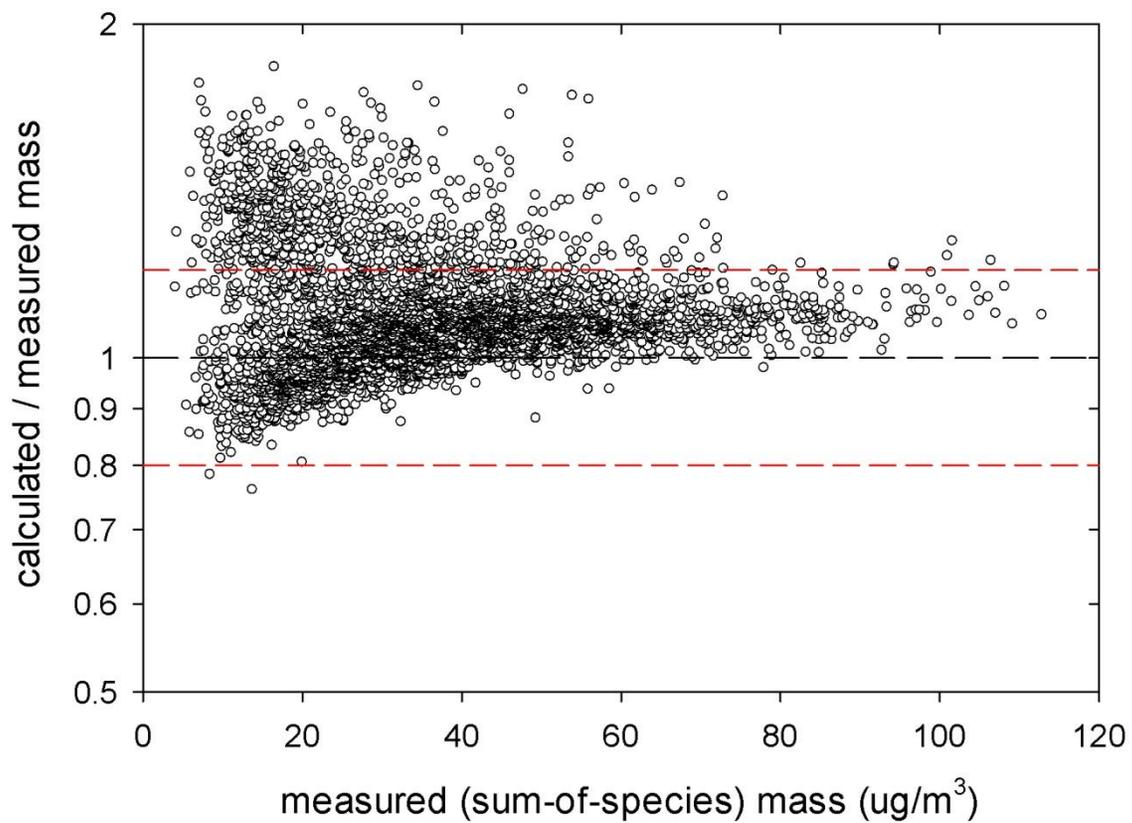
**Figure 3.20** UNMIX - Long term annual and seasonal trends of relative contribution from non-local sources at the general stations (a) and roadside station (b).



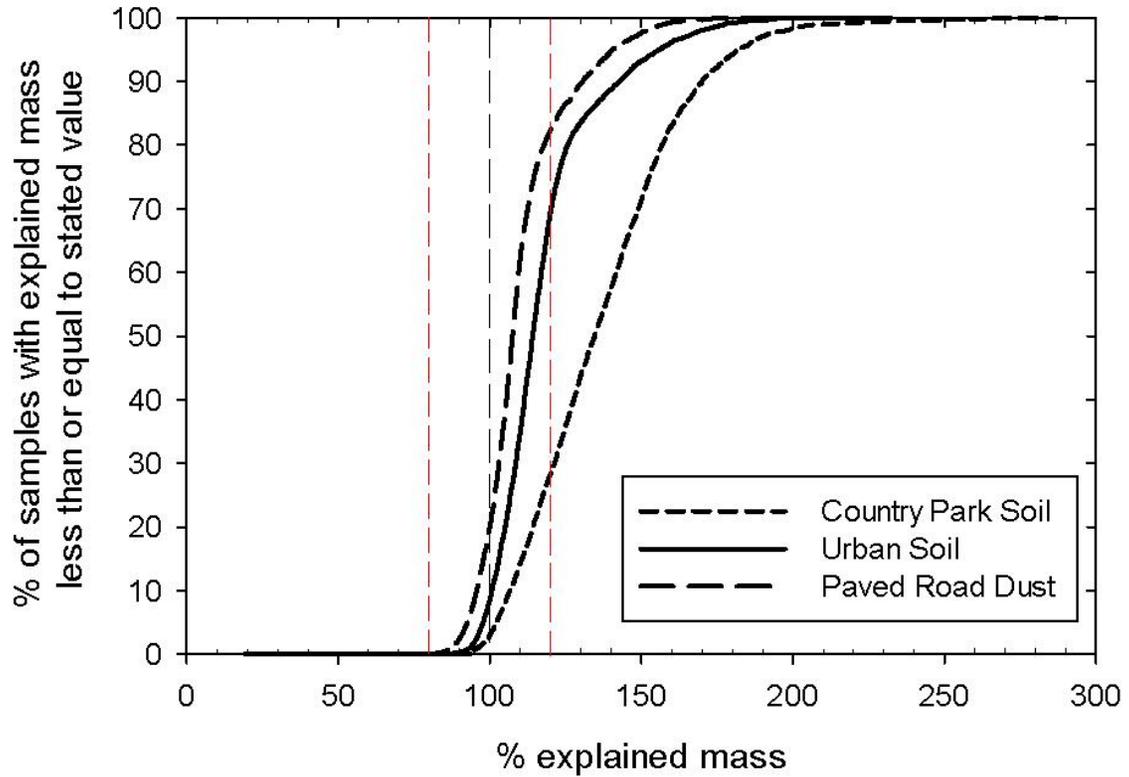
**Figure 3.21** Rescaled source profiles used in the CMB modeling



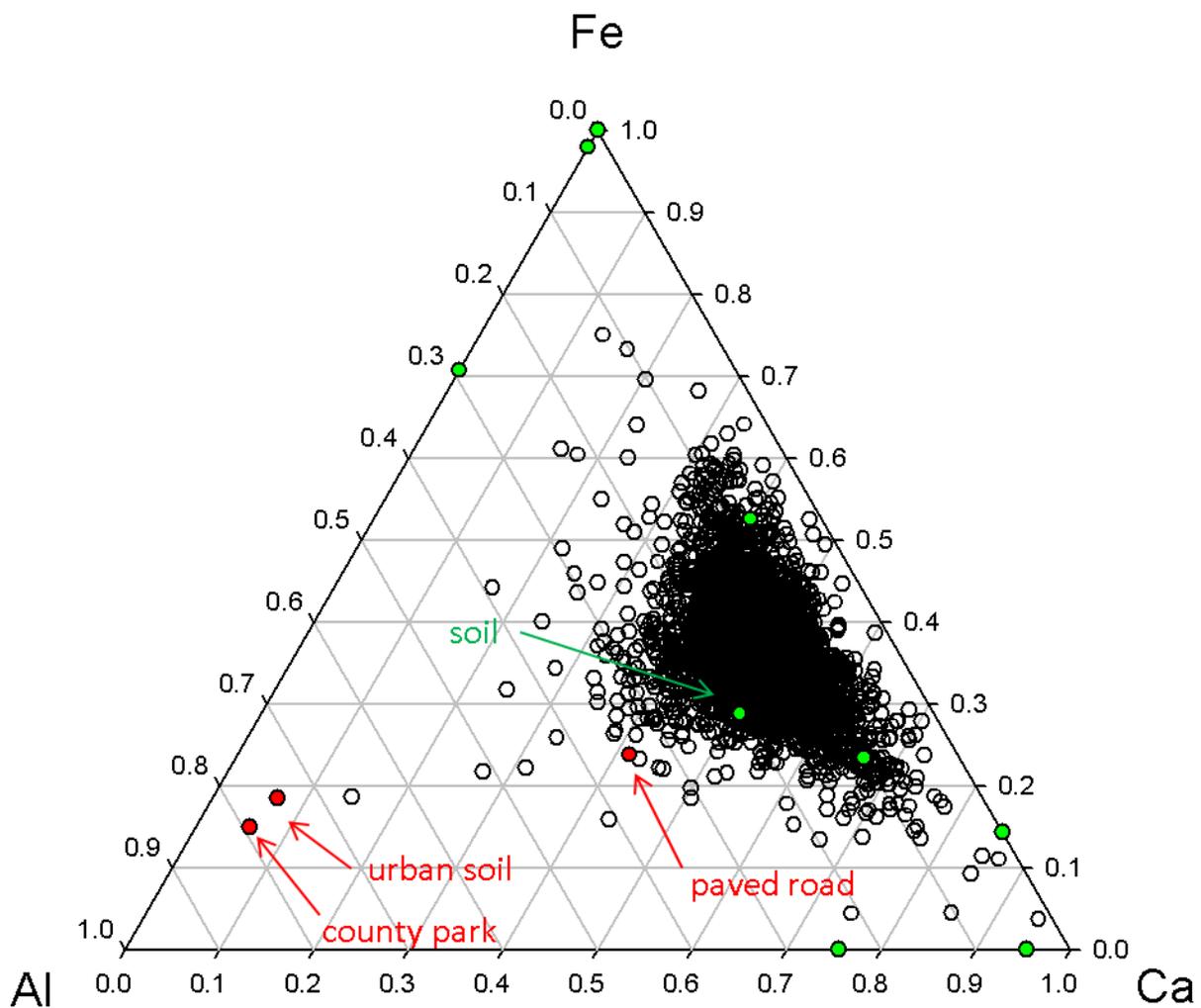
**Figure 3.22** Cumulative distribution of the ratio of the percentage mass explained (ratio of the CMB-modeled mass to the measured sum-of-species mass). The vertical red lines correspond to the mass closure within  $\pm 20\%$ .



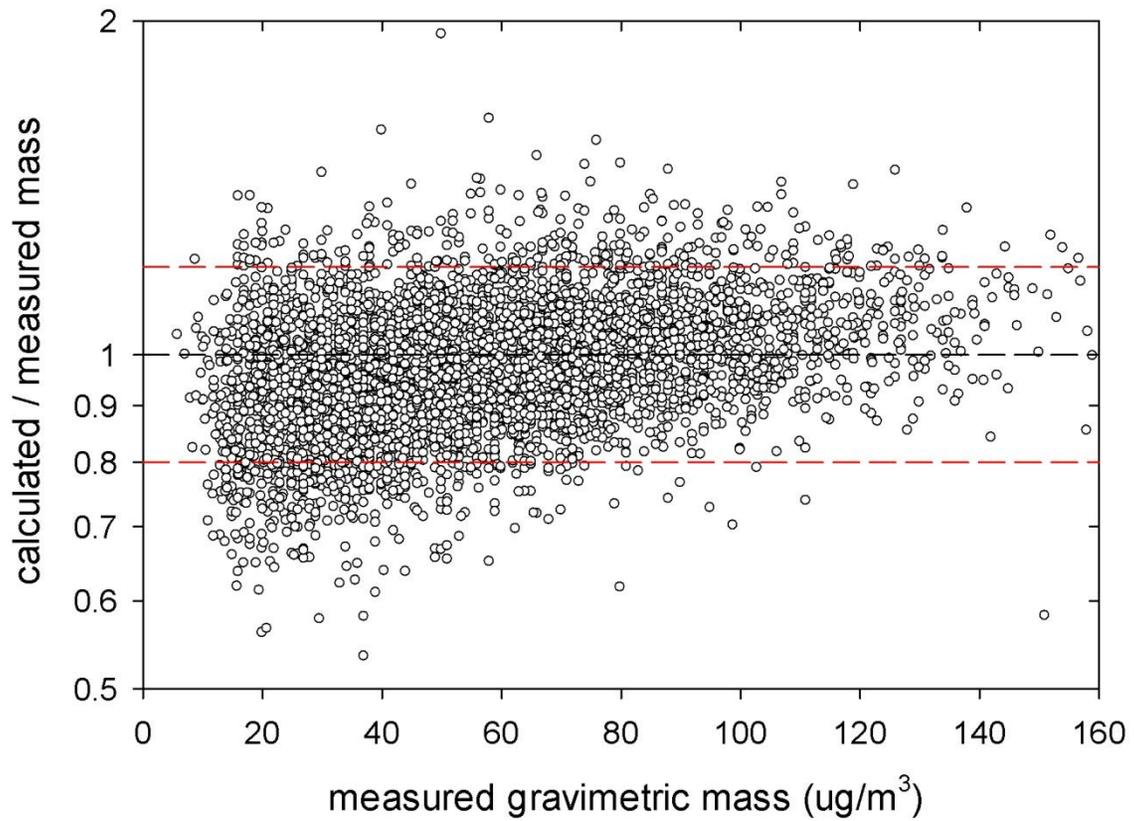
**Figure 3.23** Ratio of the CMB-modeled mass to the measured sum-of-species mass as a function of the measured sum-of-species mass. The horizontal red lines correspond to the mass closure within  $\pm 20\%$ .



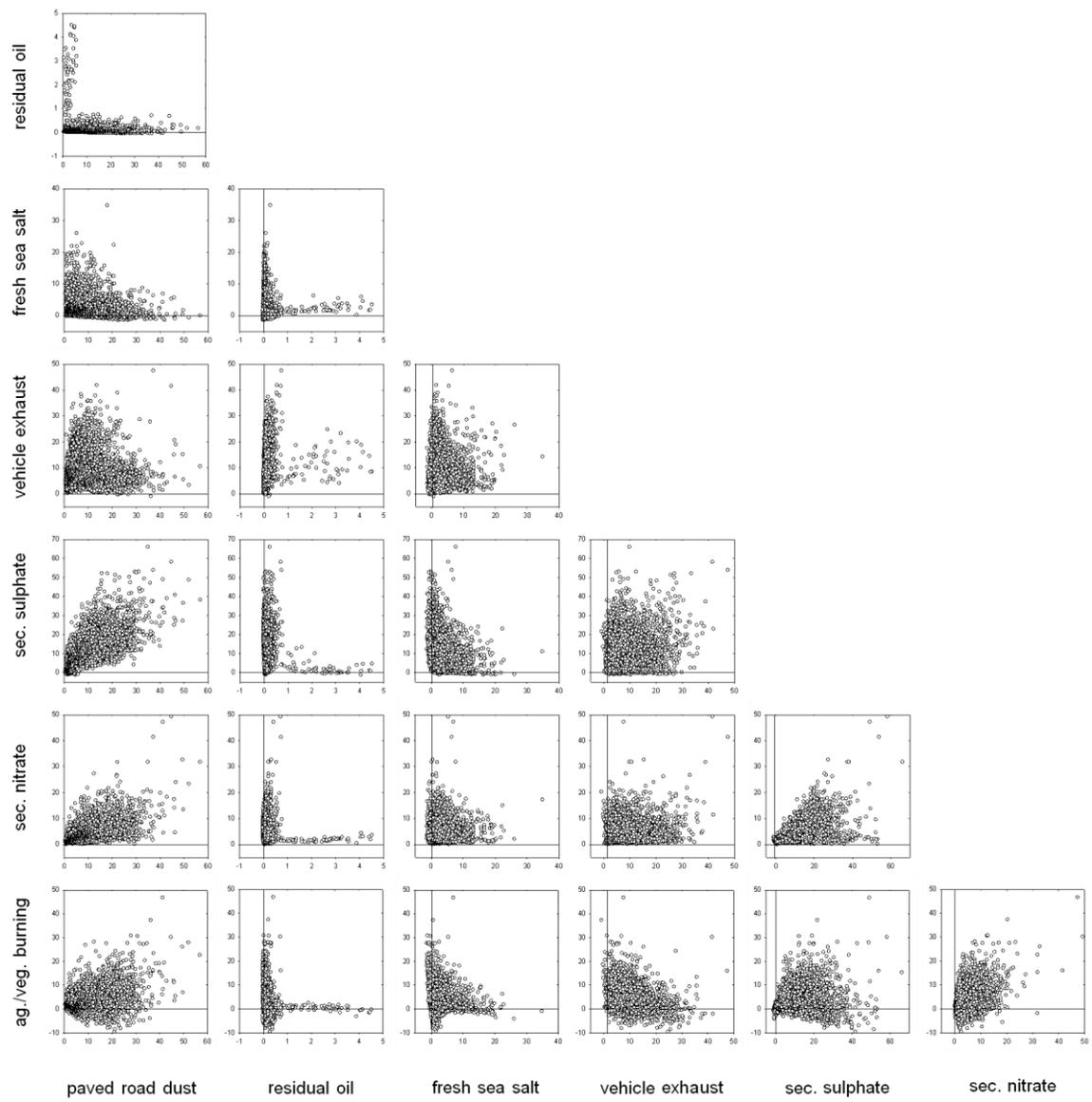
**Figure 3.24** Cumulative distribution of the ratio of the percentage mass explained (ratio of the CMB-modeled mass to the measured sum-of-species mass) for modeling with three different fugitive dust profiles: paved road dust (pink); urban soil (green); and county park soil (blue). The vertical red lines correspond to the mass closure within  $\pm 20\%$ .



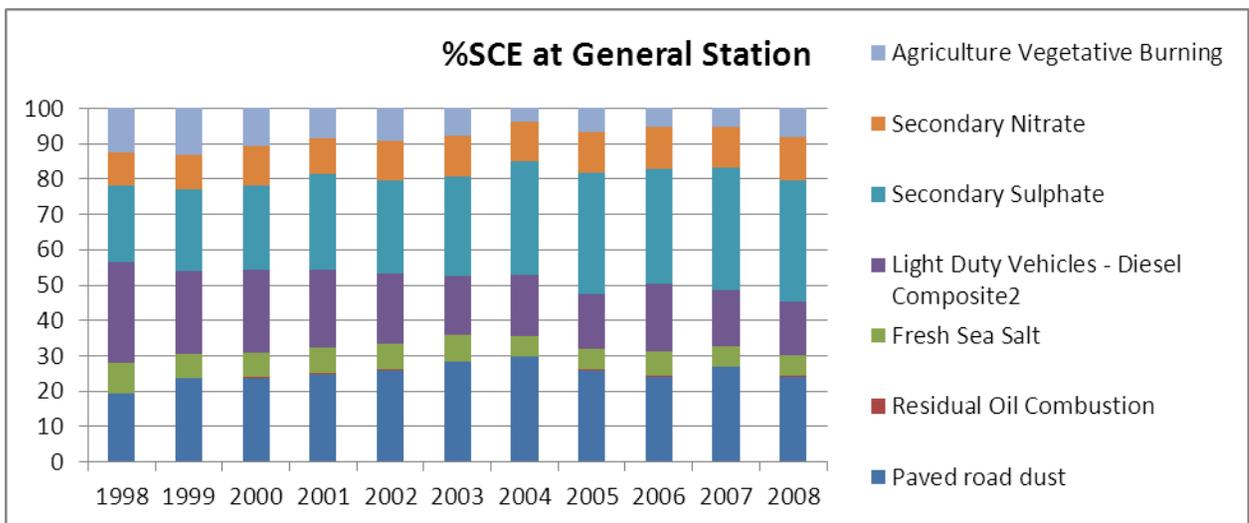
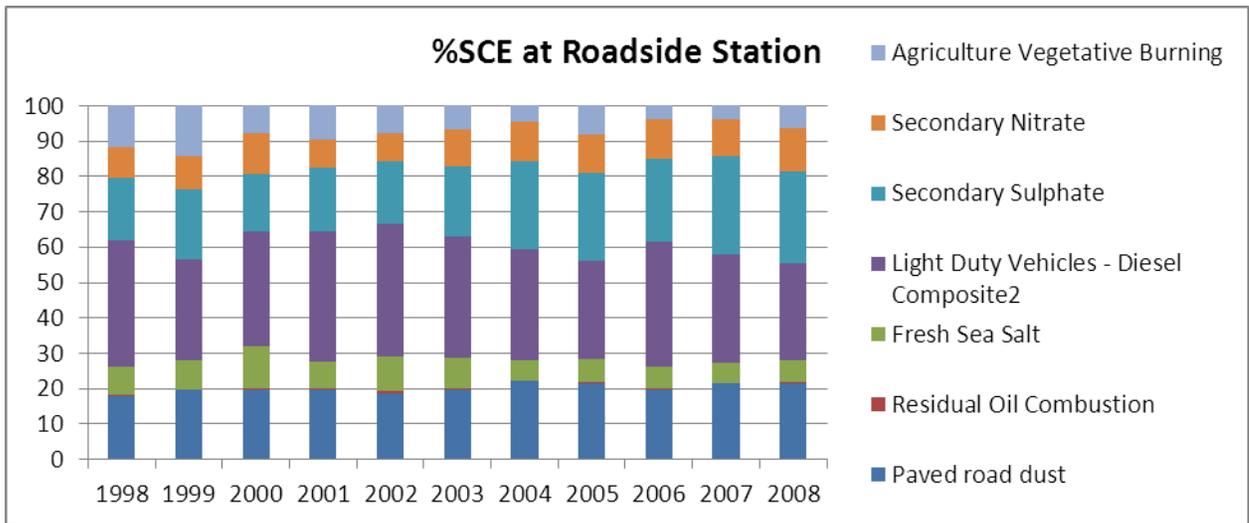
**Figure 3.25** Relative distributions of Al, Ca and Fe in the Hong Kong ambient PM<sub>10</sub> samples (open circles), PMF-resolved factors (green circles) and PM<sub>10</sub> source profiles of Ho *et al.* 2003 (red circles).



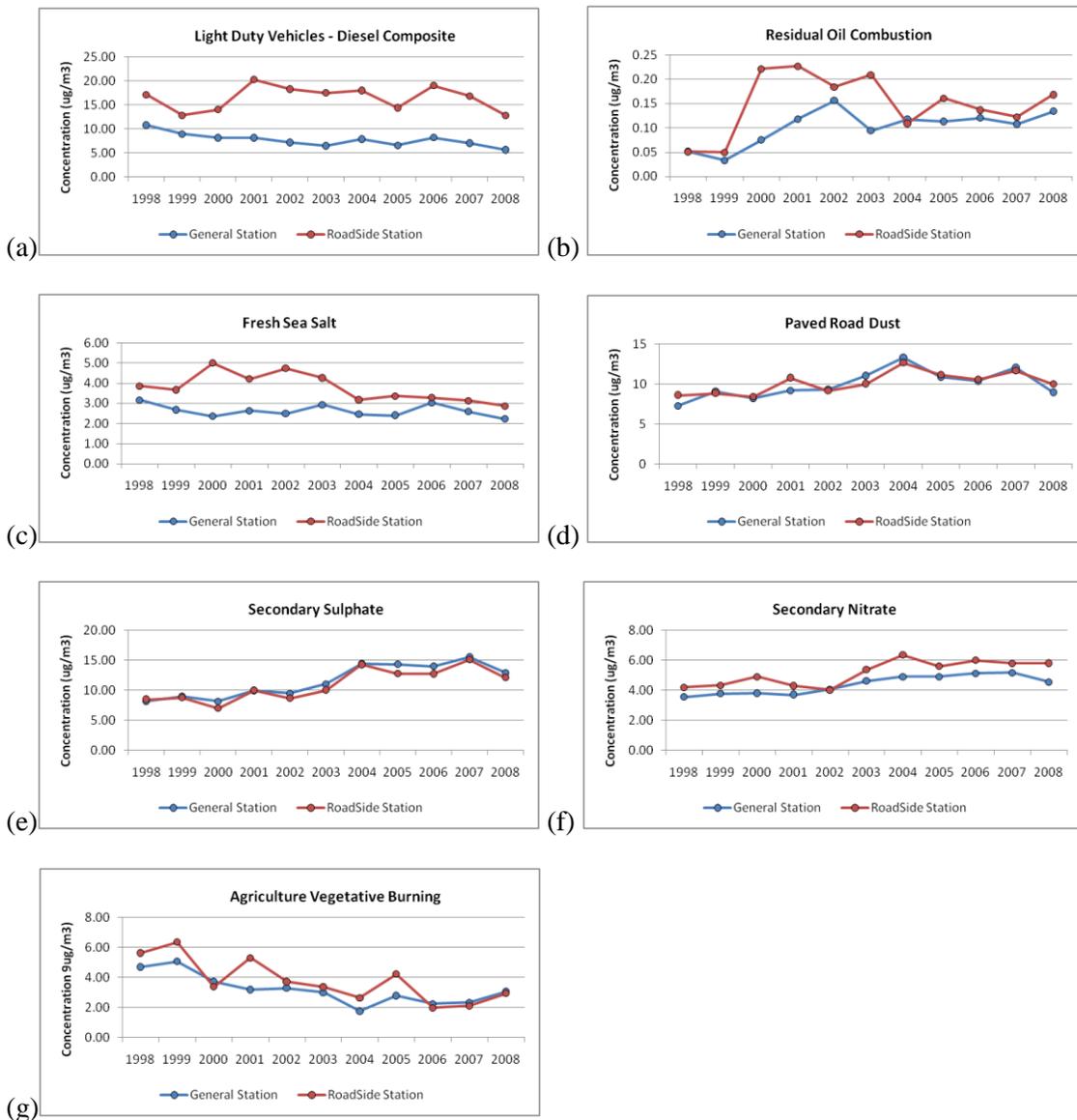
**Figure 3.26** Ratio of the CMB-modeled mass to the measured gravimetric mass as a function of the measured gravimetric mass after doubling the SCE for paved road dust. The horizontal red lines correspond to the mass closure within  $\pm 20\%$ .



