Source apportionment of particulate matter and selected volatile organic compounds with multiple time resolution data

Cheng-Pin Kuo [a,1], Ho-Tang Liao [a,1], Charles C.-K. Chou [b], Chang-Fu Wu [a,c,d,*]

[a] Institute of Occupational Medicine and Industrial Hygiene, National Taiwan University, Taipei 100, Taiwan
[b] Research Center for Environmental Changes, Academia Sinica, Taipei 115, Taiwan
[c] Department of Public Health, National Taiwan University, Taipei 100, Taiwan
[d] Institute of Environmental Health, National Taiwan University, Taipei 100, Taiwan

HIGHLIGHTS

- Hourly VOC and 12-h $PM_{2.5}$ speciation data were combined in a receptor model.
- The Multilinear Engine (ME-2) program was applied, and five factors were retrieved.
- The source contribution distribution in $PM_{2.5}$ differed from that in VOCs.
- Sources having similar VOC profiles were separated by including $PM_{2.5}$ data.

ARTICLE INFO

ABSTRACT

Fine particulate matter ($PM_{2.5}$) and volatile organic compounds (VOCs) co-exist in ambient air and contribute to adverse health effects in human populations. This study was conducted to demonstrate the feasibility of utilizing a composite data set which included $PM_{2.5}$ and VOC data with multiple time resolutions for source apportionment. Hourly VOC and 12-h $PM_{2.5}$ speciation data were combined into an improved source apportionment model to quantify different pollutant source contributions to $PM_{2.5}$ and VOC mixtures. Five factors were retrieved, including vehicle 1, vehicle 2, industrial processing, transported regional, and secondary pollution sources. The largest contributors were vehicular emissions for VOCs (62%) and $PM_{2.5}$ (35%). Nonetheless, transported regional and secondary pollution sources accounted for a noteworthy portion of $PM_{2.5}$ (27% and 25%, respectively) relative to VOCs (8% and 5%, respectively). Additional sensitivity analyses showed that excluding the $PM_{2.5}$ data or reducing the associated temporal resolution (12-h to 24-h) retrieved fewer source factors and increased the errors of source contribution estimates.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Volatile organic compounds (VOCs) and fine particulate matter ($PM_{2.5}$) are considered hazardous air pollutants (HAPs), and may cause adverse health effects (Kampa and Castanas, 2008; Seagrave et al., 2006; Stanek et al., 2011a, 2011b; Wichmann et al., 2009). They are emitted to the environment from a variety of natural and anthropogenic sources, such as biogenic emission, industrial activities, and vehicle exhaust (Brown et al., 2007; Buzcu and Fraser, 2006; Kampa and Castanas, 2008). The identification of VOC and $PM_{2.5}$ sources and establishing estimates of quantitative source contributions to environments are crucial in developing effective control strategies. One approach to obtain this information is the application of receptor-based source apportionment models, such as Chemical Mass Balance (CMB) (Cooper and Watson, 1980), Absolute Principal Component Analysis (APCA) (Thurston and Spengler, 1985), and Positive Matrix Factorization (PMF) (Paatero and Tapper, 1994). Compared to the CMB model, the multivariate models require less previous knowledge regarding the emission source profiles. Therefore, when complete emission profiles are not available, the PMF model (one of the advanced multivariate models that simultaneously retrieve source contributions and profiles from measurements with non-negativity constraint) can be applied (Paatero and Tapper, 1994; Viana et al., 2008a, 2008b).

Previous studies demonstrated the advantages of using $PM_{2.5}$ and VOC composite data in multivariate receptor models. Wu et al. (2007) reported that more accurate source contribution estimates and a more complete set of sources were obtained with the addition of VOC data to $PM_{2.5}$ measurements. Chan et al. (2011) concluded that through the

* Corresponding author at: Institute of Occupational Medicine and Industrial Hygiene, National Taiwan University, No.17 Xu-Zhou Road, Taipei 100, Taiwan. Tel./fax: +886 2 3366 8096.
    E-mail address: changfu@ntu.edu.tw (C.-F. Wu).
[1] These authors contributed equally to this work.

integration of multiple type composition data, small data sets can be applied and generate similar results to those reported in studies with larger data sets.

One issue of applying multiple type composition data in receptor modeling is having species with different temporal resolutions due to varied sampling strategies or analytical methods. Conventional receptor models require averaging high-resolution to match low-resolution data, resulting in decreased sample size and loss of information in species with high temporal resolution (Lioy et al., 1989). Consequently, Zhou et al. (2004) developed an improved source apportionment model to incorporate data with various temporal resolutions.

Previous studies incorporating different temporal resolution data focused primarily on PM components (Ogulei et al., 2005; Zhou et al., 2004). However to date, few studies have investigated the feasibility of utilizing combined $PM_{2.5}$ and VOC data sets with different temporal resolution measurements. Therefore, we conducted a simulation study using synthetic data sets that included VOCs and $PM_{2.5}$ from five sources under different time resolutions (Liao et al., 2013). Results showed high correlations and comparable percentage contributions between input data and modeling results, which demonstrated the applicability and validity of the improved model.

In the present study, an extension of our former investigation (Liao et al., 2013), we evaluated performance of the improved source apportionment model with field data. The composite data set included 12-h $PM_{2.5}$ component data and hourly VOC data. Different data structures were constructed to examine