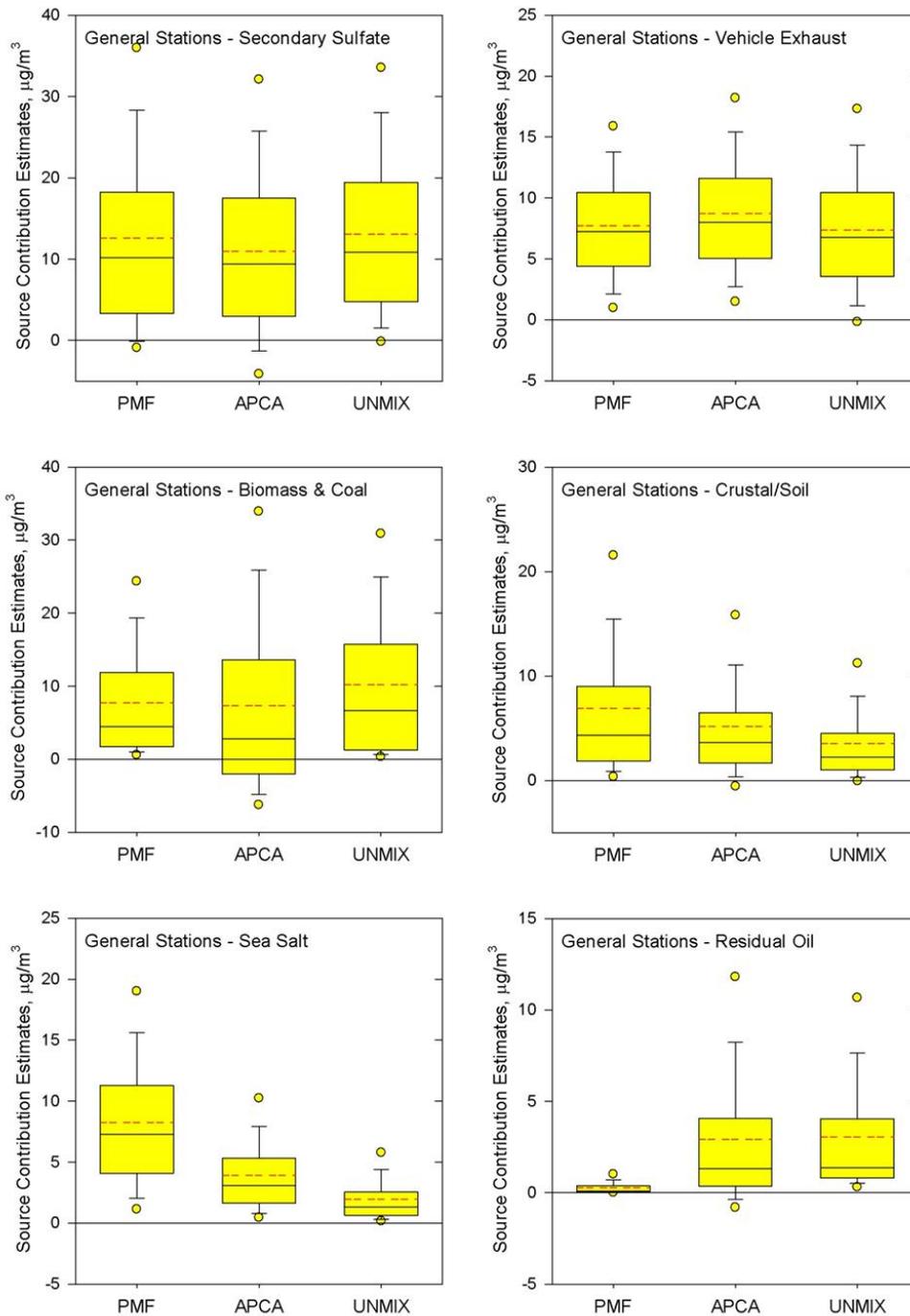
**Figure 3.27** G-space plots for the base case CMB-modeled source contribution estimates. All concentration values in  $\mu\text{g}/\text{m}^3$ .



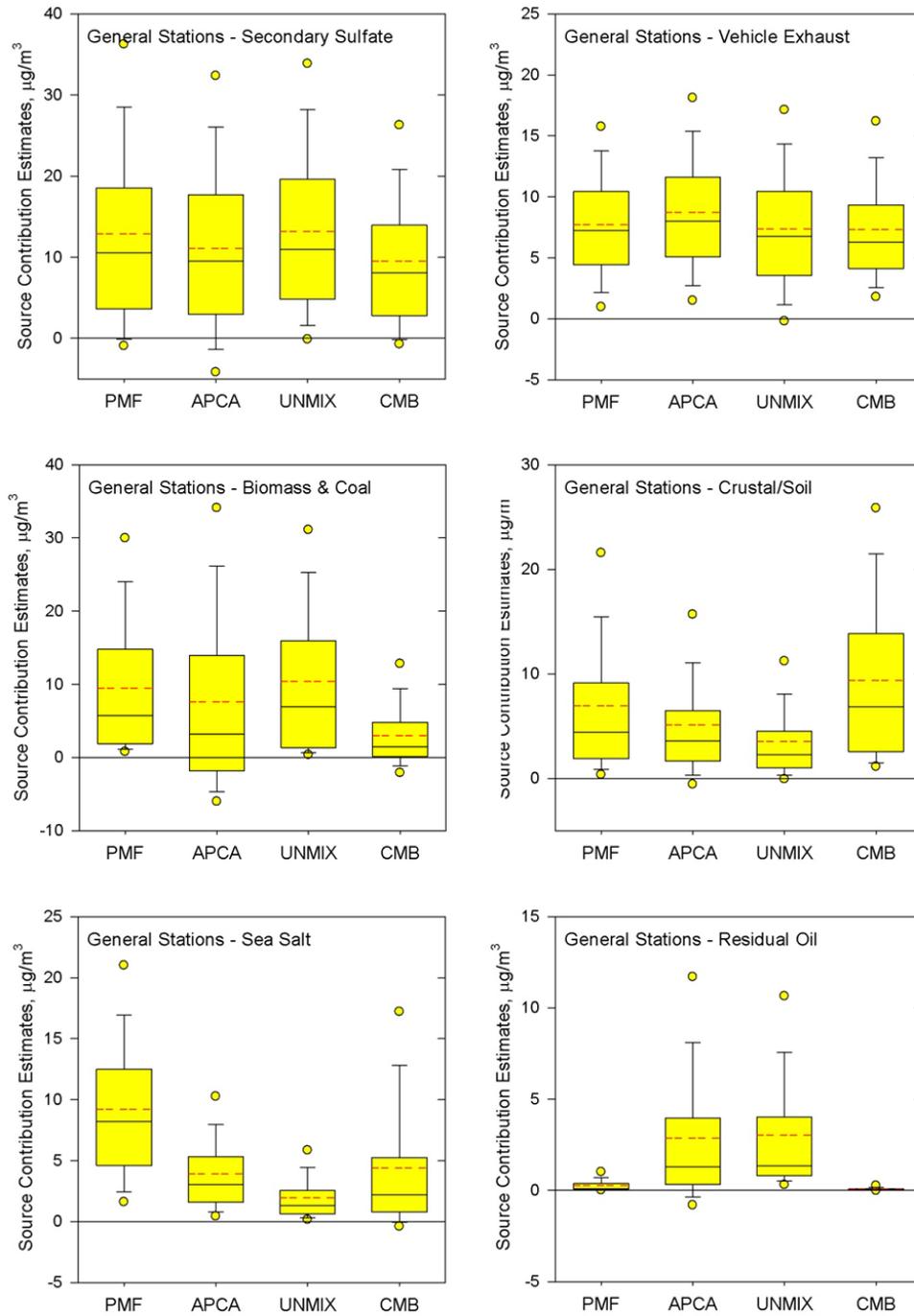
**Figure 3.28** Study-average source contribution estimates as a percentage of modeled mass for the roadside station (top) and general stations (bottom).



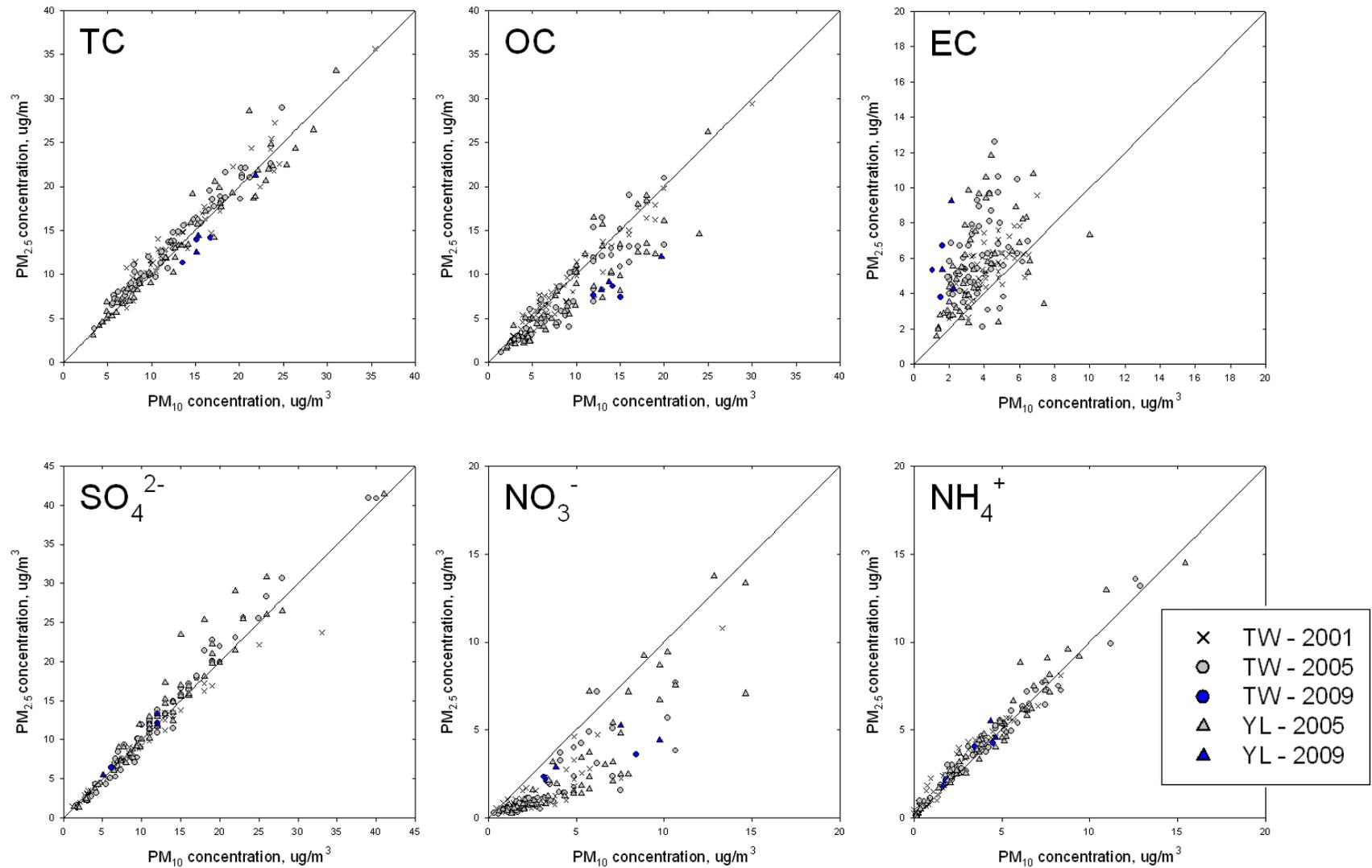
**Figure 3.29** CMB-modeled annual-average source contribution estimates at the roadside stations (red) and general stations (blue) for each emission source category.



**Figure 3.30** Distributions of PM<sub>10</sub> source contribution estimates at the general stations using PMF, APCA and UNMIX. The PMF samples have been censored to include only those samples included in the APCA and UNMIX modeling.



**Figure 3.31** Distributions of  $\text{PM}_{10}$  source contribution estimates at the general stations using PMF, APCA, UNMIX, and CMB. The PMF, APCA and UNMIX samples have been censored to include only those samples passing the CMB validation criteria.



**Figure 4.1** PM<sub>2.5</sub> versus PM<sub>10</sub> concentrations for paired sample days at the at the TW (2001, 2005, 2009) and YL (2005, 2009) stations.

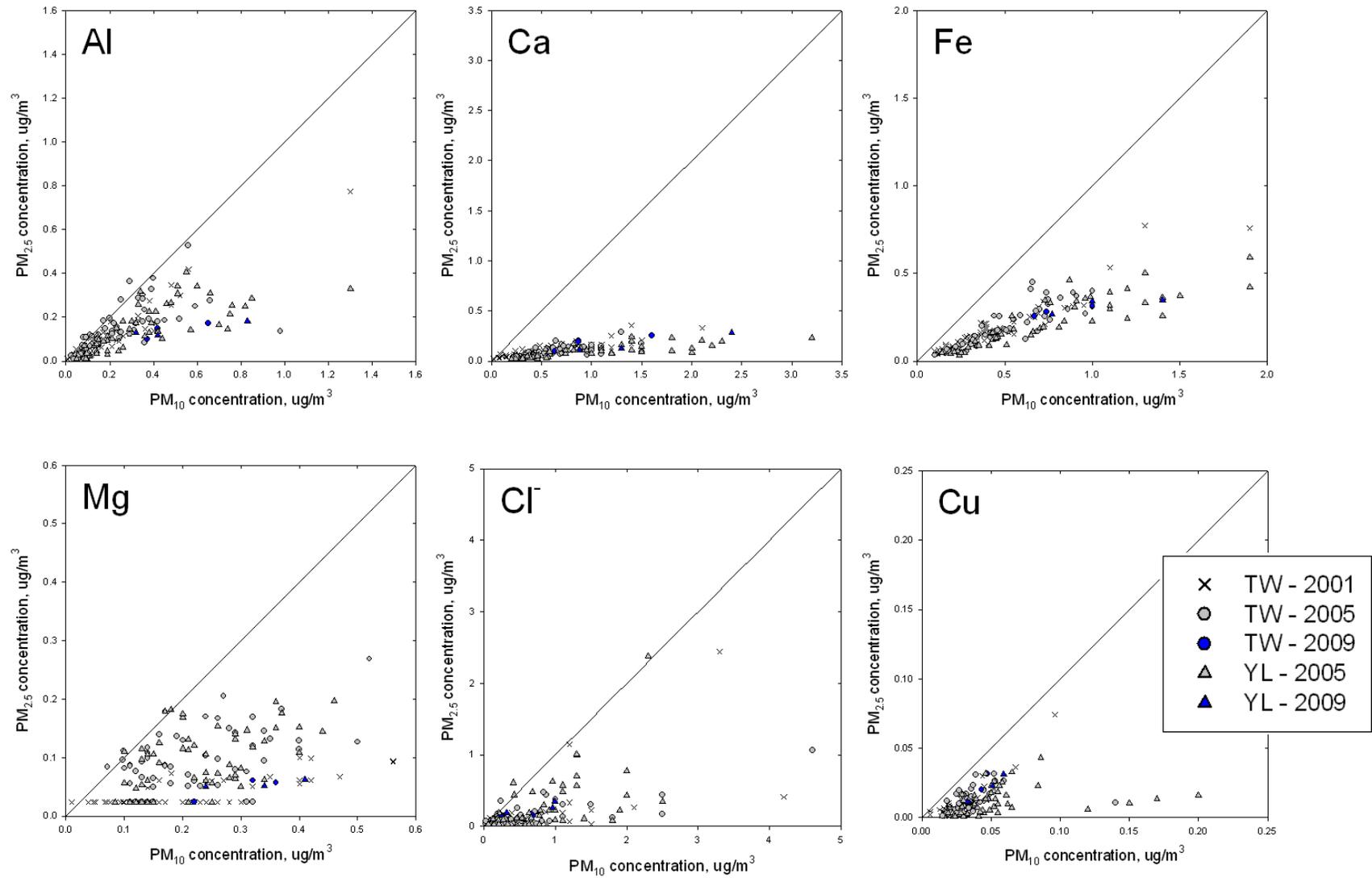


Figure 4.1 (continued)

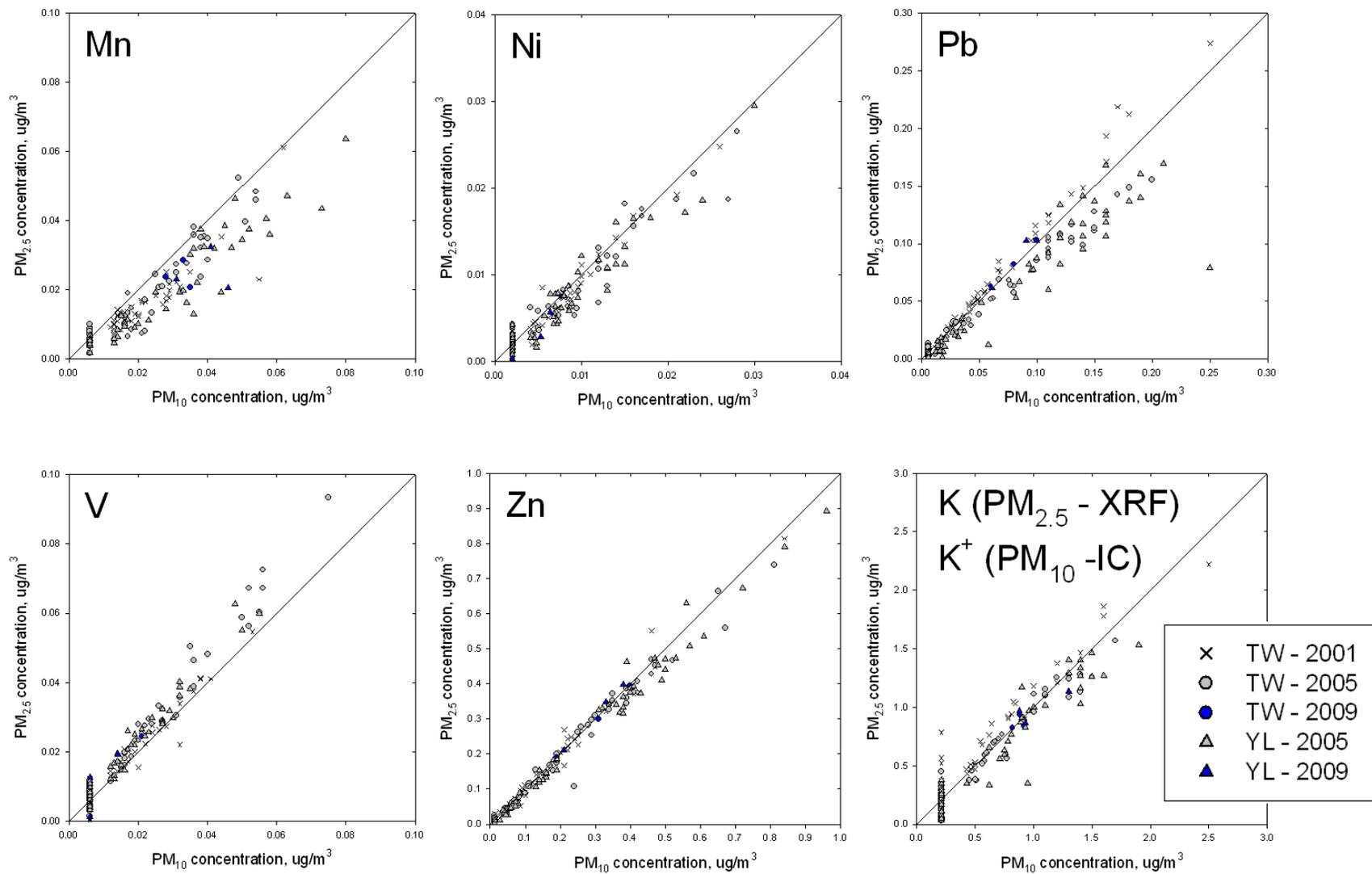
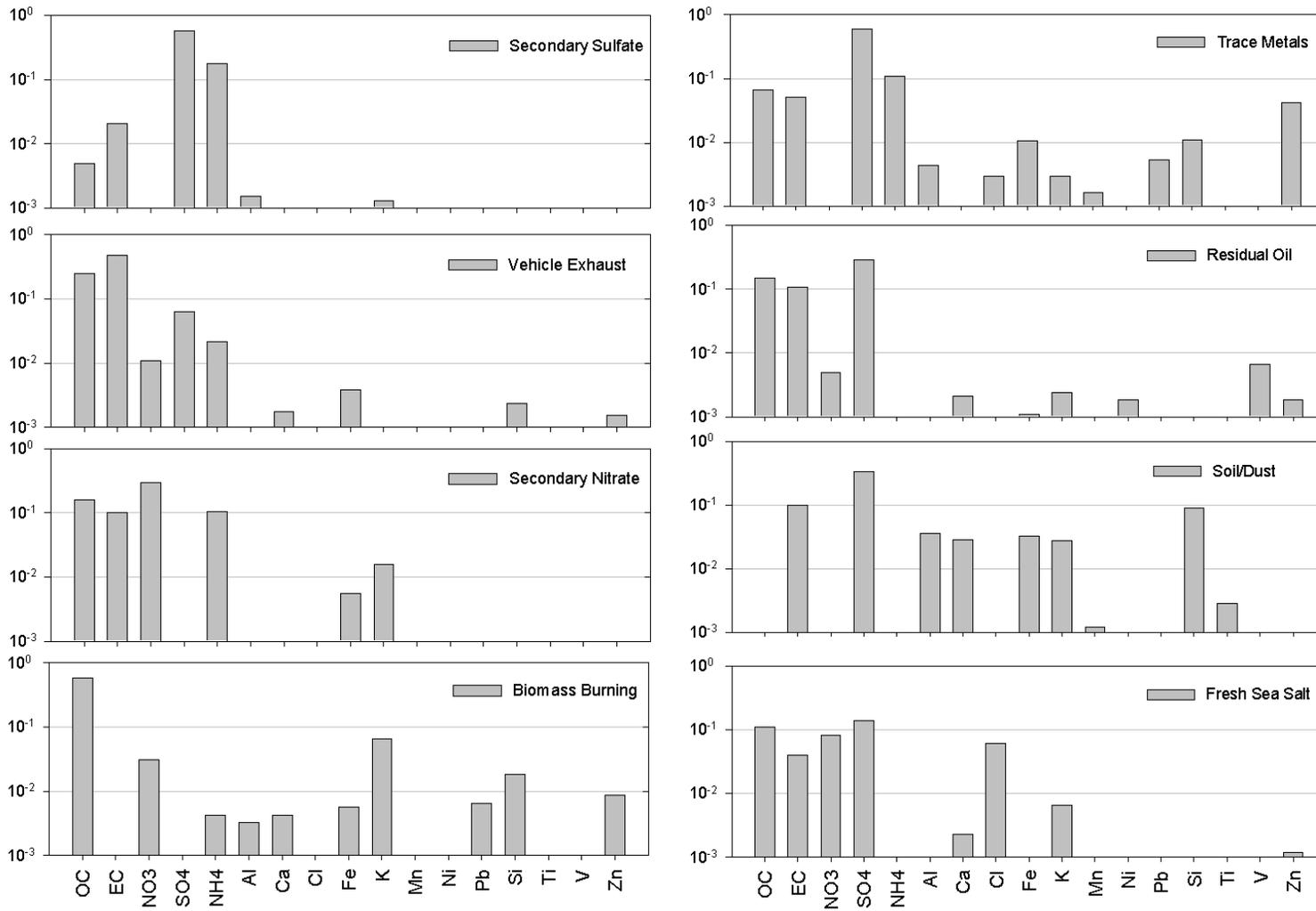
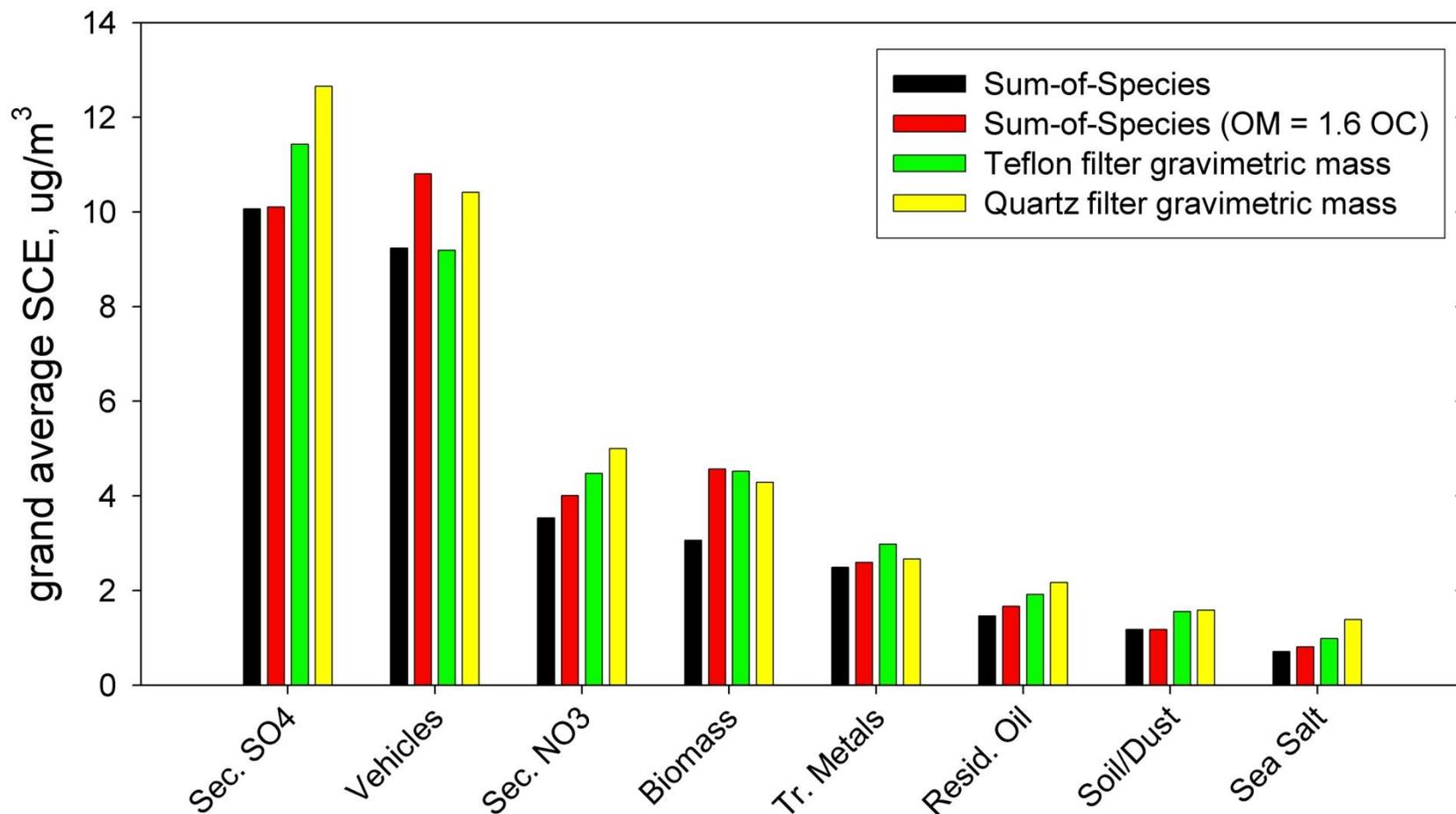


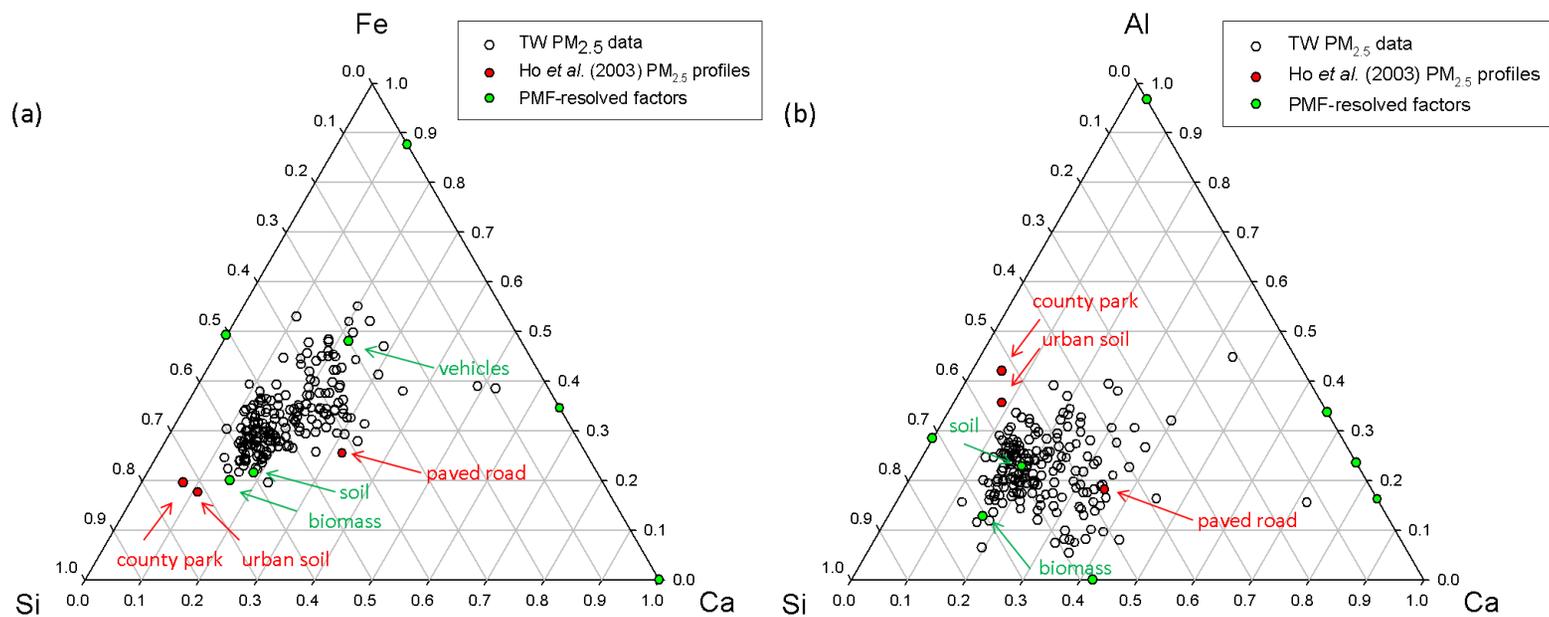
Figure 4.1 (continued)



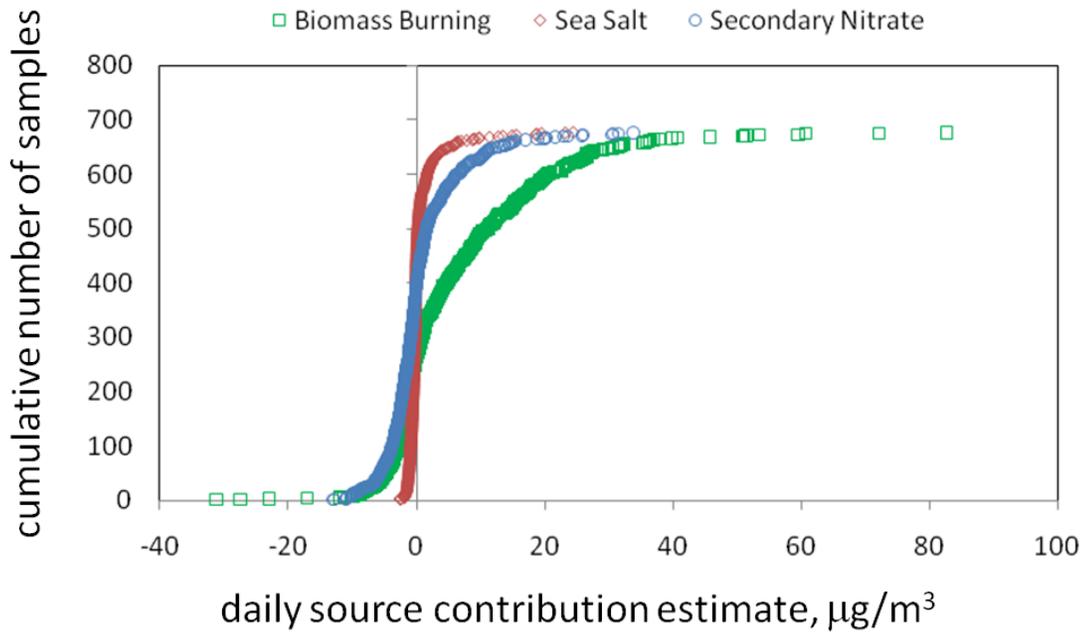
**Figure 4.2** PM<sub>2.5</sub> source profiles – expressed as mass fractions – for the eight PMF-resolved factors.



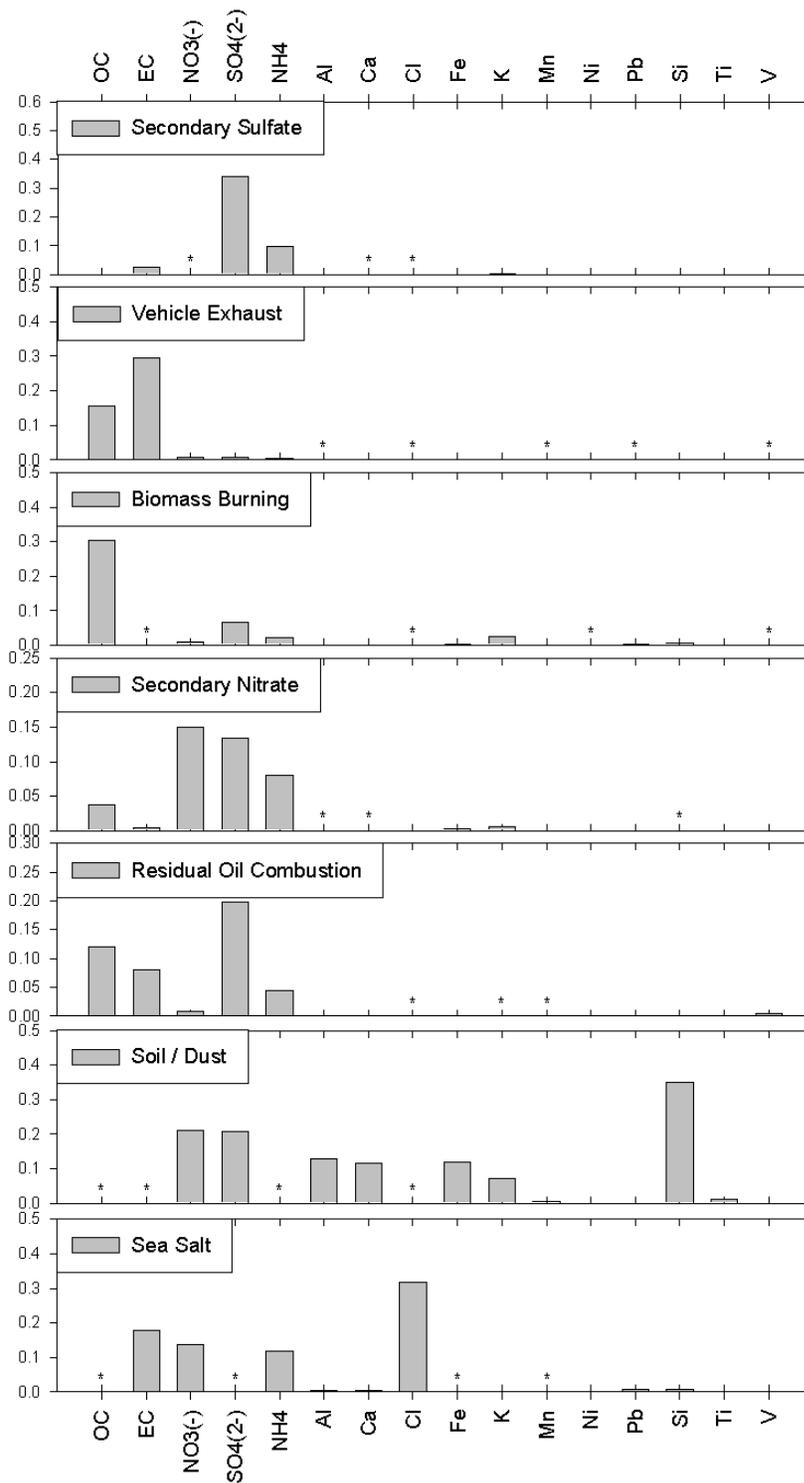
**Figure 4.3** Grand average PM<sub>2.5</sub> mass source contribution estimates – averaged over all sites and all measurement periods – for the eight PMF-resolved factors using: sum-of-species with OC; sum of species with OM = 1.6× OC; Teflon filter gravimetric mass as the Total Variable in PMF; and Quartz filter gravimetric mass as the Total Variable in PMF.



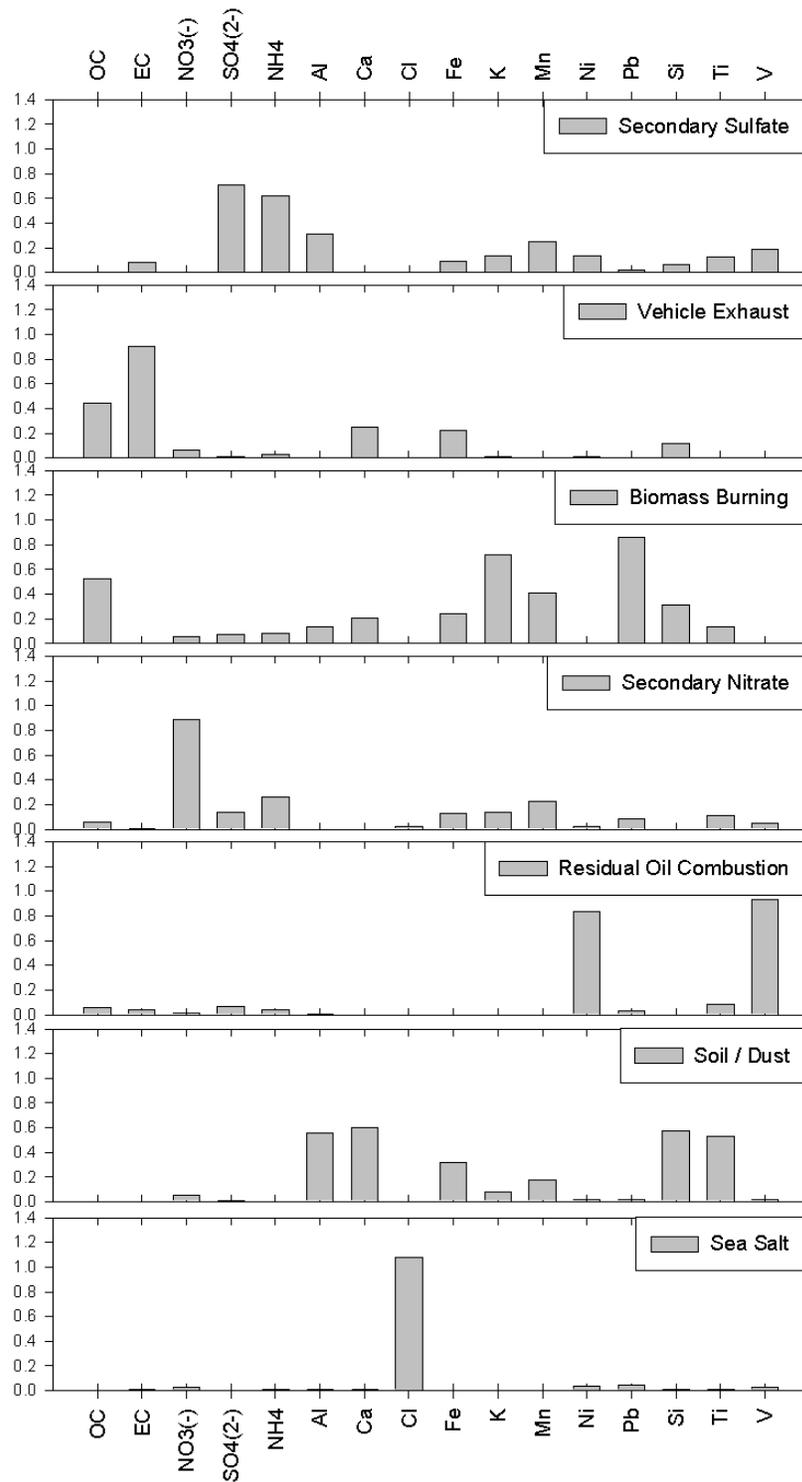
**Figure 4.4** Relative distributions of Ca, Fe, and Si (a) and Al, Ca, and Si (b) in the TW ambient samples (open circles), PMF-resolved factors (green circles) and PM<sub>2.5</sub> source profiles of Ho *et al.* 2003 (red circles).



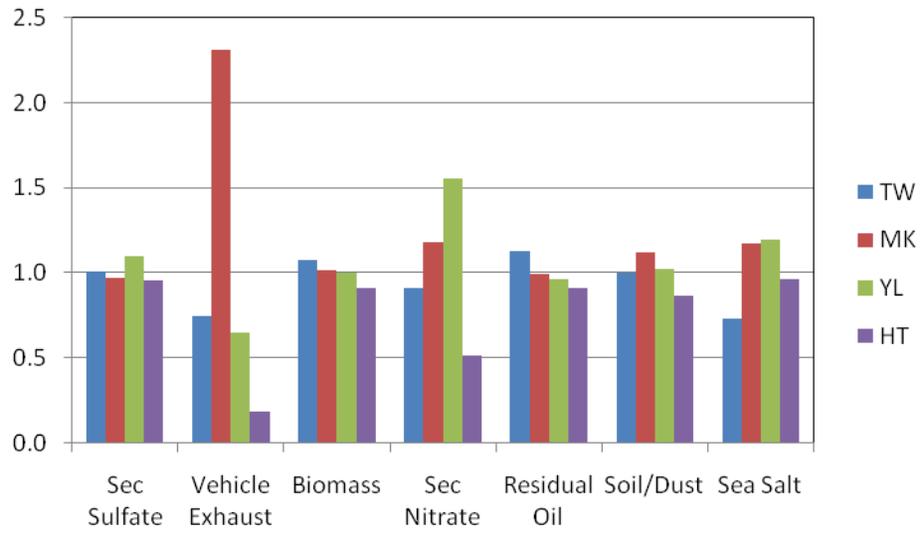
**Figure 4.5** Cumulative distribution of APCA-modeled  $\text{PM}_{2.5}$  daily source contribution estimates for biomass burning, seas salt and secondary nitrate factors in the seven-factor solution.



**Figure 4.6** Source profiles (species mass fractions) for the seven factor UNMIX solution of the Hong Kong PM<sub>2.5</sub> data set.

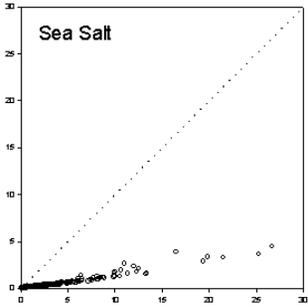
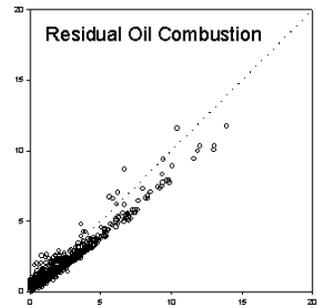
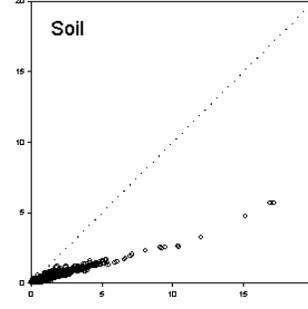
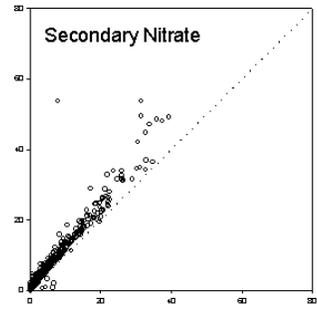
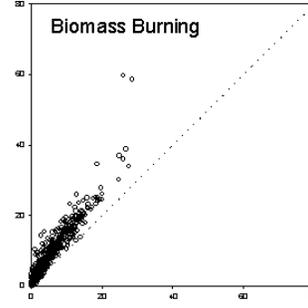
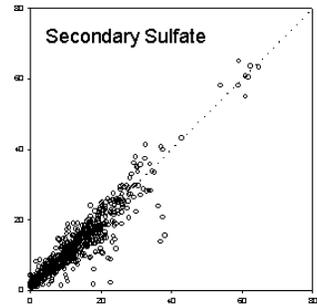
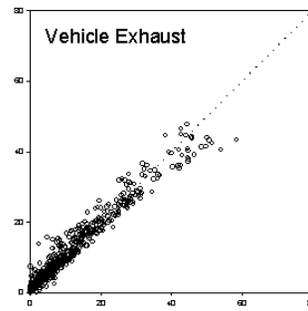
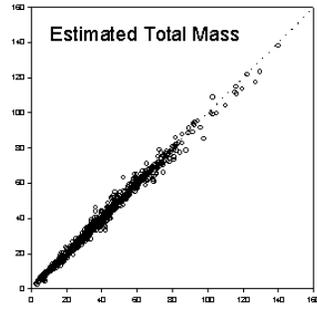


**Figure 4.7** Explained mass profiles (fraction of species mass) for the seven factor UNMIX solution of the Hong Kong PM<sub>2.5</sub> data set.



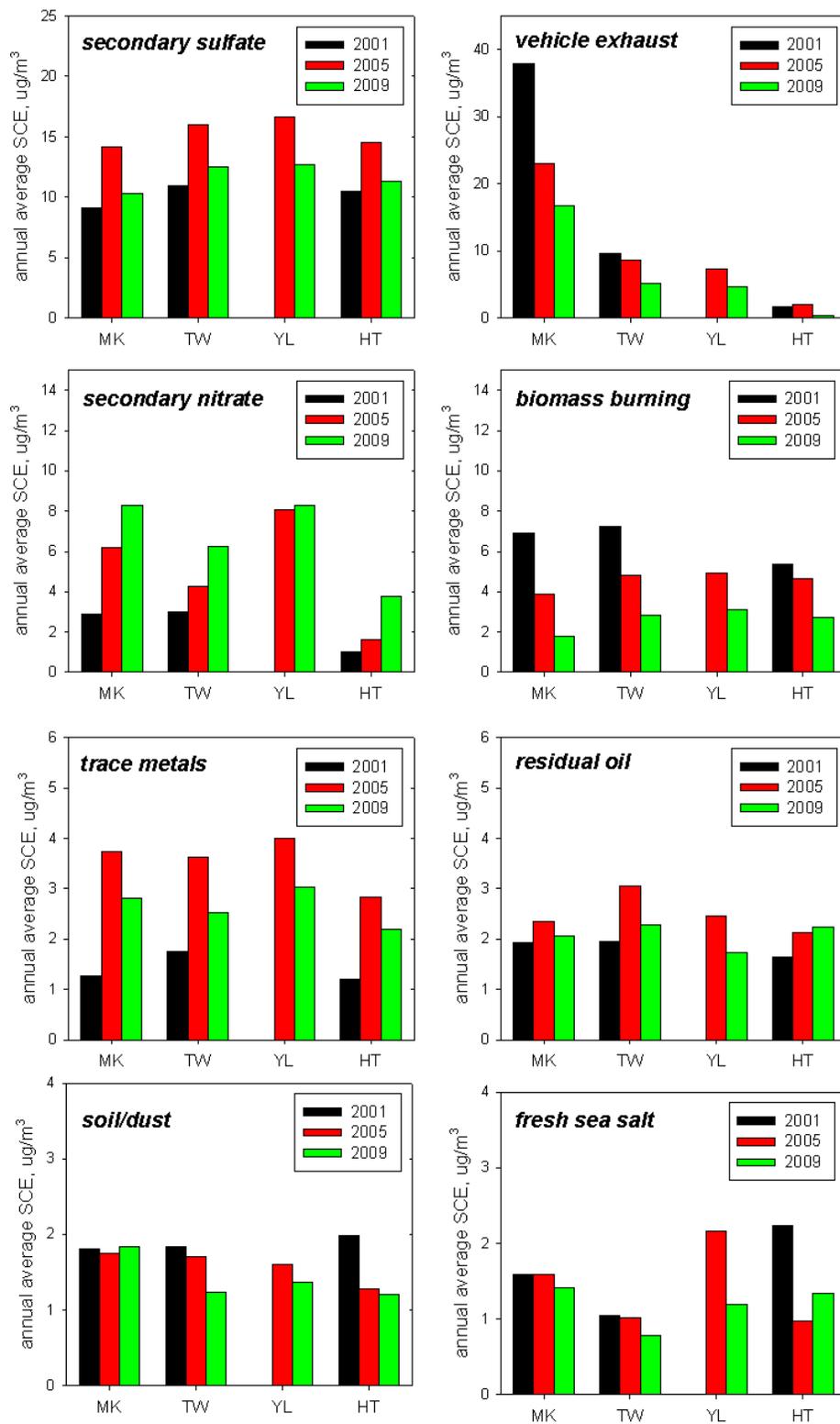
**Figure 4.8** Ratio of the site-specific average SCE to the all-sites average SCE for each of the resolved factors.

UNMIX source contribution estimates

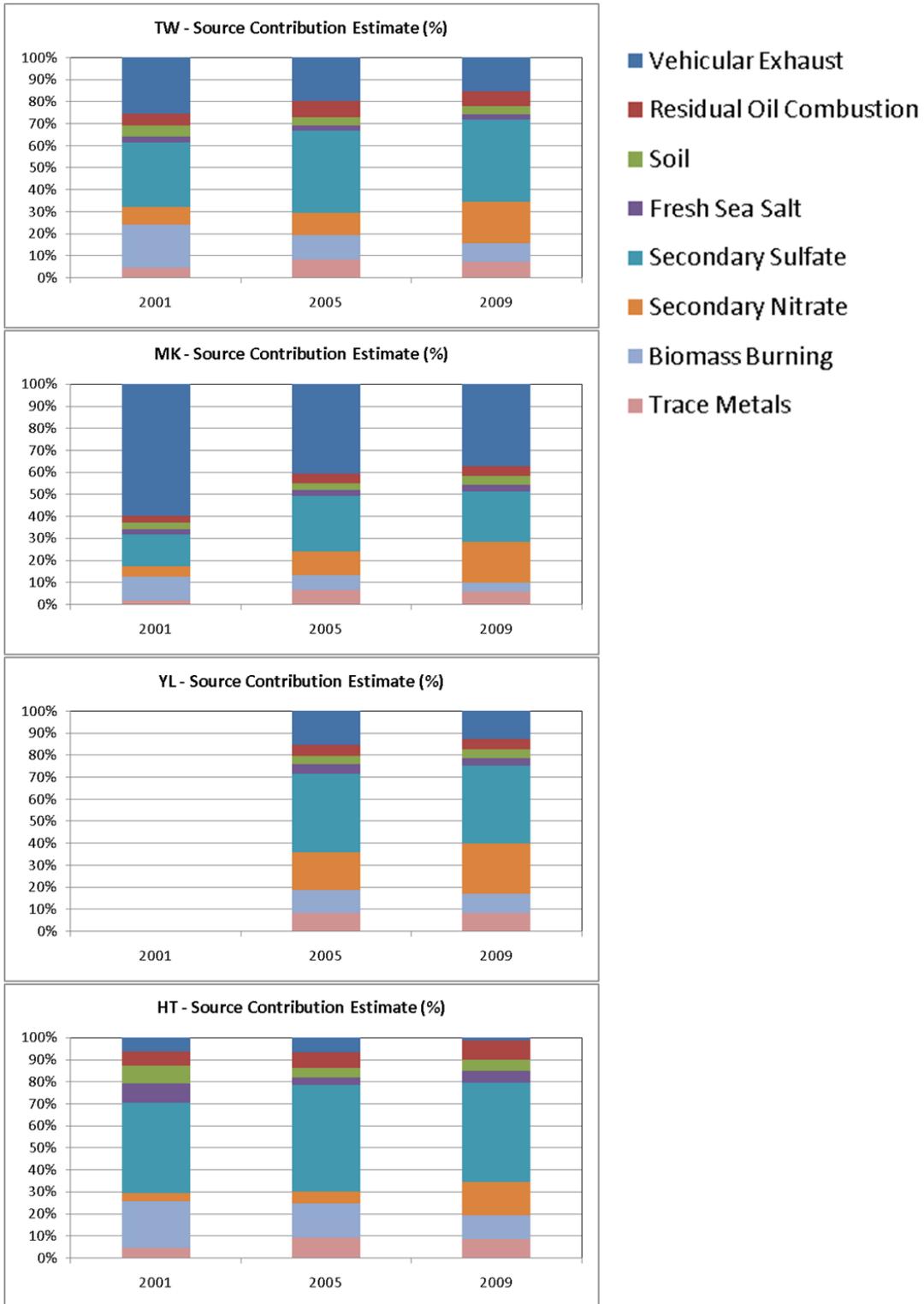


PMF source contribution estimates

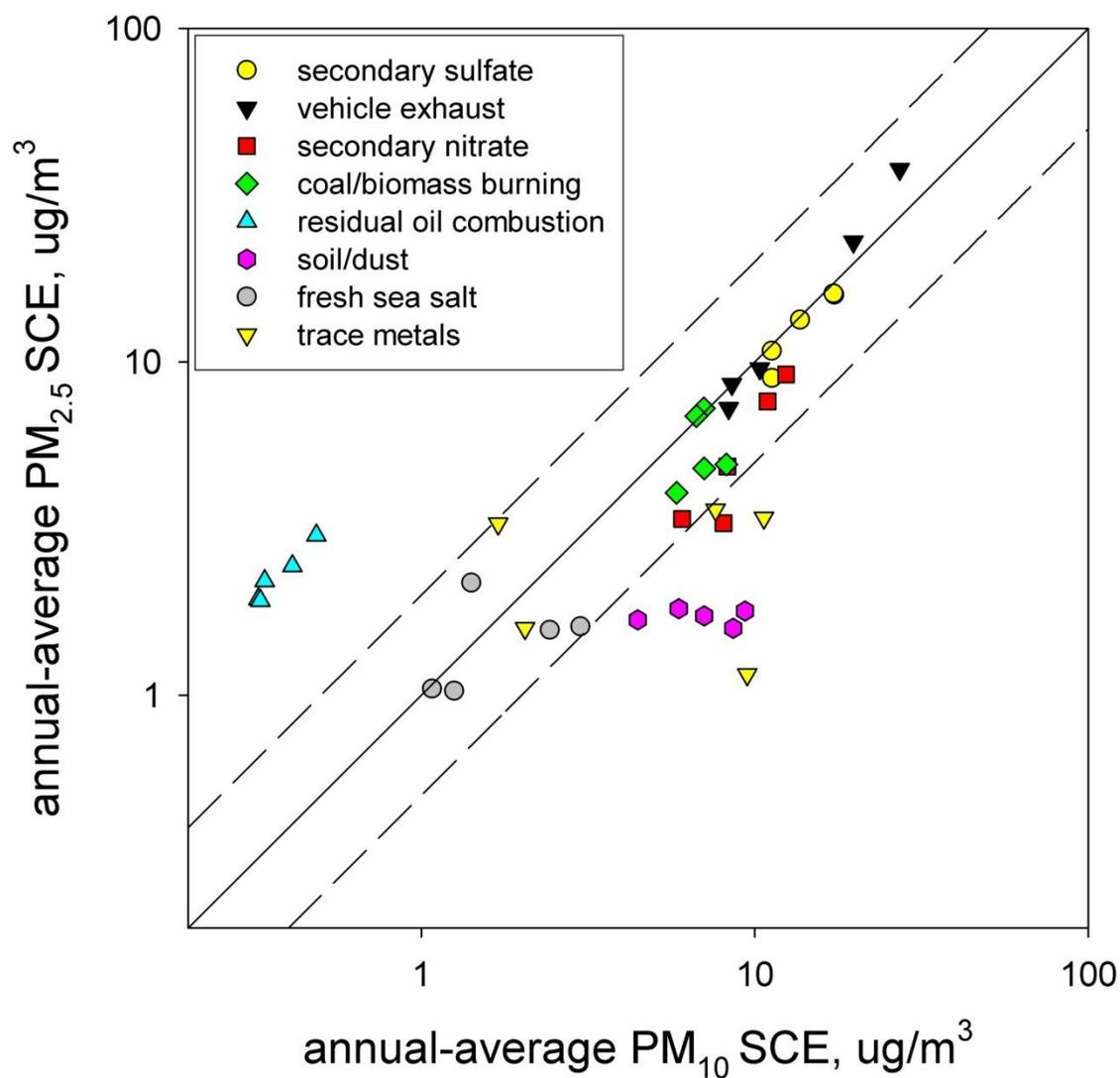
**Figure 4.9** Daily source contribution estimates for the source apportionment modeled Hong Kong PM<sub>2.5</sub> data using UNMIX and PMF. All concentration values are  $\mu\text{g}/\text{m}^3$ . The diagonal dashed line is the 1:1 line.



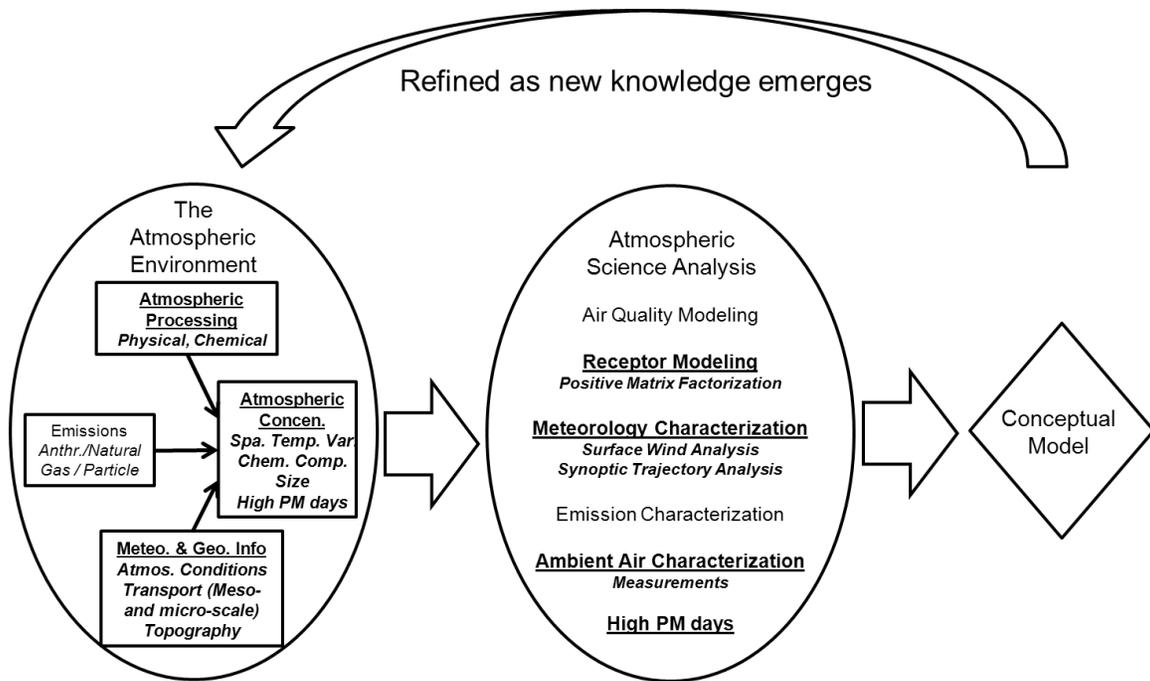
**Figure 4.10** Annual average source contribution estimates by site and year for each of the eight PMF-resolved factors.



**Figure 4.11** Annual average source contribution estimates by site and year for each of the eight PMF-resolved factors.

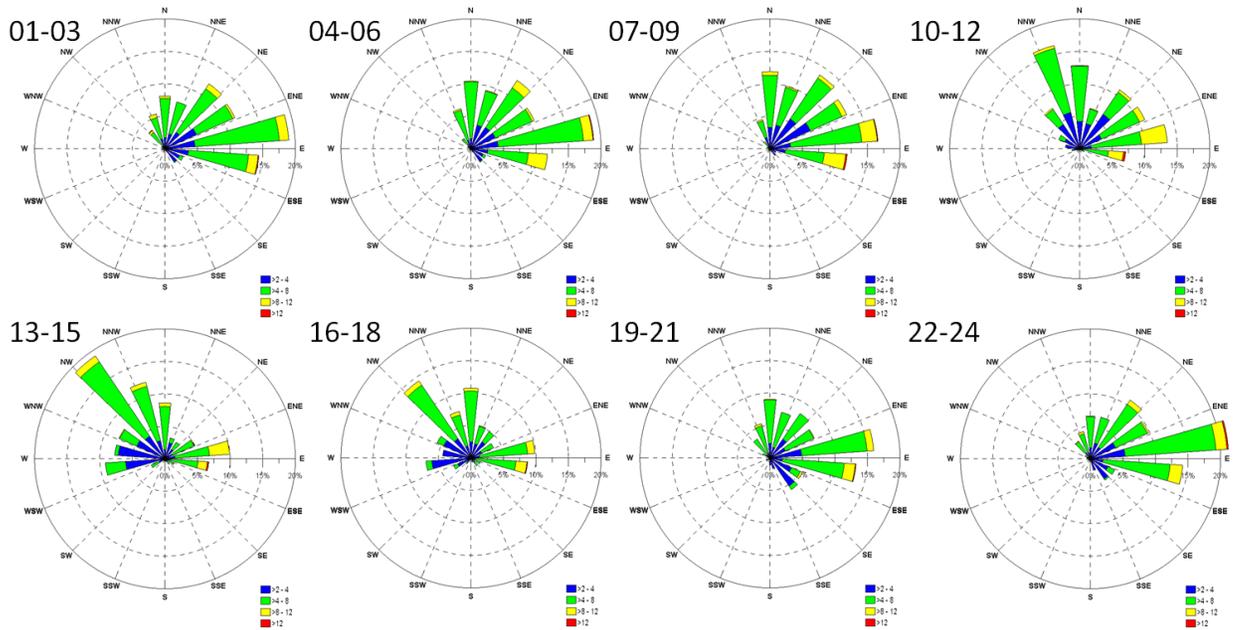


**Figure 4.12** Annual average  $PM_{2.5}$  and  $PM_{10}$  source contribution estimates (SCEs) for each of the eight PMF-resolved factors by site and year: TW (2001, 2005), YL (2005), and MK (2001, 2005). SCEs determined for PMF modeling with quartz filter gravimetric mass as the Total Variable. For TW and YL the  $PM_{2.5}$  and  $PM_{10}$  sample dates coincided and the averages were calculated using paired samples; for MK the sample dates did not coincide and the averages were calculated using all samples over the time period. The solid line is the 1:1 line; the dashed lines are 2:1 and 1:2 lines.

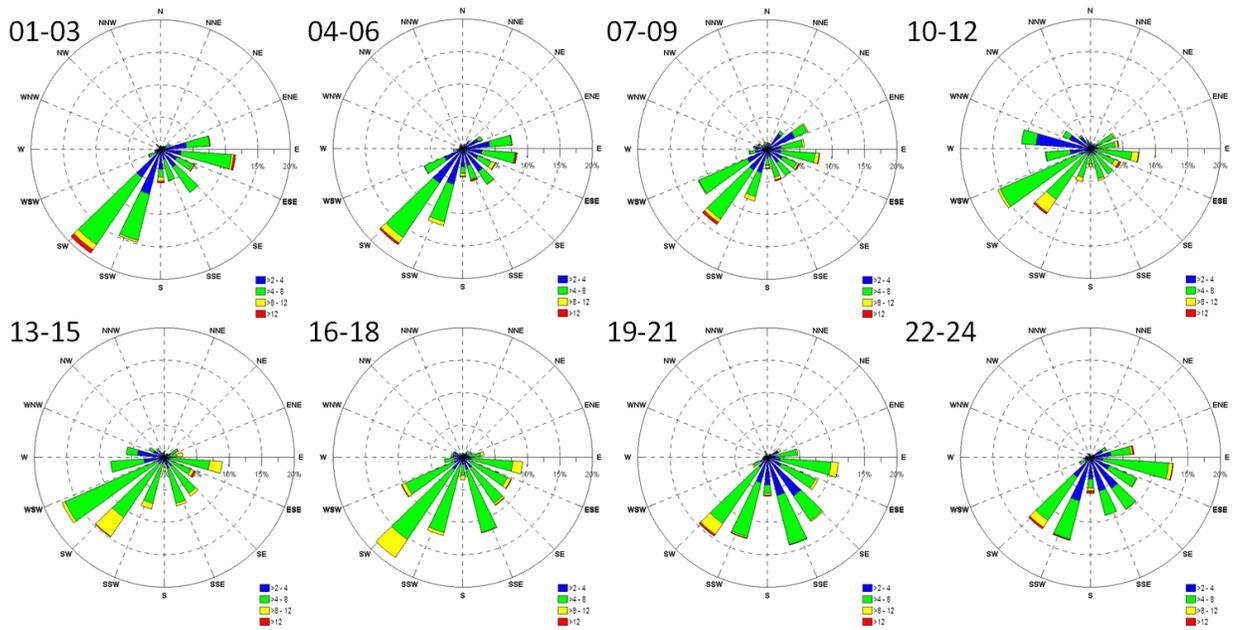


**Figure 5.1** Development of a conceptual model for Hong Kong. Elements in bold are applied in this Study

(a) JANUARY

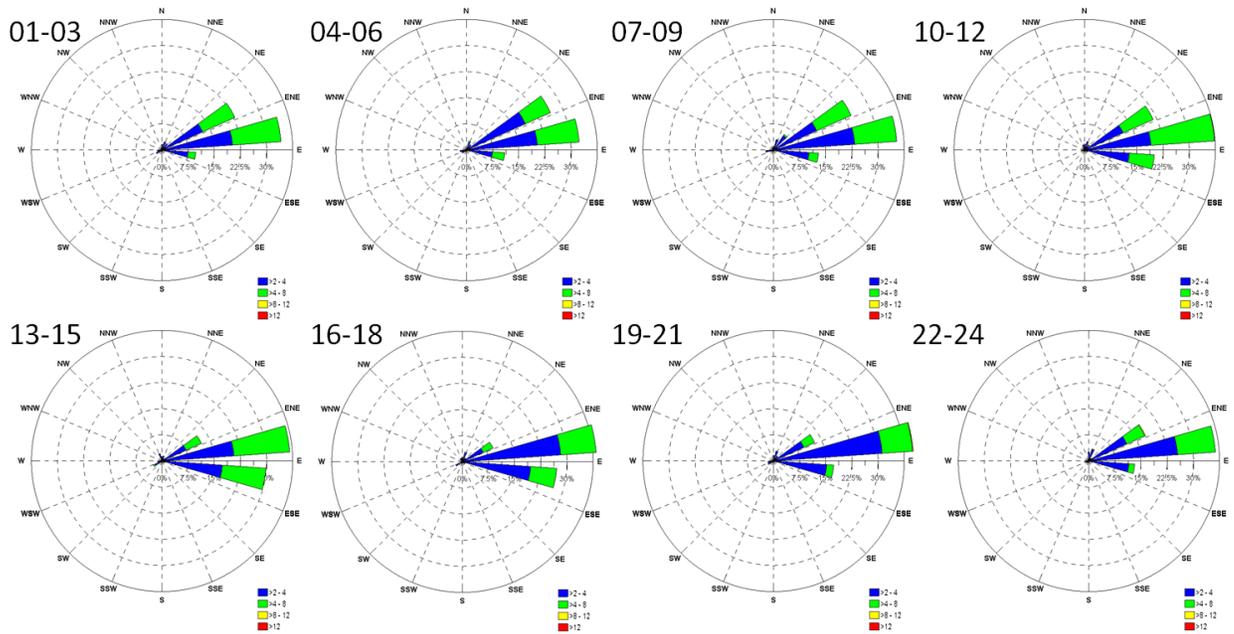


(b) JULY

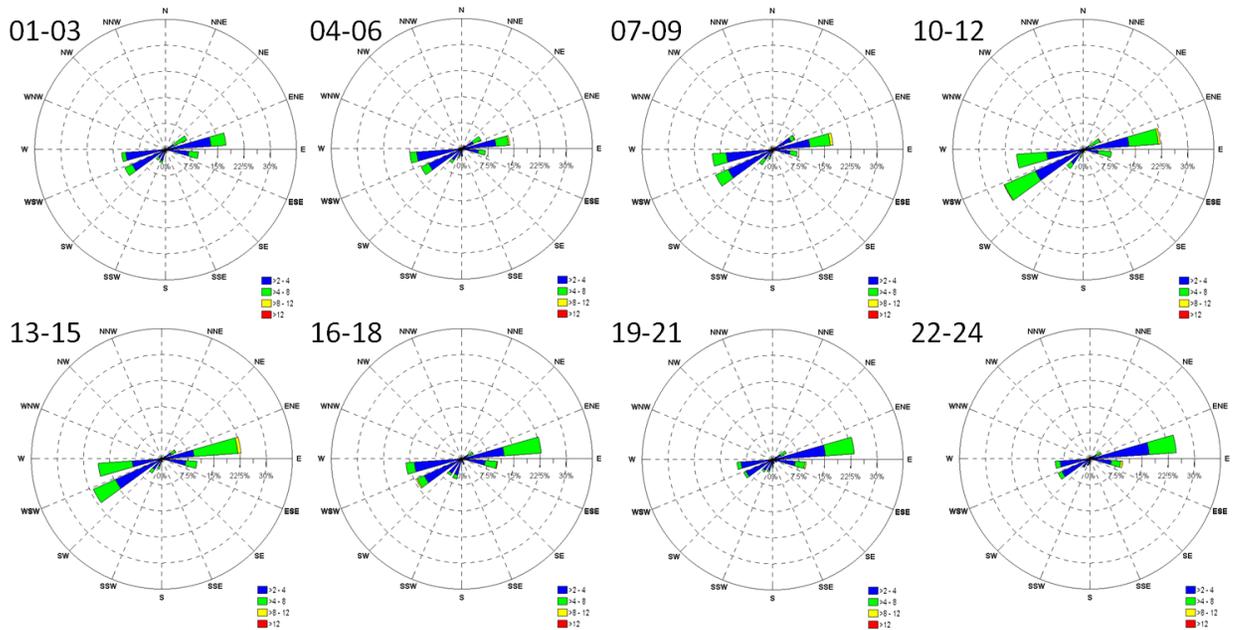


**Figure 5.2** January (a) and July (b) wind roses stratified by time of day for the Hong Kong International Airport, 1999-2008. Calm hours have been excluded.

(a) JANUARY



(b) JULY



**Figure 5.3** January (a) and July (b) wind roses stratified by time of day for the Hong Kong Observatory, 1999-2008. Calm hours have been excluded.

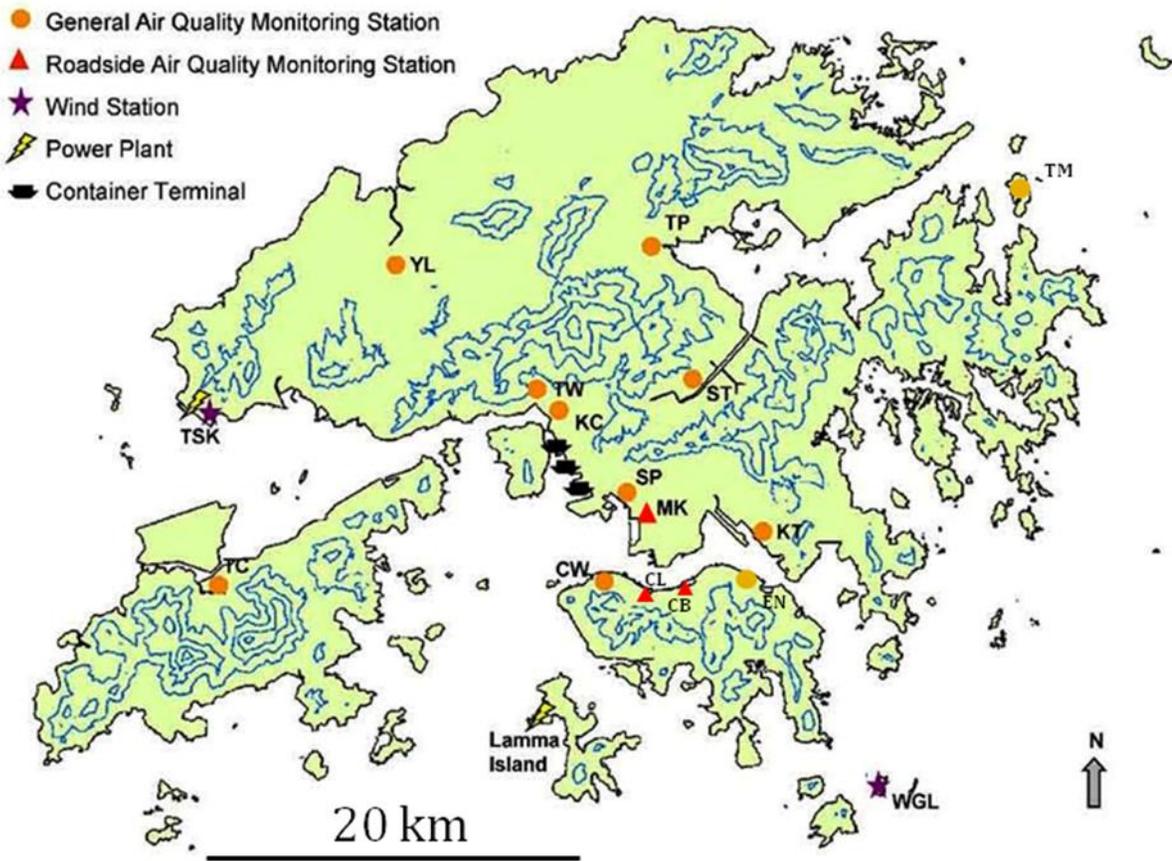
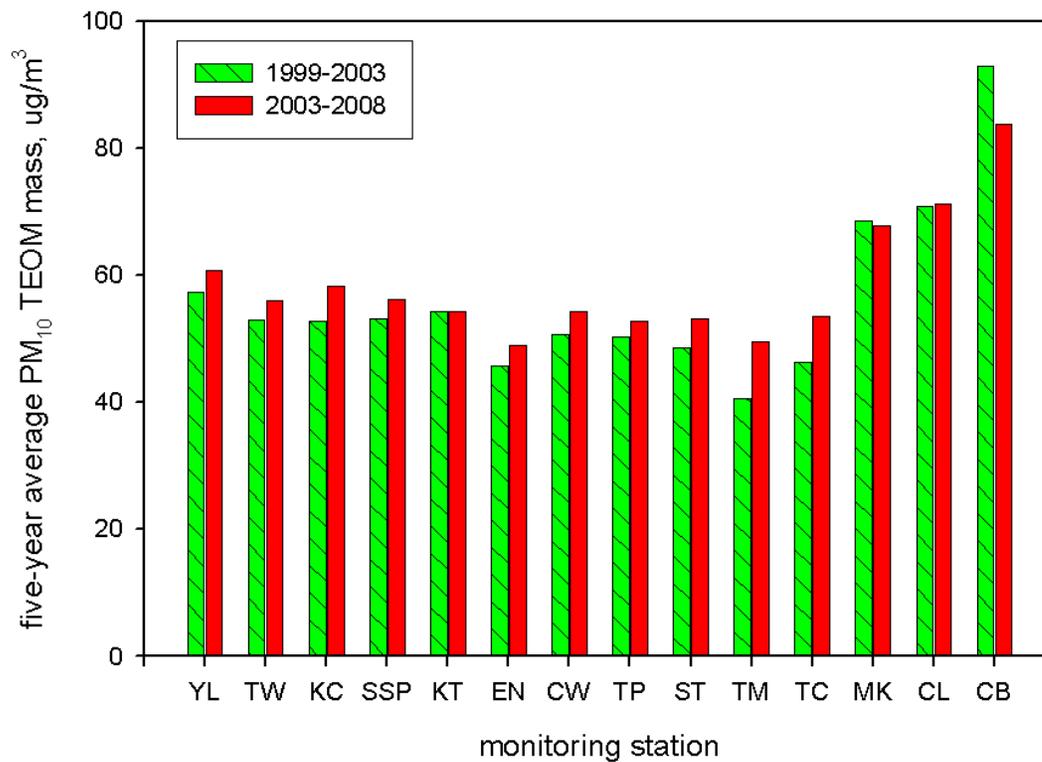
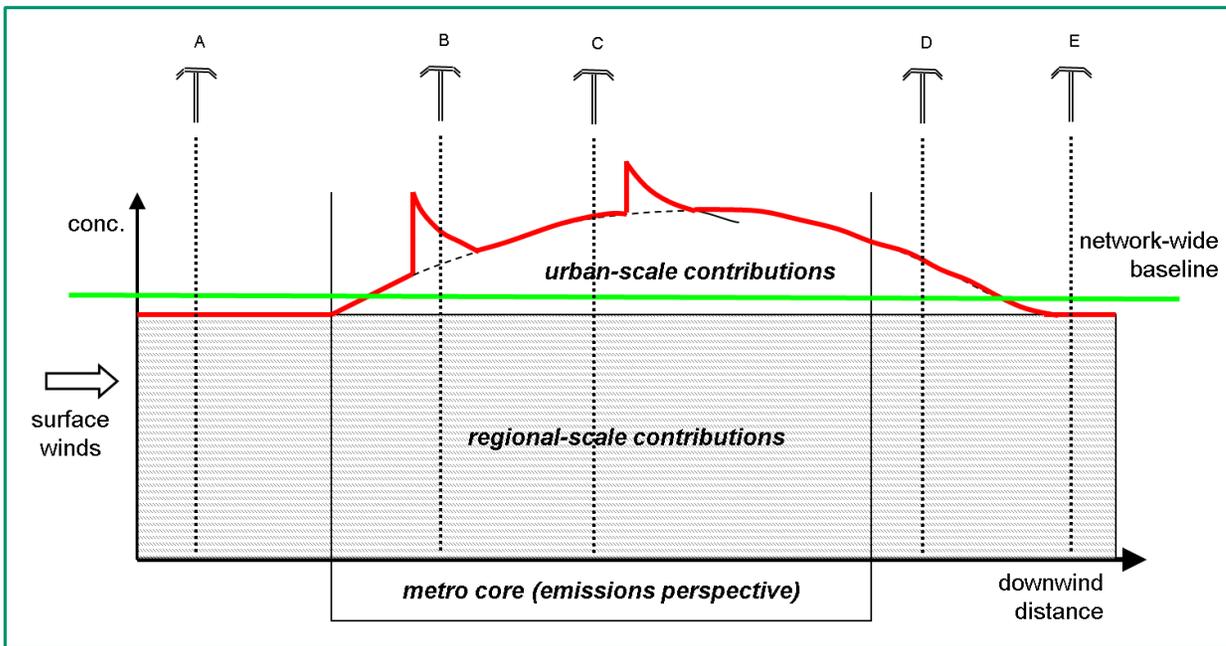
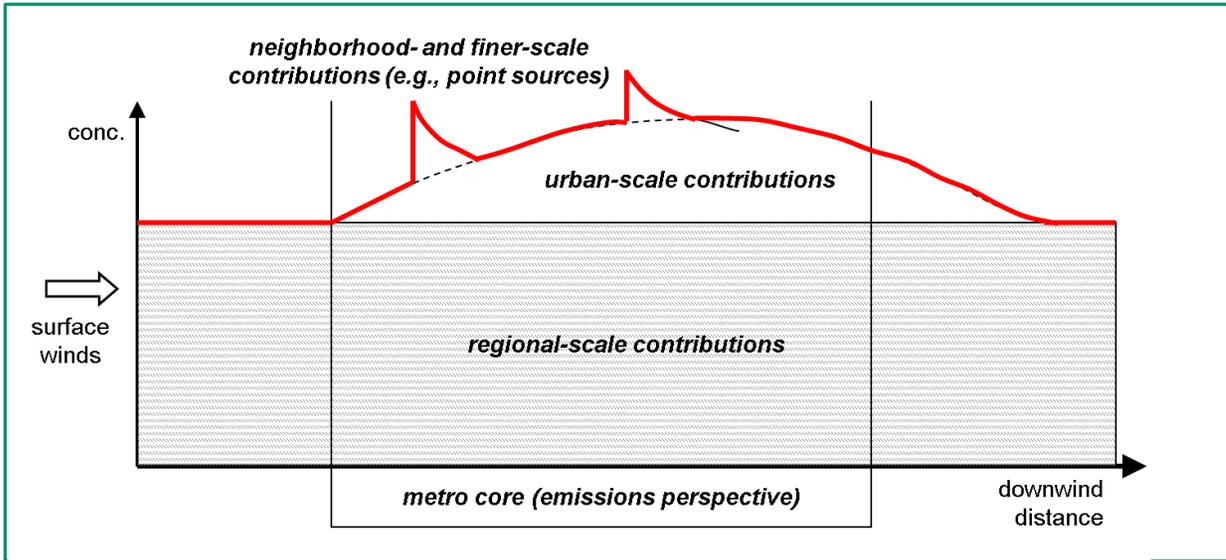


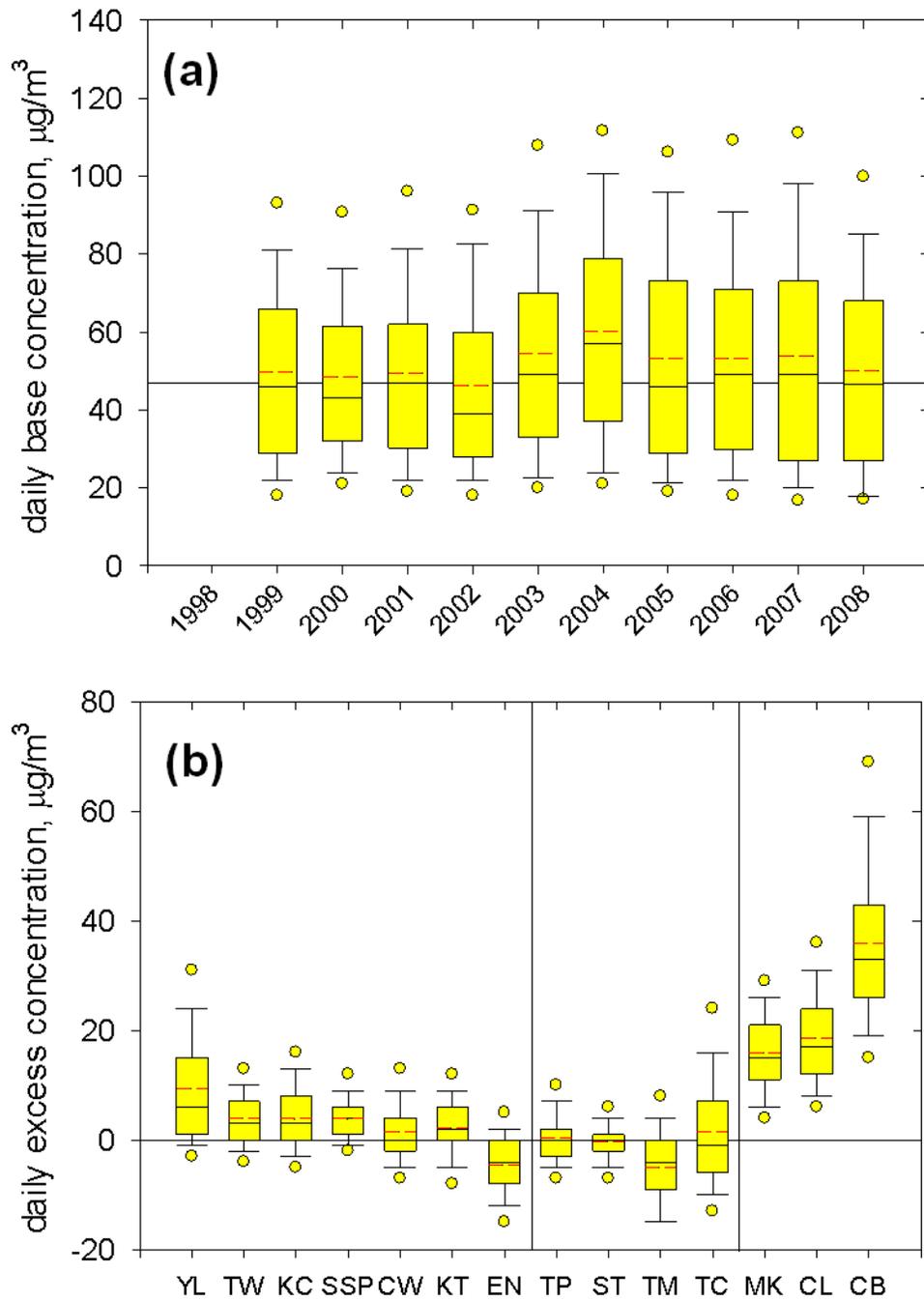
Figure 5.4 HKSAR and the Hong Kong EPD TEOM monitoring stations.



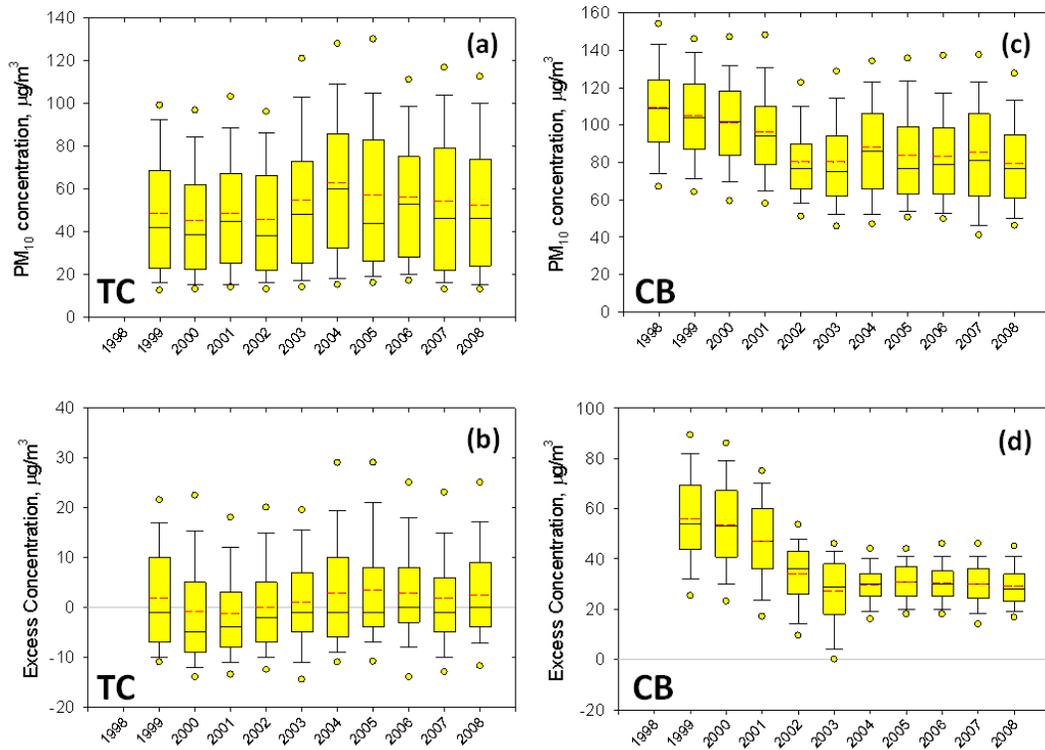
**Figure 5.5** Map of five-year average (2004-2008) PM<sub>10</sub> TEOM mass concentrations with circle area proportional to concentration (top) and five-year average concentrations for 1999-2003 and 2004-2008 (bottom).



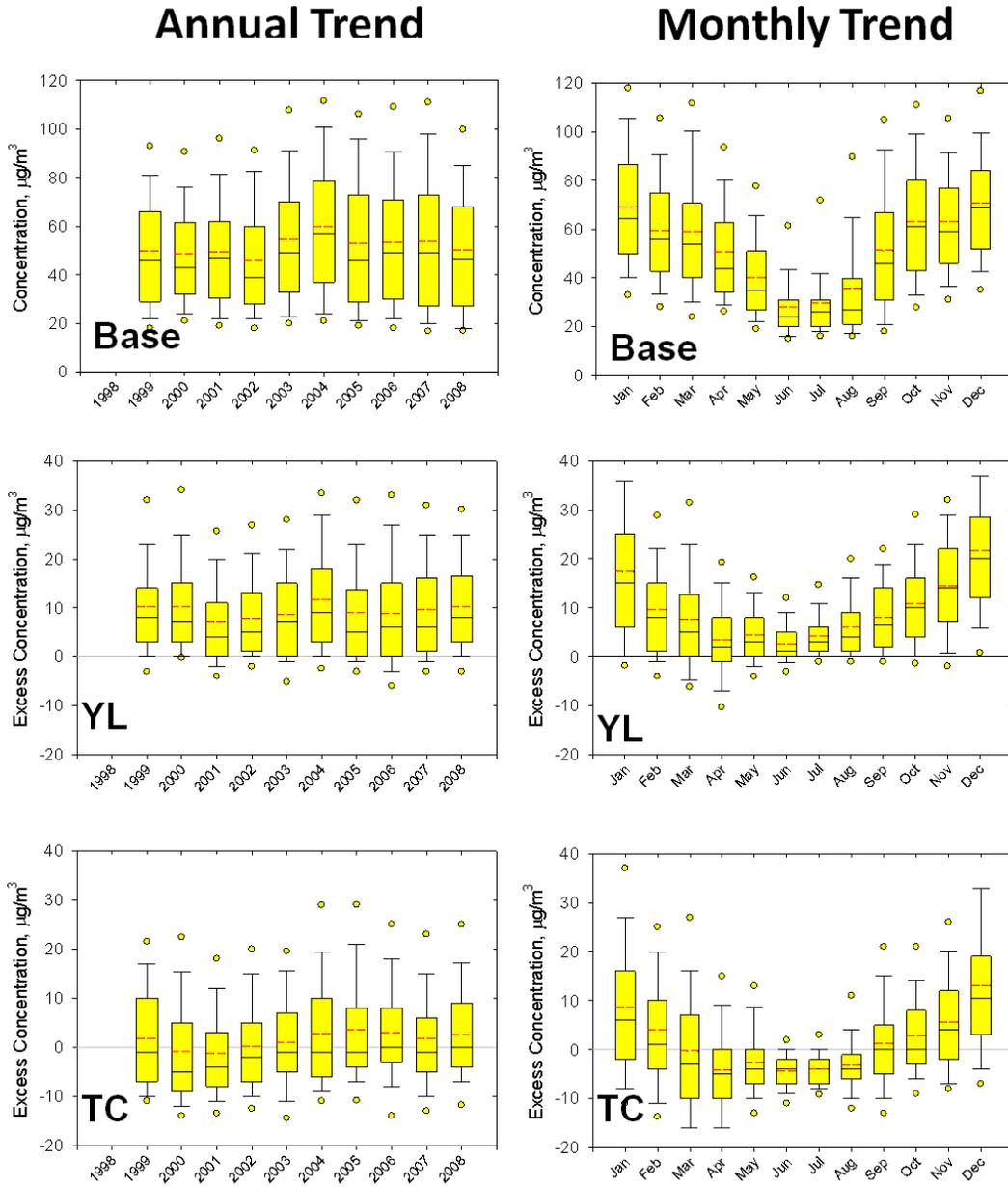
**Figure 5.6** The influence of emissions acting on various spatial scales to observed concentrations at monitoring sites. Figure adapted from Turner (2010).



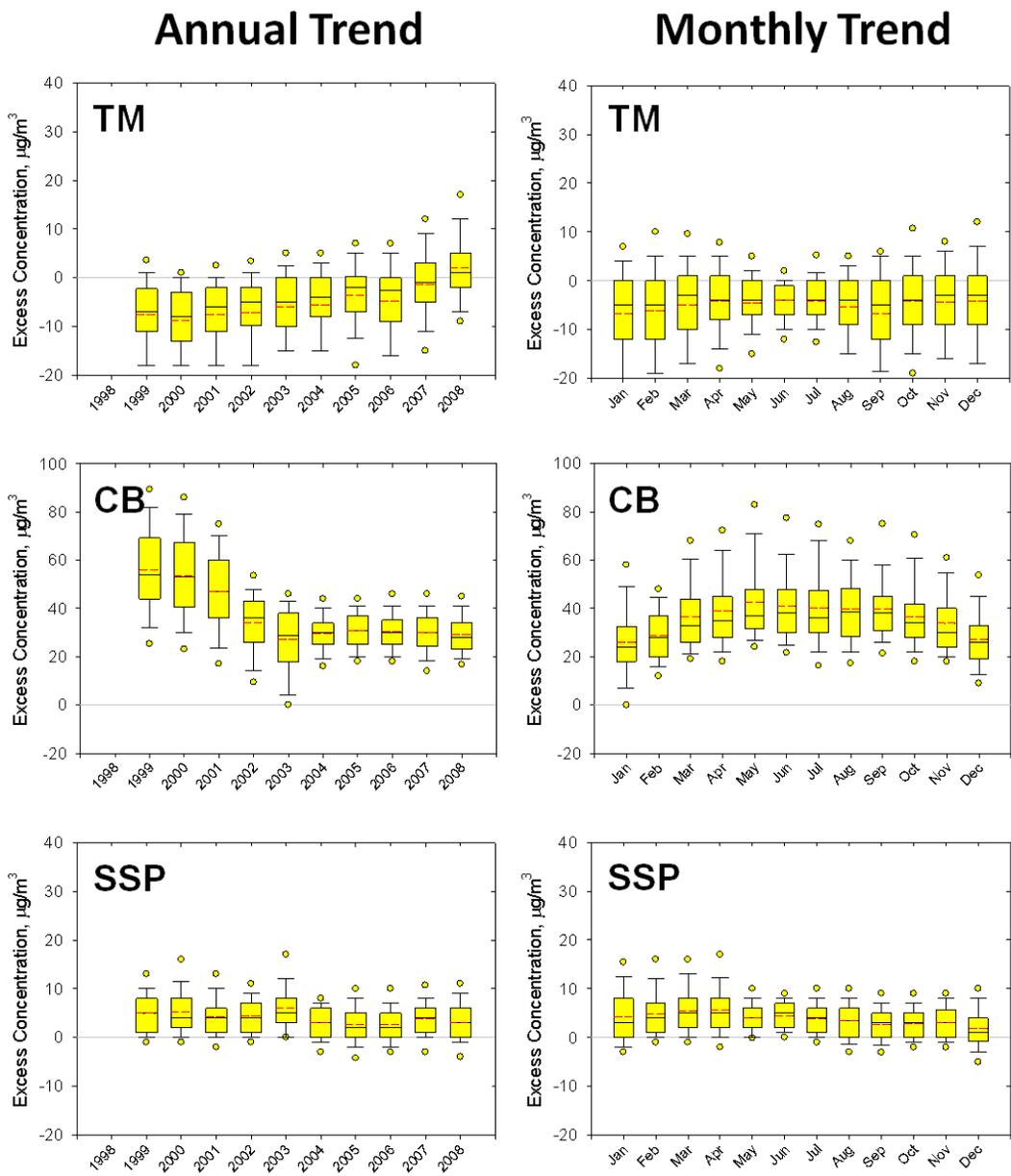
**Figure 5.7** (a) Annual distributions of the network-wide PM<sub>10</sub> TEOM mass base concentration; and (b) site-specific excess concentrations above the base concentration. In panel (a) the median based concentration for the entire data set is the black horizontal extending across the entire plot.



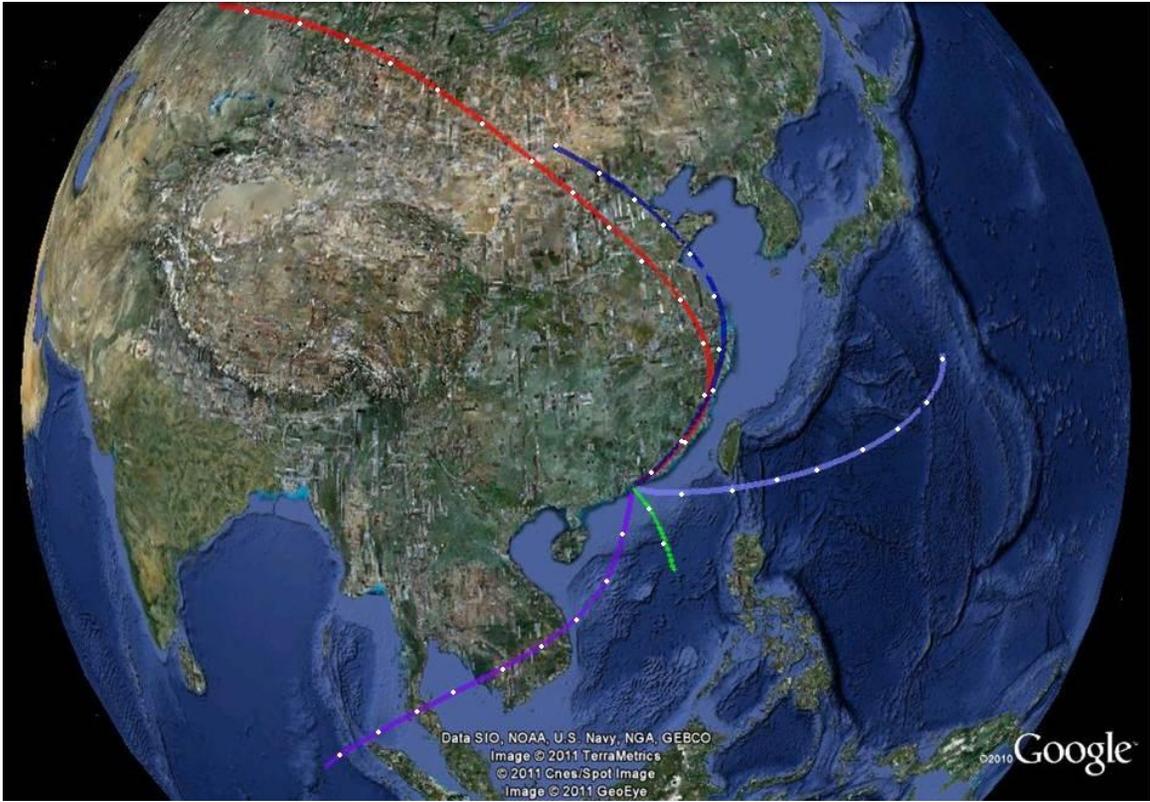
**Figure 5.8** Annual distributions of the PM<sub>10</sub> TEOM mass concentrations at the TC and CB stations. Figures (a) and (c) are the total PM<sub>10</sub> mass concentrations; figure (b) and (d) are the PM<sub>10</sub> excess mass concentrations calculated by subtracting out the day-specific network-wide base concentrations prior to constructing the distributions.



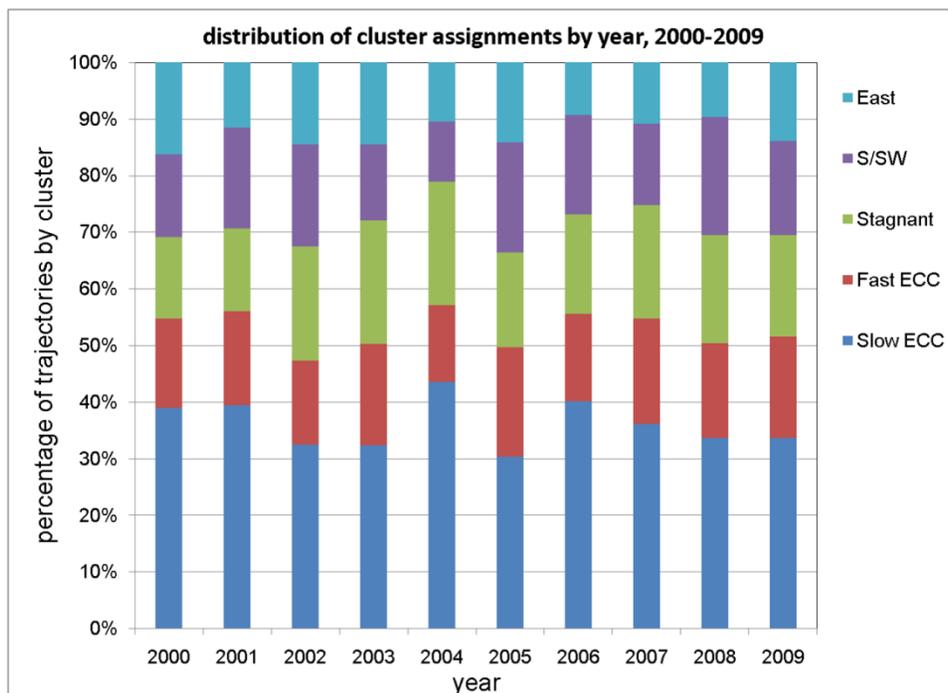
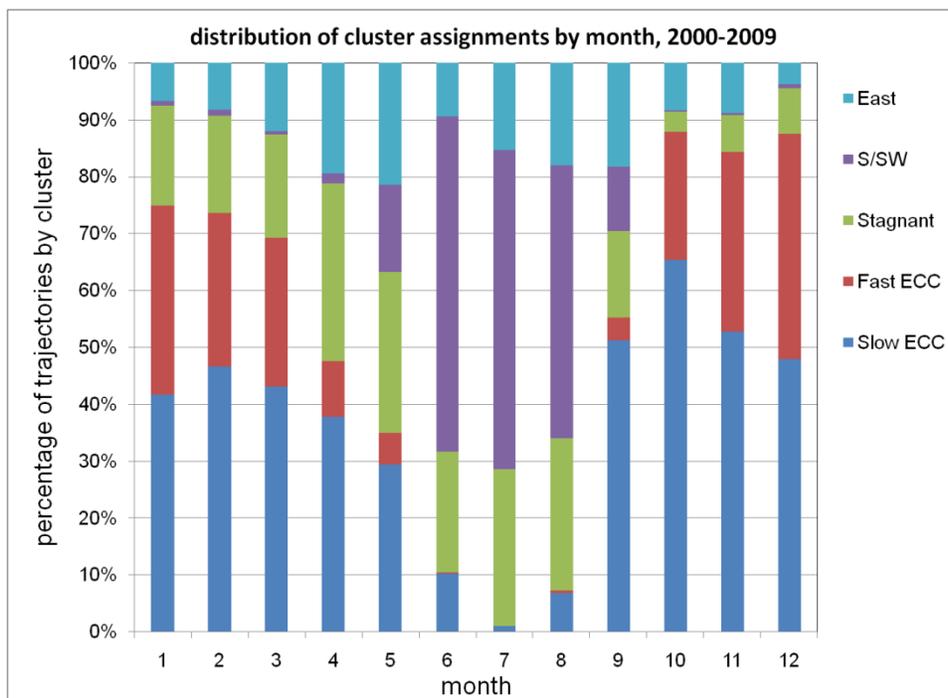
**Figure 5.9** Annual and monthly distributions of the  $\text{PM}_{10}$  TEOM mass base concentration (top panels) and excess mass concentrations at YL (middle panels) and TC (bottom panels).



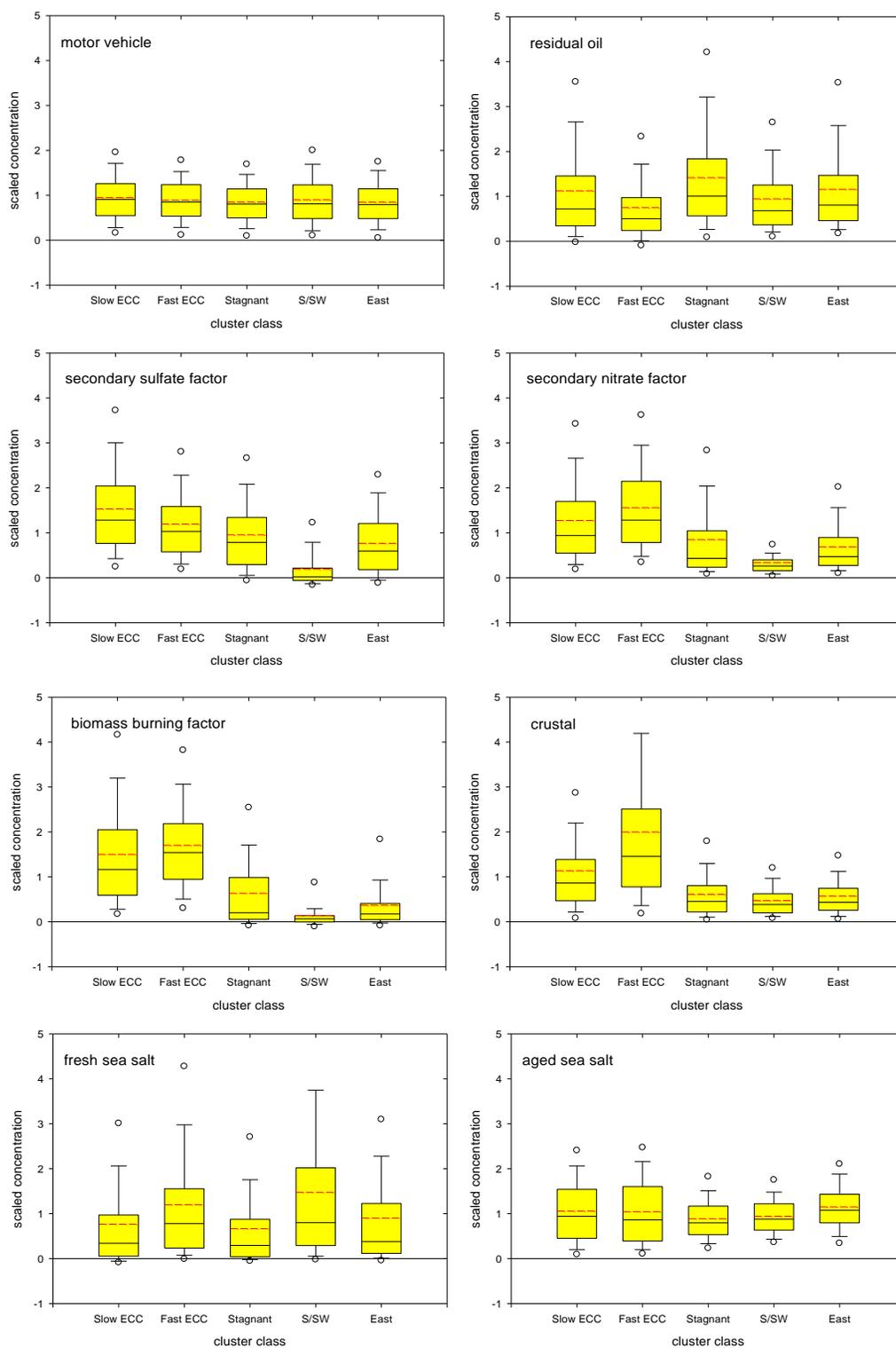
**Figure 5.9 (continued).** Annual and monthly distributions of the PM<sub>10</sub> TEOM excess mass concentrations at TM (top panels), CB (middle panels) and SSP (bottom panels).



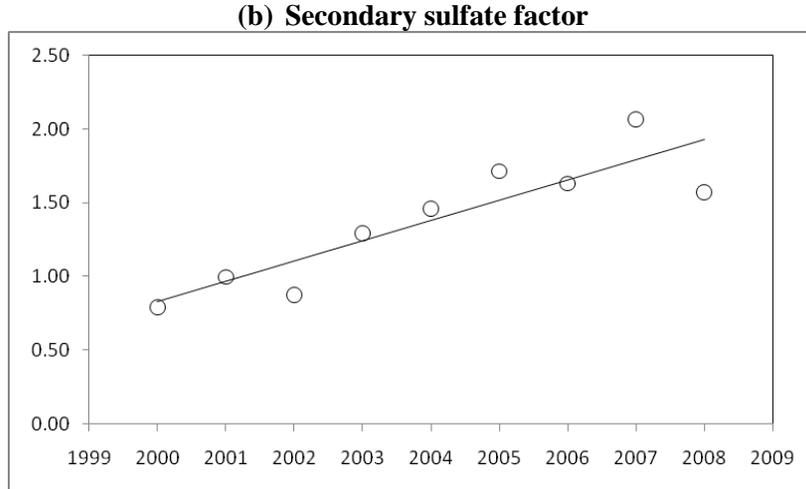
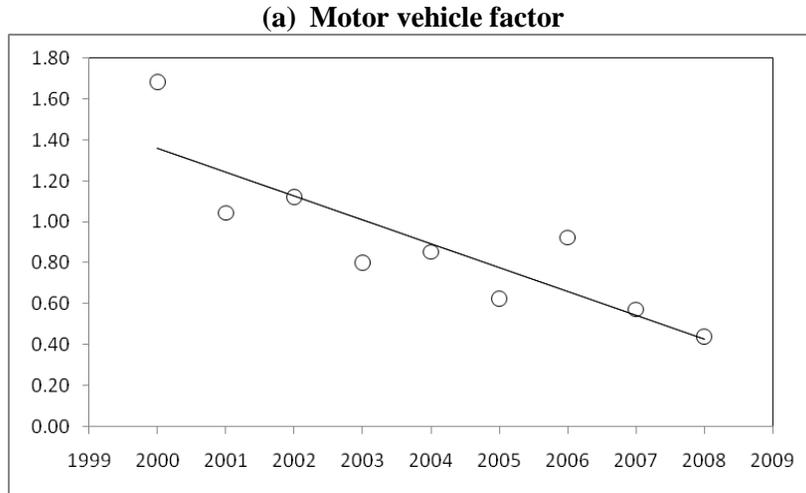
**Figure 5.10** Five air mass patterns resolved by clustering of seven-day air mass back trajectories. The centroid trajectory of each cluster is shown in this figure.



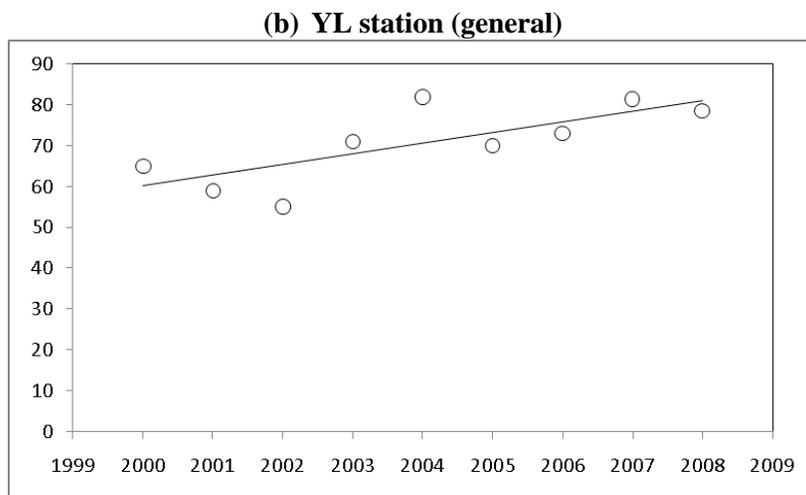
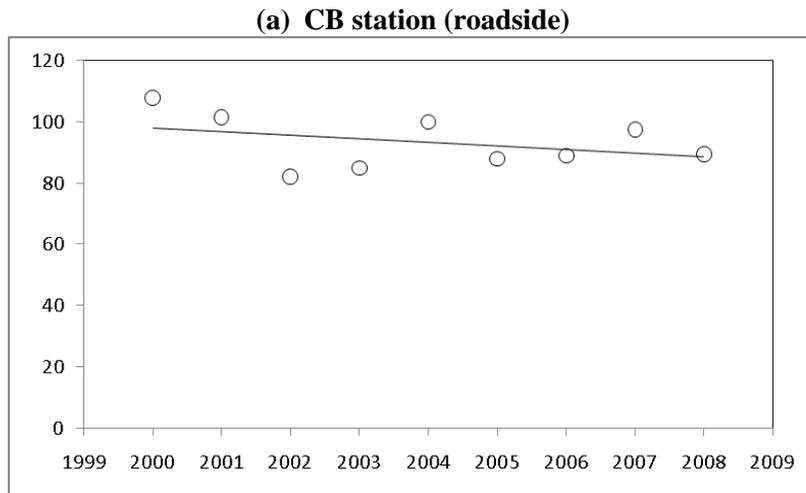
**Figure 5.11** Distributions of the air mass back trajectory cluster assignments by month (top) and year (bottom), 2000-2009. Each day included four air mass back trajectories.



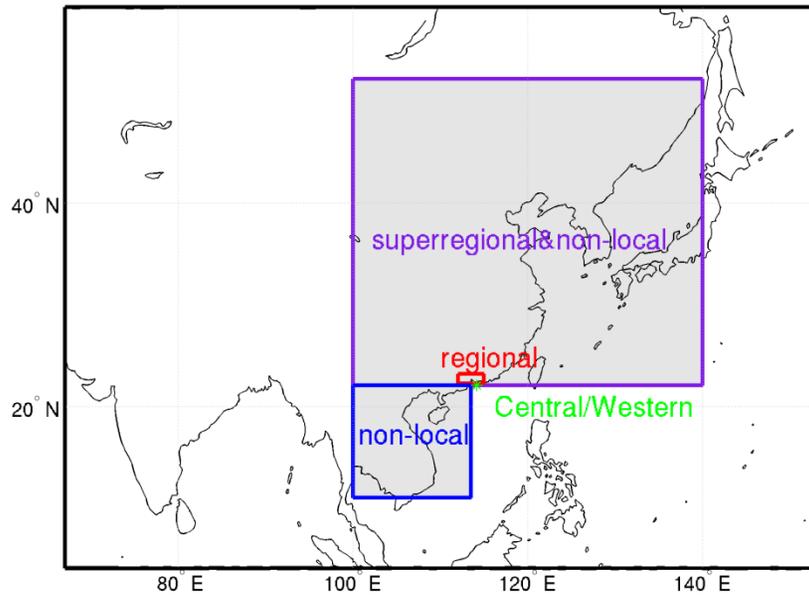
**Figure 5.12** Distribution of PMF-modeled PM<sub>10</sub> scaled source contributions estimates corresponding to each synoptic class pattern. The interior black solid line is the median, the dashed red line is the arithmetic mean, and the circles are 5<sup>th</sup> and 95<sup>th</sup> percentiles.



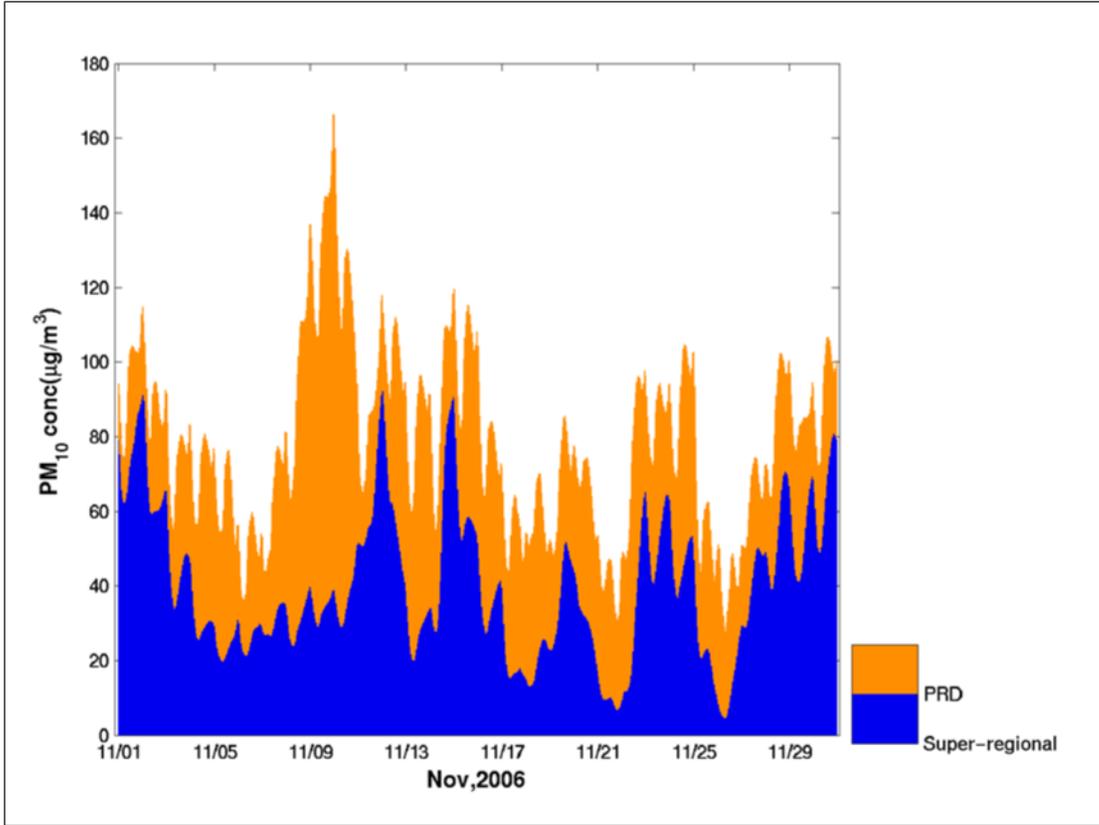
**Figure 5.13**  $\text{PM}_{10}$  annual median SCEs, after compositing over the general stations, for PMF-resolved vehicle exhaust factor (top) and secondary sulfate factor (bottom) including only those days with a transport pattern of slow moving air masses from mainland China (Slow ECC). Y-axis values are annual median SCEs reported as  $\mu\text{g}/\text{m}^3$ .



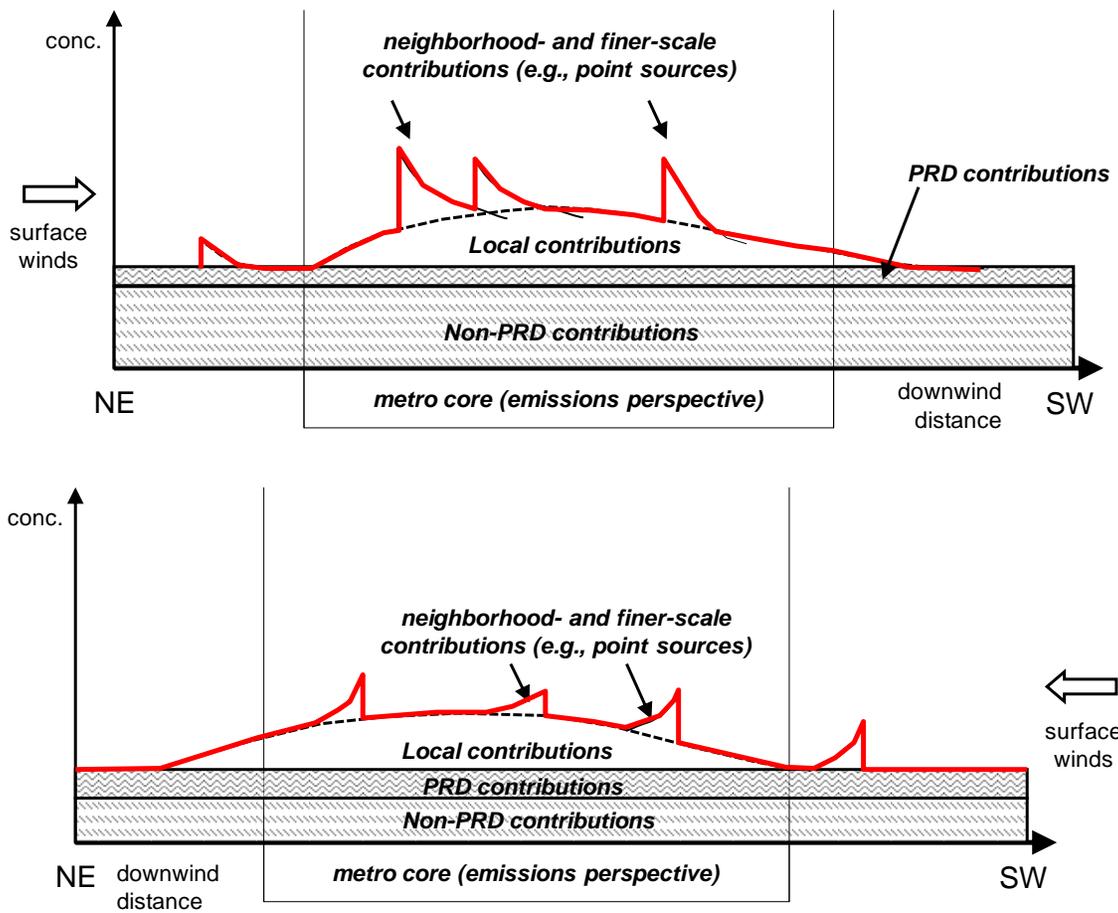
**Figure 5.14** PM<sub>10</sub> annual median mass concentrations for the CB roadside station (top) and YL general station (bottom) including only those days with a transport pattern of slow moving air masses from mainland China (Slow ECC). Y-axis values are annual median PM<sub>10</sub> mass concentrations reported as µg/m<sup>3</sup>.



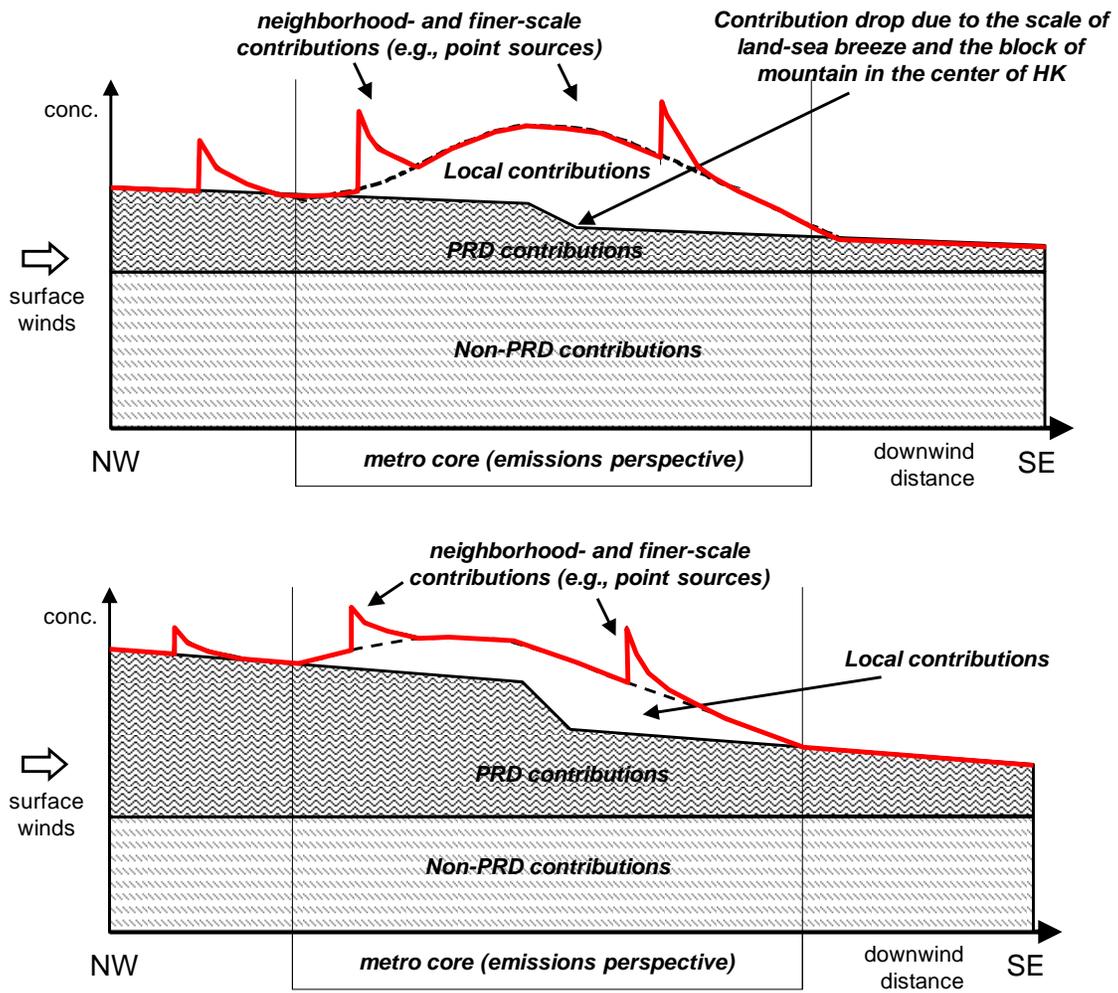
**Figure 5.15** Map of areas designated to define PRD, NON-PRD and LOCAL back trajectory classes



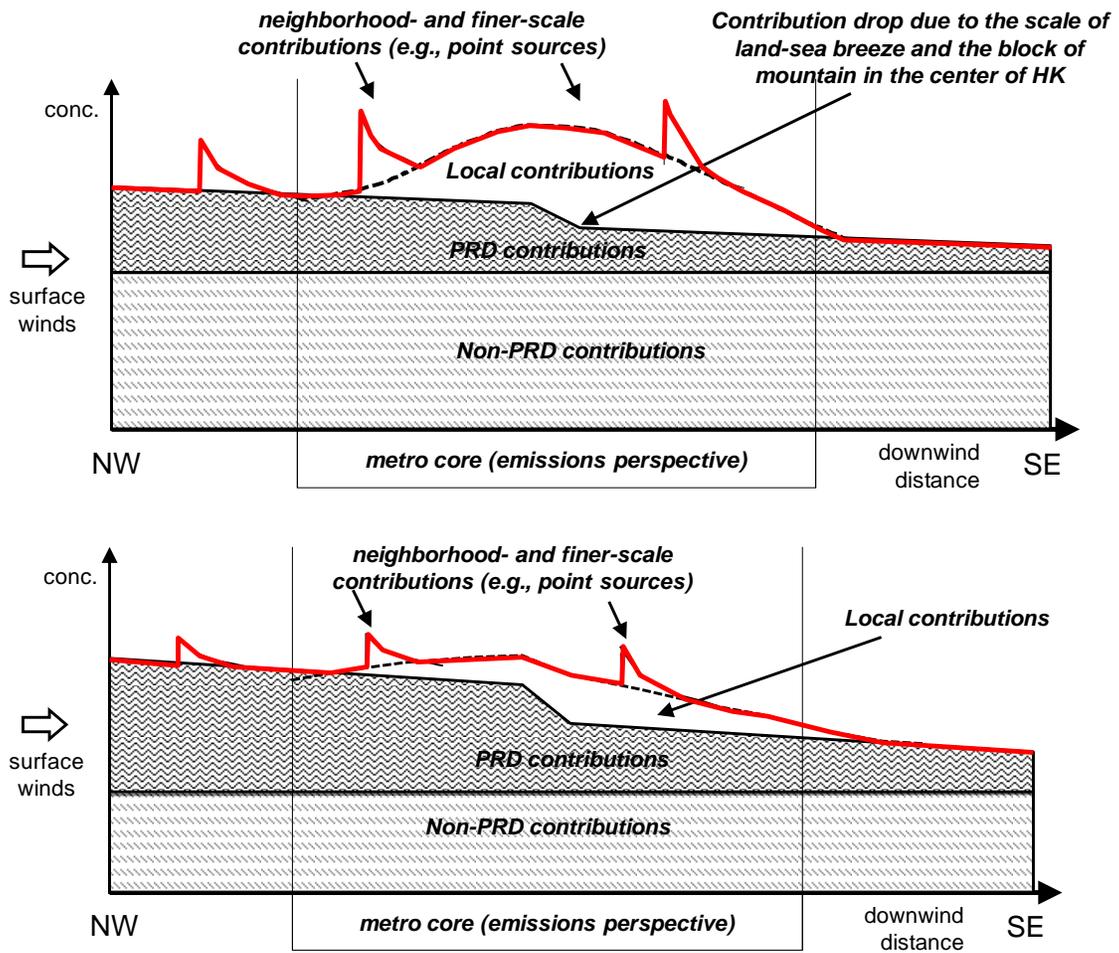
**Figure 5.16** Absolute mass contributions of PRD and non-PRD sources to the PM<sub>10</sub> in Hong Kong in November 2006, as modeled by the PSAT module in CAMx



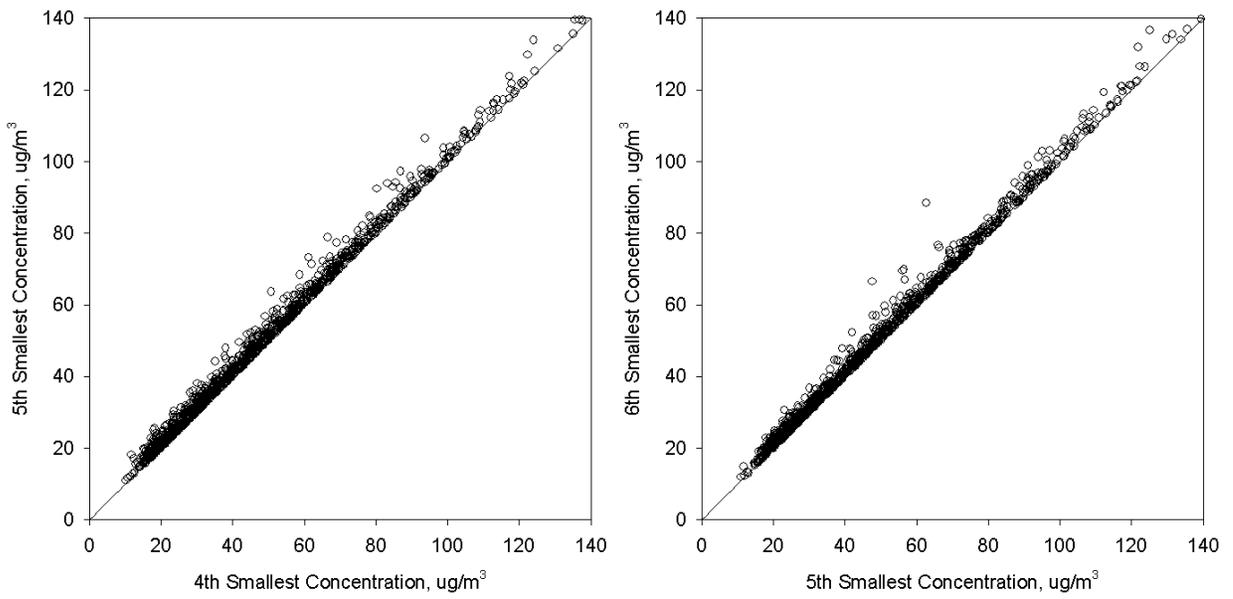
**Figure 5.17** Conceptual model of PM in Hong Kong during general conditions in wintertime (above) and summertime (below)



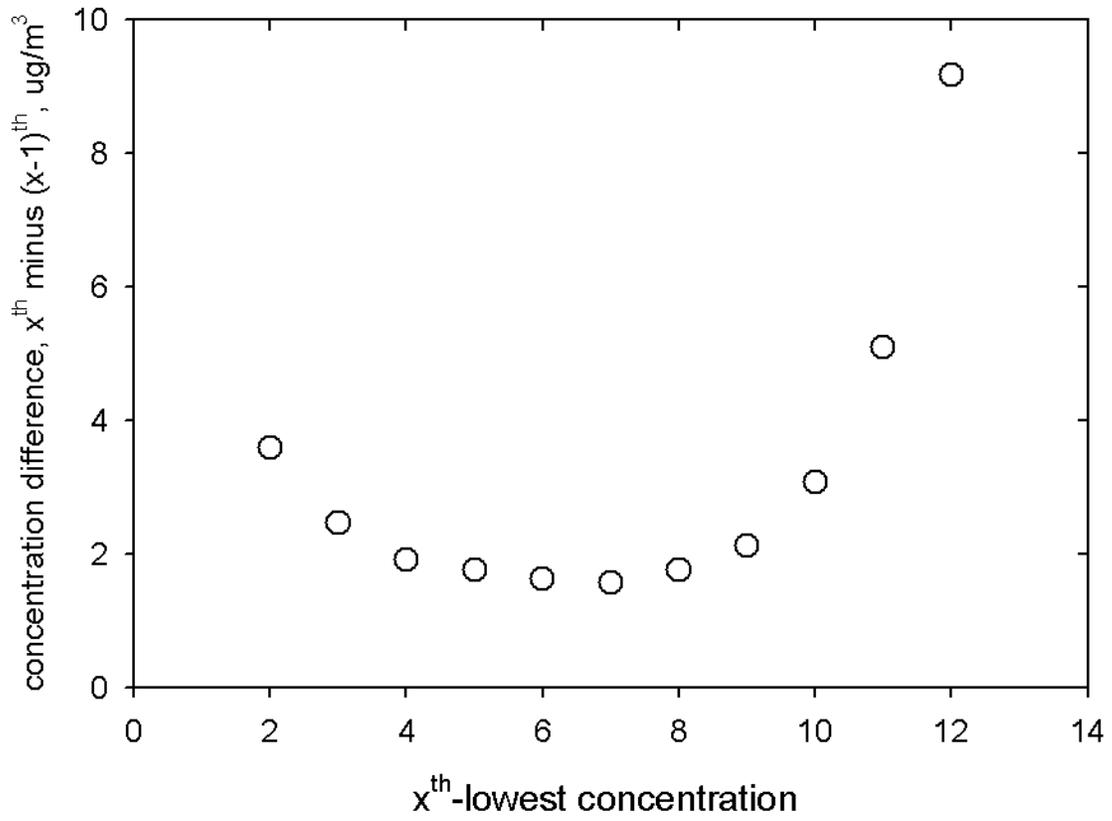
**Figure 5.18** Conceptual model of PM in Hong Kong during high PM hours [Scenario 1] in wintertime (above) and summertime (below)



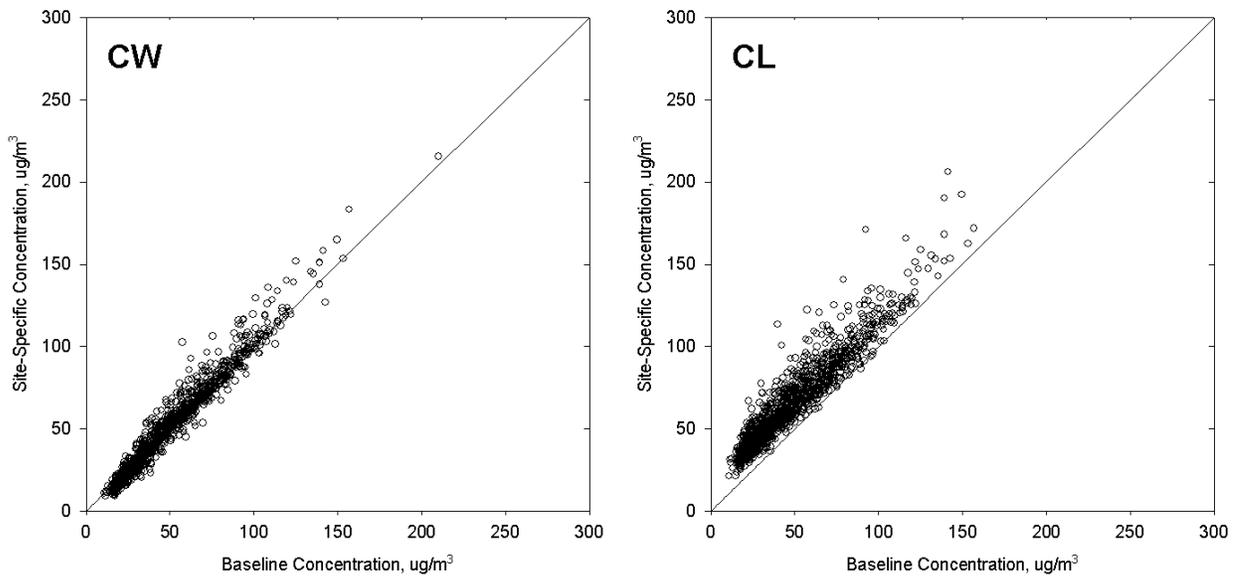
**Figure 5.19** Conceptual model of PM in Hong Kong during high PM hours [Scenario 2] in wintertime (above) and summertime (below)



**Figure C-1** Relationships between the 4<sup>th</sup>-, 5<sup>th</sup>-, and 6<sup>th</sup>-minimum of the daily-average  $\text{PM}_{10}$  mass concentrations on valid network days.



**Figure C-2** Change in the grand average base concentration upon increased the metric from the (x-1)<sup>th</sup> lowest value to the x<sup>th</sup>-lowest value.



**Figure C-3** 24-hour average PM<sub>10</sub> TEOM mass concentrations at CW (general station) and CL (roadside station) versus the daily-average base concentration defined as the 5<sup>th</sup>-minimum value across the network.

**Table 2.1** Additive and multiplicative error terms derived using the SSP collocated data set using three estimation approaches. Coefficient  $a_j$  has units  $\mu\text{g}/\text{m}^3$ .

Species	Unweighted		Weighted		Refined Weighted (Outliers Removed)	
	$a_j$	$b_j$	$a_j$	$b_j$	$a_j$	$b_j$
Al	0.0063	0.095	0.0076	0.022	0.0046	0.071
As	0.0003	0.074	0.0002	0.003	0.0001	0.050
Ca	0.0150	0.092	0.0349	0.002	0.0217	0.043
Cd	0.0001	0.066	0.0001	-0.006	0.00002	0.058
Cl <sup>-</sup>	0.0092	0.101	0.0125	0.017	0.0141	0.047
Fe	0.0075	0.054	0.0132	-0.003	0.0044	0.030
K <sup>+</sup>	0.1048	0.084	0.0568	-0.037	0.0118	0.028
Mg	0.0075	0.071	0.0076	0.001	0.0049	0.020
Mn	0.0031	0.253	0.0013	-0.025	0.0002	0.029
Na <sup>+</sup>	0.0997	0.144	0.0452	0.014	0.0264	0.018
NH <sub>4</sub> <sup>+</sup>	0.0184	0.037	0.0219	0.014	0.0169	0.028
Ni	0.0010	0.100	0.0007	-0.018	0.0003	0.050
NO <sub>3</sub> <sup>-</sup>	0.0221	0.057	0.0242	0.034	0.0429	0.026
Pb	0.0030	0.039	0.0020	0.005	0.0002	0.035
SO <sub>4</sub> <sup>2-</sup>	0.0499	0.040	0.1261	0.009	0.0597	0.025
V	0.0030	0.113	0.0012	-0.018	0.00001	0.047
Zn	0.0075	0.072	0.0043	0.002	0.0016	0.032
EC	0.0833	0.052	0.2087	-0.010	0.1148	0.023
OC	0.0833	0.055	0.2634	-0.003	0.1117	0.032

**Table 2.2** Data below the MDL, data pairs removed for the collocate precision calculations to estimate error structures, reported MDL values, effective MDL values using three times the  $a_j$  values from the “Refined Weighted” additive error term estimates, and percentage of samples below MDL in the analysis dataset. Analysis using the SSP collocated data.

Species	Count (N=517)		MDL ( $\mu\text{g}/\text{m}^3$ )		
	<MDL <sup>(1)</sup>	Removed	Reported	Effective	% of samples < MDL in the dataset
Al	1	9	0.0190 <sup>(2)</sup>	0.0137	1%
As	164	7	0.0009 <sup>(3)</sup>	0.0002	25%
Ca	0	2	0.0449 <sup>(4)</sup>	0.0651	0%
Cd	182	3	0.0004 <sup>(3)</sup>	0.0001	28%
Cl <sup>-</sup>	18	4	0.0276 <sup>(3)</sup>	0.0424	0%
Fe	0	7	0.0225 <sup>(4)</sup>	0.0133	0%
K <sup>+</sup>	234	4	0.3143 <sup>(3)</sup>	0.0354	0%
Mg	2	8	0.0225 <sup>(4)</sup>	0.0146	0%
Mn	188	3	0.0092 <sup>(3)</sup>	0.0007	27%
Na <sup>+</sup>	40	27	0.2991 <sup>(4)</sup>	0.0792	0%
NH <sub>4</sub> <sup>+</sup>	7	3	0.0553 <sup>(3)</sup>	0.0508	0%
Ni	224	2	0.0031 <sup>(3)</sup>	0.0008	36%
NO <sub>3</sub> <sup>-</sup>	0	6	0.0662 <sup>(4)</sup>	0.1286	0%
Pb	144	15	0.0090 <sup>(3)</sup>	0.0006	20%
SO <sub>4</sub> <sup>2-</sup>	0	13	0.1496 <sup>(4)</sup>	0.1792	0%
V	310	4	0.0090 <sup>(3)</sup>	0.00002	48%
Zn	39	8	0.0225 <sup>(3)</sup>	0.0048	6%
EC	0	9	0.2500 <sup>(4)</sup>	0.3443	0%
OC	0	8	0.2500 <sup>(4)</sup>	0.3352	0%

- (1) Data pairs with one-or-both concentration values below the reported MDL.
- (2) MDL taken from the table of reported MDL values of Interim Report 1
- (3) MDL estimated from collocated SSP dataset, average of “<XXX” values
- (4) Average of the Laboratory MDL range reported by HKEPD

**Table 2.3** Collocated data metrics for the 517 TW samples: number of sample pairs with one or more values below the laboratory-reported MDL; and number of sample pairs removed from the collocated precision calculations.

Species	# < MDL	# Removed
Al	3	10
As	141	0
Ca	1	4
Cd	153	4
Cl <sup>-</sup>	31	1
Fe	0	14
K <sup>+</sup>	219	2
Mg	1	8
Mn	166	35
Na <sup>+</sup>	34	14
NH <sub>4</sub> <sup>+</sup>	3	1
Ni	184	3
NO <sub>3</sub> <sup>-</sup>	0	1
Pb	115	1
SO <sub>4</sub> <sup>2-</sup>	0	4
V	241	0
Zn	18	2
EC	0	18
OM	0	16

**Table 2.4** Characteristics of the collocated PM<sub>2.5</sub> data set used to estimate error structures. N Rejected = number of data pairs excluded from the error structure coefficient determinations. Error structure coefficients (*a*,*b*) are described in the text.

Species	Analysis Method	MDL (mg/m <sup>3</sup> )	LQL (mg/m <sup>3</sup> )	N	N < MDL	% < MDL	N < LQL	% < LQL	N Rejected	Polissar Method		Weighted Regress.	
										<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>
Cl <sup>-</sup>	IC	0.0623	0.0844	40	17	43%	19	48%	3	0.0208	0.0601	0.0070	0.0446
NO <sub>3</sub> <sup>-</sup>	IC	0.0623	0.1074	40	0	0%	0	0%	0	0.0208	0.0930	0.0380	0.0566
SO <sub>4</sub> <sup>2-</sup>	IC	0.0623	1.4417	40	0	0%	0	0%	8	0.0208	0.0223	0.3993	-0.0052
NH <sub>4</sub> <sup>+</sup>	AC	0.0623	0.0415	40	1	3%	1	3%	3	0.0208	0.0331	0.0493	0.0171
Na	XRF	0.1558	0.1824	47	1	2%	1	2%	1	0.0519	0.1746	0.0272	0.1402
Mg	XRF	0.0471	0.0406	47	44	74%	40	85%	4	0.0157	0.3377	0.0036	0.3119
Al	XRF	0.0186	0.0262	47	9	7%	14	30%	0	0.0062	0.1316	0.0055	0.0859
Si	XRF	0.0150	0.0463	47	3	3%	8	17%	1	0.0050	0.0752	0.0013	0.0673
Cl	XRF	0.0020	0.0115	47	4	5%	11	23%	1	0.0007	0.2152	0.0012	0.1255
K	XRF	0.0019	0.0049	47	0	0%	0	0%	0	0.0006	0.0325	0.0005	0.0295
Ca	XRF	0.0030	0.0139	47	0	0%	0	0%	3	0.0010	0.0731	0.0015	0.0470
Ti	XRF	0.0014	0.0028	47	7	7%	12	26%	0	0.0005	0.1506	0.0005	0.1123
V	XRF	0.0003	0.0005	47	0	0%	0	0%	0	0.0001	0.0382	0.0005	0.0201
Cr	XRF	0.0016	0.0028	47	38	66%	43	91%	4	0.0005	0.2778	0.0003	0.0720
Mn	XRF	0.0035	0.0032	47	10	19%	10	21%	4	0.0012	0.1033	0.0010	0.0365
Fe	XRF	0.0032	0.0126	47	1	1%	2	4%	1	0.0011	0.0523	0.0015	0.0362
Ni	XRF	0.0005	0.0007	47	1	2%	1	2%	0	0.0002	0.1386	0.0005	0.0265
Cu	XRF	0.0018	0.0025	47	8	9%	9	19%	0	0.0006	0.0707	0.0003	0.0509
Zn	XRF	0.0016	0.0035	47	1	1%	3	6%	1	0.0005	0.0630	0.0008	0.0304
Br	XRF	0.0017	0.0014	47	7	7%	6	13%	1	0.0006	0.1234	0.0009	0.0212
Sr	XRF	0.0026	0.0015	47	43	84%	35	74%	0	0.0009	0.2157	0.0003	0.1417
Sn	XRF	0.0057	0.0076	47	27	41%	32	68%	0	0.0019	0.1857	0.0005	0.1985
Pb	XRF	0.0039	0.0031	47	12	18%	12	26%	0	0.0013	0.1146	0.0012	0.0612

MDL = minimum detection limit; LQL = lower quantifiable limit; N = number of collocated sample data pairs; N < MDL (or LQL) = number of data pairs with one or both values below the MDL (or LQL)

**Table 3.1** Sampling overview for monitoring stations.

Station	Abbreviation	Number of Valid Samples	Starting Date	Ending Date	Major data gaps	Characteristics
Yuen Long	YL	654	Jan. 2, 1998	Dec. 26, 2008		Urban / Residential
Tung Chung	TC	575	Apr. 3, 1999	Dec. 28, 2008		New Town / Residential
Tsuen Wan	TW	631	Jan. 3, 1998	Dec. 26, 2008	Feb. 6, 2003 – Apr. 23, 2003	Urban / Residential / Commercial
Sham Shui Po	SSP	586	Jan. 11, 1998	Dec. 26, 2008	Nov. 29, 2003 – Mar. 16, 2004; May 12, 2007 – Jun. 5, 2008	Urban/ Commercial
Central / Western	CW	641	Jan. 2, 1998	Dec. 26, 2008	Nov. 9, 2001 – Mar. 27, 2002	Urban / Residential
Kwun Tong	KT	598	Jan. 4, 1998	Dec. 31, 2008	Oct. 31, 1998 – Jan. 11, 1999; Apr. 13, 2002 – Feb. 1, 2003	Commercial / Residential / Near ferry pier
Mong Kok	MK	482	Jan. 2, 2001	Dec. 27, 2008		Roadside, relocated at the end of 2000

**Table 3.2** Source profiles and uncertainties derived by PMF. Source profiles are expressed in mass percentages.

	Vehicle Exhaust	Residual Oil Combustion	Fresh Sea Salt	Aged Sea Salt	Crustal Soil / Dust	Secondary Sulfate	Secondary Nitrate	Trace Metals	Coal Comb. / Biomass Burning
Al	0%	2±1%	0%	0%	2%	0%	0%	1%	0%
As	0%	0%	0%	0%	0%	0%	0%	0%	0%
Ca	1%	7±4%	1%	1%	6%	0%	0%	4±1%	1%
Cd	0%	0%	0%	0%	0%	0%	0%	0%	0%
Cl <sup>-</sup>	0%	0%	44±3%	0%	0%	0%	1%	0%	0%
Fe	1%	2±1%	0%	0%	3%	0%	0%	4±1%	0%
K <sup>+</sup>	0%	3±2%	0%	1%	0%	0%	0%	0%	6%
Mg	0%	0%	2%	2%	1%	0%	0%	0%	0%
Mn	0%	0%	0%	0%	0%	0%	0%	0%	0%
Na <sup>+</sup>	0%	0%	19±1%	16%	0%	0%	0%	0%	0%
NH <sub>4</sub> <sup>+</sup>	0%	0±3%	2%	0%	0%	18±1%	9%	0±1%	0±1%
Ni	0%	1±1%	0%	0%	0%	0%	0%	0%	0%
NO <sub>3</sub> <sup>-</sup>	0%	0±2%	0±2%	6±1%	4±1%	0%	34±1%	0±1%	0±1%
Pb	0%	0%	0%	0%	0%	0%	0%	1%	0%
SO <sub>4</sub> <sup>2-</sup>	4%	11±11%	0%	39±1%	5±1%	51±2%	0%	11±4%	10±2%
V	0%	4±3%	0%	0%	0%	0%	0%	0%	0%
Zn	0%	1±1%	0%	0%	0%	0%	0%	5±1%	0%
EC	34±1%	0±9%	7±1%	4±1%	0±1%	0±1%	0±1%	0±1%	6±2%
OC	32±1%	31±17%	4±2%	0%	0±1%	5±1%	23±1%	5±5%	34±2%
Unidentified	28%	38%	20%	30%	78%	25%	32%	68%	42%

**Table 3.3** Tracers applied for source identification

Source	Tracers
Vehicle Exhaust	EC, OC
Residual Oil Combustion	Ni, V
Fresh Sea Salt	Na <sup>+</sup> , Mg <sup>2+</sup> , Cl <sup>-</sup>
Aged Sea Salt	Na <sup>+</sup> , Mg <sup>2+</sup> , NO <sub>3</sub> <sup>-</sup> , SO <sub>4</sub> <sup>2-</sup>
Crustal Soil / Dust	Al, Ca, Mg, Fe
Secondary Sulfate	NH <sub>4</sub> <sup>+</sup> , SO <sub>4</sub> <sup>2-</sup>
Secondary Nitrate	NH <sub>4</sub> <sup>+</sup> , NO <sub>3</sub> <sup>-</sup>
Trace Metals	Zn
Coal Combustion / Biomass Burning	As, Cd, Pb, K <sup>+</sup> , OC

**Table 3.4** Annual-, wintertime-, summertime-, and transition season- averaged source contributions (unit:  $\mu\text{g}/\text{m}^3$ ) and contribution percentages for general stations from 1998 to 2008 and for roadside station from 2001 to 2008

	General stations (YL, TC, TW, SSP, CW & KT)					Roadside station (MK)				
	Annual	Winter	Summer	Transitional	Win/Sum	Annual	Winter	Summer	Transitional	Win/Sum
Vehicle Exhaust	8.5 (15%)	8.2 (11%)	8.9 (25%)	8.4 (15%)	0.92	21.3 (30%)	19.0 (22%)	23.3 (47%)	21.4 (28%)	0.82
Residual Oil	0.2 (0.4%)	0.2 (0.3%)	0.3 (1%)	0.2 (0.4%)	0.74	0.3 (0.4%)	0.2 (0.2%)	0.3 (1%)	0.3 (0.4%)	0.57
Fresh Sea Salt	1.8 (3%)	2.1 (3%)	1.8 (5%)	1.4 (2%)	1.2	2.3 (3%)	2.5 (3%)	2.8 (6%)	1.7 (2%)	0.90
Aged Sea Salt	7.1 (13%)	6.2 (8%)	6.5 (19%)	8.6 (15%)	0.95	7.2 (10%)	6.5 (7%)	6.5 (13%)	8.5 (11%)	1.0
Crustal Soil / Dust	6.7 (12%)	9.7 (13%)	3.2 (9%)	7.3 (13%)	3.0	7.5 (11%)	10.3 (12%)	4.0 (8%)	8.4 (11%)	2.6
Secondary Sulfate	12.3 (22%)	16.1 (22%)	6.9 (20%)	14.1 (25%)	2.3	13.4 (19%)	16.4 (19%)	6.9 (14%)	17.1 (22%)	2.4
Secondary Nitrate	8.3 (15%)	14.9 (20%)	3.1 (9%)	7.2 (13%)	4.8	11.0 (16%)	18.2 (21%)	4.5 (9%)	10.6 (14%)	4.0
Trace Metals	2.6 (5%)	4.1 (5%)	1.0 (3%)	2.9 (5%)	4.0	2.8 (4%)	4.1 (5%)	0.9 (2%)	3.4 (4%)	4.4
Coal Comb. / Biomass Burning	7.3 (13%)	12.7 (17%)	2.6 (7%)	7.1 (12%)	5.0	6.2 (9%)	10.4 (12%)	1.7 (3%)	6.7 (9%)	6.1
Local sources	17.6 (32%)	16.7 (22%)	17.4 (50%)	18.6 (32%)	1.0	31.0 (44%)	28.2 (33%)	32.9 (66%)	32.0 (41%)	0.86
Non-local sources	37.2 (67%)	57.6 (77%)	16.8 (48%)	38.6 (67%)	3.4	41.0 (58%)	59.3 (69%)	18.1 (36%)	46.3 (60%)	3.3
Residual*	0.5 (1%)	0.6 (1%)	0.8 (2%)	0.2 (0.3%)		-1.1 (-2%)	-0.9 (-1%)	-1.2 (-3%)	-1.1 (-1%)	
Total	55.3	74.8	35.1	57.4		70.9	86.6	49.7	77.1	

\* Residual refers to particle mass that is not apportioned into identified sources by PMF, i.e. the total mass in the residual matrix.

**Table 3.5** APCA-modeled source contribution estimates for 4-to-9 factor solutions, mapped onto the PMF-resolved source categories. All concentration values in  $\mu\text{g}/\text{m}^3$ .

(a) Constrained multivariable regression (constant term not included)

No. of Factors	Coal & Biomass Comb.	Sea Salt	Residual Oil	Crustal Soil	Vehicle Exhaust	Secondary Sulfate	Secondary Nitrate	OM Factor	Chloride Factor
9	7.84	4.15	3.24	5.07	8.45	8.30	-0.49	4.83	-0.36
8	8.22	3.83	3.07	5.26	10.64	9.20	1.59	-1.03	
7	8.99	3.78	3.10	5.14	10.26	8.41	1.16		
6	7.08	4.01	2.91	5.27	10.43	11.26			
5	17.09	6.53	4.63	4.18	8.23				
4	14.62	6.50	10.45	8.25					

(b) Unconstrained multivariable regression (constant term included)

No. of Factors	Coal & Biomass Comb.	Sea Salt	Residual Oil	Crustal Soil	Vehicle Exhaust	Secondary Sulfate	Secondary Nitrate	OM Factor	Chloride Factor	Constant
9	7.81	4.01	3.21	5.04	8.30	8.26	-0.49	4.80	-0.38	0.52
8	8.08	3.27	2.94	5.11	9.78	8.96	1.57	-1.09		2.45
7	8.88	3.36	3.00	5.00	9.60	8.17	1.16			1.92
6	7.05	3.84	2.86	5.19	10.17	11.14				0.82
5	16.70	5.75	4.36	3.99	7.53					2.75
4	13.76	4.52	8.51	7.08						7.21

**Table 3.6** Source profiles derived from APCA analysis for a six factor solution. The values are expressed as mass percentages.

Species	Secondary Sulfate	Vehicular Exhaust	Coal & Biomass Comb.	Crustal Soil	Sea Salt	Residual Oil
Al	0.21	-0.09	0.64	3.26	0.32	0.25
As	0.01	0.00	0.03	0.02	-0.01	0.02
Ca	0.14	0.94	1.85	7.99	1.51	0.75
Cd	0.00	0.00	0.01	0.01	0.00	0.00
Cl <sup>-</sup>	-1.97	0.56	0.30	-0.50	29.69	-1.48
Fe	0.58	1.18	1.26	4.71	0.11	0.73
K <sup>+</sup>	1.03	0.64	3.49	2.56	0.26	-0.05
Mg	0.25	-0.03	-0.06	1.87	3.46	-0.04
Mn	0.03	0.01	0.07	0.16	-0.01	0.02
Na <sup>+</sup>	2.22	0.18	-1.45	1.96	28.65	0.83
NH <sub>4</sub> <sup>+</sup>	16.22	-0.20	10.59	3.63	-2.24	13.02
Ni	0.01	0.01	0.01	0.01	0.00	0.11
NO <sub>3</sub> <sup>-</sup>	10.14	3.04	8.32	8.63	24.09	7.62
Pb	0.11	0.04	0.41	0.28	-0.10	0.10
SO <sub>4</sub> <sup>2-</sup>	46.89	1.54	26.73	26.66	12.74	34.65
V	0.01	0.02	0.01	0.00	-0.01	0.26
EC	0.50	34.66	0.90	4.68	3.55	10.08
OM	23.61	57.51	46.88	34.08	-2.01	33.12

**Table 3.7** Source profiles derived from UNMIX for a six factor solution. The values are expressed as mass percentages.

Species	Secondary Sulfate	Vehicle Exhaust	Coal & Biomass Comb.	Crustal Soil	Sea Salt	Residual Oil
Al	0.19	-0.09	0.42	4.75	0.24	0.50
As	0.00	-0.01	0.05	0.00	0.00	0.01
Ca	0.17	1.08	1.39	11.24	2.39	1.38
Cd	0.00	0.00	0.01	0.00	0.00	0.00
Cl <sup>-</sup>	-0.77	0.22	0.11	-0.55	53.17	-0.51
Fe	0.62	1.16	0.99	6.57	-0.12	1.14
K <sup>+</sup>	0.39	0.81	4.40	1.03	0.96	-1.38
Mg	0.51	0.17	-0.43	3.16	4.50	0.37
Mn	0.02	0.01	0.08	0.19	-0.01	0.01
Na <sup>+</sup>	4.00	1.81	-3.24	5.67	38.07	2.87
NH <sub>4</sub> <sup>+</sup>	13.97	-0.80	11.85	-1.71	-11.88	10.08
Ni	0.01	0.01	0.00	0.01	0.00	0.12
NO <sub>3</sub> <sup>-</sup>	14.34	1.08	5.22	12.89	23.01	5.98
Pb	0.01	0.05	0.54	0.05	-0.01	-0.05
SO <sub>4</sub> <sup>2-</sup>	39.57	6.50	28.74	21.80	-15.75	31.30
V	0.01	0.01	-0.01	0.01	0.00	0.31
EC	3.40	33.23	-0.19	9.11	9.59	18.37
OM	23.58	54.78	50.08	25.77	-4.16	29.48

**Table 3.8** Study-average source contribution estimates for the APCA and UNMIX six factor solutions. All concentration values in  $\mu\text{g}/\text{m}^3$ .

Model	Secondary Sulfate	Vehicle Exhaust	Coal & Biomass Comb.	Crustal Soil	Sea Salt	Residual Oil
APCA	11.3	10.4	7.1	5.3	4.0	2.9
UNMIX	13.3	10.1	8.9	3.6	2.0	3.1

**Table 3.9** Annual, summer, winter and transition (trans.) averaged source contribution estimates from January 1998 to December 2008 for the general stations and the roadside station. W/S represents the winter to summer ratio. Table 5(a): APCA; Table 5(b) UNMIX. All concentration values in  $\mu\text{g}/\text{m}^3$ .

5(a) APCA	General Stations (YL, TC, TW, SSP, CW, KT)					Roadside Station (MK)				
	Annual	Summer	Winter	Trans.	W/S	Annual	Summer	Winter	Trans.	W/S
Secondary Sulphate	11.7	7.2	13.8	14.5	1.9	13.4	7.2	16.9	16.6	2.3
Vehicle Exhaust	8.4	8.0	9.2	7.7	1.1	21.4	23.7	20.1	20.3	0.8
Coal & Biomass Combust.*	8.7	3.0	16.4	7.2	5.4	7.2	2.0	13.1	6.7	6.4
Crustal Soil	5.3	3.2	7.0	6.3	2.2	6.0	3.8	7.6	7.4	2.0
Sea Salt	4.0	3.6	4.4	3.8	1.2	4.6	4.7	4.8	4.3	1.0
Residual Oil	3.1	3.9	2.6	2.9	0.7	3.0	3.6	1.9	3.5	0.5

(\*) Coal Combustion & Biomass Burning

5(b) UNMIX	General Station					Roadside Station				
	Annual	Summer	Winter	Trans.	W/S	Annual	Summer	Winter	Trans.	W/S
Secondary Sulfate	13.5	7.6	18.1	15.1	2.37	14.9	7.1	20.9	17.1	2.9
Vehicular Exhaust	7.0	7.5	6.8	6.5	0.91	18.9	22.1	16.7	17.7	0.8
Coal & Biomass Combust.*	10.3	4.3	17.2	10.2	4.04	9.5	3.1	15.2	10.7	4.8
Crustal Soil	3.6	1.9	5.3	4.1	2.86	3.8	2.0	5.3	4.6	2.7
Sea Salt	2.0	1.8	2.2	1.8	1.22	2.4	2.6	2.5	2.1	1.0
Residual Oil	3.2	3.6	2.9	3.0	0.80	3.2	3.5	2.3	3.7	0.7

(\*) Coal Combustion & Biomass Burning

**Table 3.10** Rescaled source profiles for the CMB modeling. The original profiles were rescaled such that the species measured in this study summed to unity for each profile.

Species	Paved Road Dust		Residual Oil Combustion		Fresh Sea Salt		Light Duty Vehicle-Diesel Exhaust		Secondary Sulphate		Secondary Nitrate		Agricultural Vegetative Burning	
	Mass Fraction	Uncertainty	Mass Fraction	Uncertainty	Mass Fraction	Uncertainty	Mass Fraction	Uncertainty	Mass Fraction	Uncertainty	Mass Fraction	Uncertainty	Mass Fraction	Uncertainty
Al	1.79E-01	6.23E-02	6.90E-03	1.86E-03	0	0	3.19E-04	1.52E-03	0	0	0	0	9.77E-04	1.72E-04
As	9.65E-06	5.19E-04	0	0	0	0	2.00E-05	1.20E-04	0	0	0	0	8.81E-07	4.49E-05
Ca	2.14E-01	5.51E-02	2.06E-02	5.11E-03	1.22E-02	2.12E-03	2.69E-04	3.95E-04	0	0	0	0	3.92E-03	4.56E-04
Cd	5.79E-05	1.11E-03	0	0	0	0	0	0	0	0	0	0	8.37E-05	2.34E-04
Cl <sup>-</sup>	1.49E-02	5.57E-03	0	0	5.75E-01	9.88E-02	4.69E-04	6.61E-04	0	0	0	0	4.59E-02	5.22E-03
Fe	1.23E-01	2.62E-02	3.87E-02	6.47E-03	0	0	2.30E-04	9.34E-04	0	0	0	0	1.98E-03	2.24E-04
K <sup>+</sup>	5.22E-02	1.44E-02	3.65E-03	8.22E-04	1.18E-02	2.02E-03	1.30E-04	4.46E-04	0	0	0	0	5.91E-02	6.87E-03
Mg	5.03E-03	2.02E-03	0	0	0	0	0	0	0	0	0	0	3.96E-04	4.33E-04
Mn	2.45E-03	5.46E-04	5.99E-04	1.21E-04	0	0	2.99E-05	2.10E-04	0	0	0	0	1.67E-05	4.41E-05
Na <sup>+</sup>	1.73E-02	5.32E-03	4.56E-02	1.30E-02	3.21E-01	5.48E-02	0	0	0	0	0	0	6.29E-03	8.63E-04
NH <sub>4</sub> <sup>+</sup>	1.27E-03	2.46E-03	0	0	0	0	0	0	2.73E-01	1.00E-03	2.25E-01	1.00E-03	1.28E-03	5.47E-04
Ni	1.88E-04	1.08E-04	6.98E-02	1.22E-02	0	0	9.98E-06	8.99E-05	0	0	0	0	2.91E-05	1.61E-05
NO <sub>3</sub> <sup>-</sup>	9.36E-04	2.59E-03	8.47E-03	3.20E-03	3.71E-05	6.06E-06	0	0	0	0	7.75E-01	1.00E-03	5.84E-03	8.42E-04
Pb	2.56E-03	6.44E-04	1.43E-03	4.80E-04	0	0	3.39E-04	7.52E-04	0	0	0	0	0	0
SO <sub>4</sub> <sup>2-</sup>	2.02E-02	5.52E-03	6.26E-01	6.52E-02	8.02E-02	1.38E-02	0	0	7.27E-01	1.00E-03	0	0	3.22E-02	3.67E-03
V	4.83E-05	1.28E-03	4.48E-02	7.68E-03	0	0	5.99E-05	6.59E-04	0	0	0	0	0	0
EC	3.06E-02	1.92E-02	3.15E-02	1.76E-02	0	0	6.37E-01	2.42E-01	0	0	0	0	5.90E-02	1.41E-02
OC	3.36E-01	1.17E-01	1.02E-01	4.43E-02	0	0	3.62E-01	1.50E-01	0	0	0	0	7.83E-01	8.99E-02

**Table 3.11** Study-average source contribution estimates (SCEs) for the CMB modeling and PMF modeling. The PMF SCEs have been adjusted to include only those samples with CMB modeling results meeting the mass closure criterion and thus differ from the SCEs reported by Lau *et al.* (2010). All concentration values in  $\mu\text{g}/\text{m}^3$ .

Source	General		Roadside	
	CMB	PMF	CMB	PMF
Vehicle Exhaust	7.7	6.2	17.1	15.2
Residual Oil	0.1	0.2	0.2	0.2
Sea Salt	2.6	5.9	3.6	6.8
Crustal Soil/Dust	10.1	1.4	10.7	1.7
Secondary Sulfate	11.8	11.2	11.9	11.1
Secondary Nitrate	4.4	6.2	5.4	7.9
Biomass Burning	3.1	4.7	3.3	3.9
Total	39.7	35.7	52.2	46.7

**Table 4.1**  $\text{PM}_{2.5}$ -to- $\text{PM}_{10}$  ratio of total PM mass and concentration of each species, expressed as the slope of the straight line that best fit the data by “least squares” method

Species	Ratio	Species	Ratio	Species	Ratio
PM mass	0.68	Mg	0.19	Pb	0.76
Al	0.37	Mn	0.75	$\text{SO}_4^{2-}$	1.04
Ca	0.11	$\text{Na}^+$	0.10	V	1.12
$\text{Cl}^-$	0.28	$\text{NH}_4^+$	0.94	EC	0.66
Fe	0.32	Ni	0.89	OC	0.86
$\text{K}^+$	0.95	$\text{NO}_3^-$	0.72		

**Table 4.2** Grand average PM<sub>2.5</sub> mass source contribution estimates ( $\mu\text{g}/\text{m}^3$ ) across all sites and all periods for the eight PMF-resolved factors using: sum-of-species with OC; sum of species with OM = 1.6× OC (denoted SSOM on the last column); Teflon filter gravimetric mass as the Total Variable in PMF; and Quartz filter gravimetric mass as the Total Variable in PMF.

Source Category	Sum of Species (OC)	Sum of Species (OM)	Teflon Gravimetric Mass	Quartz Gravimetric Mass	Ratio (SSOM/Quartz)
Secondary Sulfate	10.1	10.1	11.4	12.7	1.25
Vehicular Exhaust	9.2	10.8	9.2	10.4	0.96
Secondary Nitrate	3.5	4.0	4.5	5.0	1.25
Biomass Burning	3.1	4.6	4.5	4.3	0.94
Trace Metals	2.5	2.6	3.0	2.7	1.03
Residual Oil	1.5	1.7	1.9	2.2	1.30
Soil/Dust	1.2	1.2	1.6	1.6	1.35
Fresh Sea Salt	0.7	0.8	1.0	1.4	1.71

**Table 4.3** APCA-modeled source contribution estimates for 5-to-9 factor solutions, mapped onto the PMF-resolved source categories. All concentration values in  $\mu\text{g}/\text{m}^3$ .

No. of Factors	Vehicle Exhaust	Crustal Soil	Biomass Burning	Residual Oil	Sea Salt	Secondary Sulfate	Secondary Nitrate	Trace Metals	Unknown
9	10.03	7.05	4.47	3.51	0.36	12.79	0.80	0.24	-0.02
8	10.02	7.26	4.88	3.48	0.35	12.50	0.69	0.06	
7	10.35	6.80	6.25	3.54	0.38	11.44	0.46		
6	10.14	7.04	6.12	3.64	0.72	11.44			
5	10.21	6.30	16.99	5.20	0.10				

**Table 4.4** Species loadings onto the seven factor UNMIX solution for the PM<sub>2.5</sub> data. All concentration values in µg/m<sup>3</sup>.

Species	Secondary Sulfate	Vehicle Exhaust	Biomass Burning	Secondary Nitrate	Residual Oil Combustion	Soil/Dust	Salt
OC	-0.2350	3.2200	3.8000	0.4160	0.4140	-0.2970	-0.0080
EC	0.5280	6.0300	-0.1090	0.0488	0.2760	-0.1370	0.0634
NO <sub>3</sub> <sup>-</sup>	-0.1840	0.1180	0.0988	1.6800	0.0293	0.1040	0.0485
SO <sub>4</sub> <sup>-</sup>	7.6200	0.1300	0.8380	1.5000	0.6790	0.1030	-0.1110
NH <sub>4</sub> <sup>+</sup>	2.1800	0.0976	0.2770	0.8950	0.1530	-0.1510	0.0418
Al	0.0357	-0.0025	0.0154	0.0000	0.0008	0.0635	0.0010
Ca	-0.0028	0.0233	0.0199	-0.0047	0.0005	0.0573	0.0013
Cl	-0.0007	-0.0020	-0.0038	0.0023	-0.0020	-0.0023	0.1120
Fe	0.0157	0.0408	0.0439	0.0235	0.0009	0.0583	-0.0012
K	0.0593	0.0046	0.3240	0.0594	-0.0217	0.0352	-0.0119
Mn	0.0033	0.0000	0.0054	0.0030	-0.0006	0.0023	-0.0001
Ni	0.0008	0.0000	-0.0003	0.0001	0.0046	0.0001	0.0002
Pb	0.0012	-0.0028	0.0447	0.0042	0.0015	0.0011	0.0022
Si	0.0183	0.0350	0.0924	-0.0228	0.0001	0.1730	0.0027
Ti	0.0012	0.0000	0.0012	0.0011	0.0008	0.0050	0.0001
V	0.0031	0.0000	-0.0034	0.0008	0.0155	0.0003	0.0004
<b>Quartz Filter</b>							
<b>Total Mass</b>	<b>12.30</b>	<b>10.80</b>	<b>7.08</b>	<b>6.58</b>	<b>1.87</b>	<b>0.48</b>	<b>0.21</b>

**Table 4.5** PM<sub>2.5</sub> mass source contribution estimates for the eight PMF-resolved factors stratified by site and year reported in µg/m<sup>3</sup> (percentages in parenthesis). The quartz fiber filter gravimetric mass was used as the Total Variable in the PMF modeling.

Site & Period	Secondary Sulfate		Vehicle Exhaust		Secondary Nitrate		Biomass Burning		Trace Metals		Residual Oil		Soil/Dust		Fresh Sea Salt		Total
<b>All Sites</b>	12.7	(31.5)	10.4	(25.9)	5.0	(12.5)	4.3	(10.7)	2.7	(6.6)	2.2	(5.4)	1.6	(4.0)	1.4	(3.4)	40.2
2000-01	10.2	(24.2)	16.4	(38.9)	2.3	(5.5)	6.5	(15.4)	1.4	(3.3)	1.8	(4.4)	1.9	(4.5)	1.6	(3.8)	42.1
2004-05	15.4	(34.7)	10.2	(23.1)	5.0	(11.4)	4.6	(10.3)	3.5	(8.0)	2.5	(5.6)	1.6	(3.6)	1.4	(3.2)	44.3
2008-09	11.7	(33.4)	6.8	(19.3)	6.7	(19.0)	2.6	(7.4)	2.6	(7.5)	2.1	(5.9)	1.4	(4.0)	1.2	(3.4)	35.1
<b>TW</b>	13.2	(34.9)	7.6	(20.2)	4.6	(12.1)	4.8	(12.8)	2.7	(7.0)	2.4	(6.4)	1.6	(4.1)	0.9	(2.5)	37.9
2000-01	11.0	(29.3)	9.6	(25.6)	3.0	(8.0)	7.3	(19.4)	1.8	(4.7)	2.0	(5.2)	1.8	(4.9)	1.0	(2.8)	37.3
2004-05	16.0	(37.2)	8.6	(19.9)	4.3	(9.9)	4.8	(11.2)	3.6	(8.4)	3.1	(7.1)	1.7	(3.9)	1.0	(2.4)	43.1
2008-09	12.5	(37.4)	5.1	(15.3)	6.2	(18.5)	2.8	(8.4)	2.5	(7.5)	2.3	(6.8)	1.2	(3.7)	0.8	(2.4)	33.6
<b>MK</b>	11.3	(20.6)	25.4	(46.4)	5.9	(10.8)	4.0	(7.4)	2.6	(4.8)	2.1	(3.9)	1.8	(3.3)	1.5	(2.8)	54.7
2000-01	9.1	(14.4)	37.9	(59.8)	2.9	(4.6)	6.9	(10.9)	1.3	(2.0)	1.9	(3.0)	1.8	(2.9)	1.6	(2.5)	63.4
2004-05	14.2	(25.0)	23.1	(40.7)	6.2	(10.9)	3.9	(6.8)	3.7	(6.6)	2.4	(4.1)	1.8	(3.1)	1.6	(2.8)	56.8
2008-09	10.3	(22.8)	16.8	(37.1)	8.3	(18.3)	1.8	(3.9)	2.8	(6.2)	2.1	(4.6)	1.8	(4.0)	1.4	(3.1)	45.3
<b>YL</b>	14.6	(35.3)	5.9	(14.2)	8.2	(19.8)	4.0	(9.6)	3.5	(8.4)	2.1	(5.0)	1.5	(3.6)	1.7	(4.0)	41.3
2000-01	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
2004-05	16.7	(35.3)	7.3	(15.4)	8.1	(17.1)	4.9	(10.5)	4.0	(8.5)	2.5	(5.2)	1.6	(3.4)	2.2	(4.6)	47.2
2008-09	12.7	(35.3)	4.6	(12.8)	8.3	(23.0)	3.1	(8.7)	3.0	(8.4)	1.7	(4.8)	1.4	(3.8)	1.2	(3.3)	36.0
<b>HT</b>	12.1	(45.1)	1.3	(4.9)	2.2	(8.1)	4.2	(15.6)	2.1	(7.8)	2.0	(7.5)	1.5	(5.5)	1.5	(5.5)	26.9
2000-01	10.5	(40.9)	1.6	(6.4)	1.0	(4.0)	5.4	(21.0)	1.2	(4.7)	1.7	(6.5)	2.0	(7.8)	2.2	(8.7)	25.5
2004-05	14.5	(48.4)	2.0	(6.8)	1.6	(5.4)	4.6	(15.5)	2.8	(9.4)	2.1	(7.1)	1.3	(4.2)	1.0	(3.2)	30.0
2008-09	11.3	(45.0)	0.3	(1.3)	3.8	(15.1)	2.7	(10.8)	2.2	(8.7)	2.2	(9.0)	1.2	(4.8)	1.3	(5.4)	25.1

**Table 5.1** Change in PMF-modeled PM<sub>10</sub> annual median of the scaled source contribution estimate from a linear least-squares regression of annual median source contribution estimate on year after stratifying the data by air mass transport pattern. Values are shown only for those cases when the change is statistically different from zero at the 95% confidence level.\*

Air Mass	Vehicle Exhaust	Crustal/Soil	Smelting (Zn)	Secondary Sulfate
Slow ECC	-0.12 ± 0.06	-0.04 ± 0.03		0.14 ± 0.06
Fast ECC	-0.11 ± 0.05			0.10 ± 0.05
Stagnant	-0.06 ± 0.05			
S/SW	-0.08 ± 0.04		-0.02 ± 0.01	
East	-0.08 ± 0.04			

(\*) No significant changes for residual oil, fresh sea salt, aged sea salt, secondary nitrate, and biomass burning.

**Table 5.2** Change in PM<sub>10</sub> mass ( $\mu\text{g}/\text{m}^3/\text{year}$ ) from a linear least-squares regression of annual median concentration on year after stratifying the data by air mass transport pattern. Values are shown only for those cases when the change is statistically different from zero at the 95% confidence level.\*

Air Mass	YL	SSP	KT	EN	TM	TC	MK	CL	CB
Slow ECC	2.6 ± 2.1				2.7 ± 1.8	2.6 ± 1.8			
Fast ECC	1.9 ± 1.8				2.3 ± 1.3				
Stagnant		-1.4 ± 1.1	-2.5 ± 1.9						-4.0 ± 2.3
S/SW					0.9 ± 0.6		-1.5 ± 0.8		-3.3 ± 1.4
East		-1.6 ± 1.5	-2.2 ± 1.3	-1.4 ± 1			-3.1 ± 1.6	-2.0 ± 1.1	-4.5 ± 3.5

(\*) No significant changes for CW, KC, ST, TP, and TW.

**Table 5.3** Average concentrations in PM<sub>10</sub> mass (unit:  $\mu\text{g}/\text{m}^3$ ) from local, PRD and non-PRD trajectory classes at TW in 2006 and in summer and winter.

<b>Trajectory Class</b>		<b>LOCAL</b>	<b>PRD</b>	<b>NON-PRD</b>
Total	Average PM <sub>10</sub> concentration	27	58~72	64~67
	Average wind speed (m/s)	2.2	1.5~1.8	2.6~2.8
Summer	Average PM <sub>10</sub> concentration	23	57~64	40~50
	Average wind speed (m/s)	2.2	1.5~2.3	2.7~2.8
Winter	Average PM <sub>10</sub> concentration	31	0~74	59~63
	Average wind speed (m/s)	2.2	1.4~1.6	2.3~2.8

**Table 5.4** Absolute contributions (in unit  $\mu\text{g}/\text{m}^3$ ) and contribution percentages of local, PRD and non-PRD sources in summer and winter and in the entire year of 2006 at TW.

<b>Source Category</b>		<b>Local</b>	<b>PRD</b>	<b>Non-PRD</b>
Total	Absolute Contribution ( $\mu\text{g}/\text{m}^3$ )	17~18	3~13	33~43
	Contribution Percentage	27~28%	4~20%	52~66%
Summer	Absolute Contribution ( $\mu\text{g}/\text{m}^3$ )	16	4~18	12~26
	Contribution Percentage	35%	8~38%	27~56%
Winter	Absolute Contribution ( $\mu\text{g}/\text{m}^3$ )	21~23	0~12	29~42
	Contribution Percentage	33~36%	0~19%	45~67%

**Table 5.5** Average concentrations in PM<sub>10</sub> mass (unit:  $\mu\text{g}/\text{m}^3$ ) from local, PRD and non-PRD trajectory classes at TW in the high mass concentration hours [Scenario 1] and the percentage of occurrence of high mass concentration hours in 2006 and in summer and winter.

<b>Trajectory Class</b>		<b>LOCAL</b>	<b>PRD</b>	<b>NON-PRD</b>
	Average PM <sub>10</sub> concentration	27	118~127	127
Total	Average wind speed (m/s)	2.2	1.9~5.4	2.4~2.9
	% of Occurrence	0%	3~19%	13~16%
	Average PM <sub>10</sub> concentration	23	118~133	123~135
Summer	Average wind speed (m/s)	2.2	3.2~5.4	1.7~2.5
	% of Occurrence	0%	13~19%	2~9%
	Average PM <sub>10</sub> concentration	31	0~131	133~138
Winter	Average wind speed (m/s)	2.2	0~1.0	1.5~2.5
	% of Occurrence	0%	0~21%	6~10%

**Table 5.6** Absolute contributions (in unit  $\mu\text{g}/\text{m}^3$ ) and contribution percentages of local, PRD and non-PRD sources in the high mass concentration hours [Scenario 1] in 2006 at TW.

<b>Source Category</b>		<b>Local</b>	<b>PRD</b>	<b>Non-PRD</b>
Total	Absolute Contribution ( $\mu\text{g}/\text{m}^3$ )	17~18	2~40	70~107
	Contribution Percentage	14%	2~31%	55~85%
Summer	Absolute Contribution ( $\mu\text{g}/\text{m}^3$ )	13	21~106	13~98
	Contribution Percentage	10%	16~81%	10~74%
Winter	Absolute Contribution ( $\mu\text{g}/\text{m}^3$ )	32~39	0~50	45~102
	Contribution Percentage	24~29%	0~37%	33~76%

**Table 5.7** Average concentrations in PM<sub>10</sub> mass (unit:  $\mu\text{g}/\text{m}^3$ ) from local, PRD and non-PRD trajectory classes at TW in the high mass concentration hours [Scenario 2] in 2006 and in summer and winter.

<b>Trajectory Class</b>		<b>LOCAL</b>	<b>PRD</b>	<b>NON-PRD</b>
Summer	Average PM <sub>10</sub> concentration	23	99~108	84~100
	Average wind speed (m/s)	2.2	2.7~4.4	2.3~2.4
Winter	Average PM <sub>10</sub> concentration	31	0~132	134~138
	Average wind speed (m/s)	2.2	0~1.0	1.6~2.5

**Table 5.8** Absolute contributions of PM<sub>10</sub> (in unit  $\mu\text{g}/\text{m}^3$ ) and contribution percentages of local, PRD and non-PRD sources in the high mass concentration hours [Scenario 2] in 2006 at TW.

<b>Source Category</b>		<b>Local</b>	<b>PRD</b>	<b>Non-PRD</b>
Summer	Absolute Contribution ( $\mu\text{g}/\text{m}^3$ )	14~15	13~63	23~72
	Contribution Percentage	14~15%	13~63%	23~72%
Winter	Absolute Contribution ( $\mu\text{g}/\text{m}^3$ )	31~38	0~51	46~103
	Contribution Percentage	23~28%	0~38%	34~77%

**Table 5.9** Average concentrations in PM<sub>2.5</sub> mass (unit:  $\mu\text{g}/\text{m}^3$ ) from local, PRD and non-PRD trajectory classes at TW in 2006 and in summer and winter.

<b>Trajectory Class</b>		<b>LOCAL</b>	<b>PRD</b>	<b>NON-PRD</b>
Total	Average PM <sub>10</sub> concentration	19	43~54	46~49
	Average wind speed (m/s)	2.2	1.5~1.8	2.6~2.8
Summer	Average PM <sub>10</sub> concentration	16	42~50	29~38
	Average wind speed (m/s)	2.2	1.5~2.3	2.7~2.8
Winter	Average PM <sub>10</sub> concentration	23	0~55	44~47
	Average wind speed (m/s)	2.2	1.4~1.6	2.3~2.8

**Table 5.10** Absolute contributions of PM<sub>2.5</sub> (in unit  $\mu\text{g}/\text{m}^3$ ) and contribution percentages of local, PRD and non-PRD sources in summer and winter and in the entire year of 2006 at TW.

<b>Source Category</b>		<b>Local</b>	<b>PRD</b>	<b>Non-PRD</b>
Total	Absolute Contribution ( $\mu\text{g}/\text{m}^3$ )	12~13	2~10	24~31
	Contribution Percentage	26~28%	4~21%	51~67%
Summer	Absolute Contribution ( $\mu\text{g}/\text{m}^3$ )	11	3~14	9~20
	Contribution Percentage	32%	9~41%	27~59%
Winter	Absolute Contribution ( $\mu\text{g}/\text{m}^3$ )	16~17	0~9	21~32
	Contribution Percentage	33~35%	0~20%	45~67%

**Table 5.11** Average concentrations in PM<sub>2.5</sub> mass (unit:  $\mu\text{g}/\text{m}^3$ ) from local, PRD and non-PRD trajectory classes at TW in the high mass concentration hours [Scenario 1] and the percentage of occurrence of high mass concentration hours in 2006 and in summer and winter.

<b>Trajectory Class</b>		<b>LOCAL</b>	<b>PRD</b>	<b>NON-PRD</b>
	Average PM <sub>10</sub> concentration	19	96~99	95~97
Total	Average wind speed (m/s)	2.2	1.7~5.5	2.3~2.7
	% of Occurrence	0%	2~18%	11~15%
	Average PM <sub>10</sub> concentration	16	96~103	104
Summer	Average wind speed (m/s)	2.2	2.6~5.5	1.7~2.1
	% of Occurrence	0%	8~22%	2~11%
	Average PM <sub>10</sub> concentration	23	0~101	101~102
Winter	Average wind speed (m/s)	2.2	0~1.1	1.6~2.3
	% of Occurrence	0%	0~20%	7~11%

**Table 5.12** Absolute contributions of PM<sub>2.5</sub> (in unit  $\mu\text{g}/\text{m}^3$ ) and contribution percentages of local, PRD and non-PRD sources in the high mass concentration hours [Scenario 1] in 2006 at TW.

<b>Source Category</b>		<b>Local</b>	<b>PRD</b>	<b>Non-PRD</b>
Total	Absolute Contribution ( $\mu\text{g}/\text{m}^3$ )	13~14	1~33	51~82
	Contribution Percentage	14%	1~34%	52~85%
Summer	Absolute Contribution ( $\mu\text{g}/\text{m}^3$ )	10~11	10~83	10~82
	Contribution Percentage	10~11%	10~81%	9~80%
Winter	Absolute Contribution ( $\mu\text{g}/\text{m}^3$ )	22~25	0~39	37~79
	Contribution Percentage	22~25%	0~38%	36~78%

**Table 5.13** Average concentrations in PM<sub>2.5</sub> mass (unit: µg/m<sup>3</sup>) from local, PRD and non-PRD trajectory classes at TW in the high mass concentration hours [Scenario 2] in 2006 and in summer and winter.

<b>Trajectory Class</b>		<b>LOCAL</b>	<b>PRD</b>	<b>NON-PRD</b>
Summer	Average PM <sub>10</sub> concentration	16	74~82	67~78
	Average wind speed (m/s)	2.2	2.7~4.4	2.4
Winter	Average PM <sub>10</sub> concentration	23	0~103	102
	Average wind speed (m/s)	2.2	0~1.1	1.6~2.3

**Table 5.14** Absolute contributions of PM<sub>2.5</sub> (in unit µg/m<sup>3</sup>) and contribution percentages of local, PRD and non-PRD sources in the high mass concentration hours [Scenario 2] in 2006 at TW.

<b>Source Category</b>		<b>Local</b>	<b>PRD</b>	<b>Non-PRD</b>
Summer	Absolute Contribution (µg/m <sup>3</sup> )	10	9~51	17~58
	Contribution Percentage	12~13%	12~66%	21~75%
Winter	Absolute Contribution (µg/m <sup>3</sup> )	22~26	0~38	39~80
	Contribution Percentage	22~25%	0~37%	38~78%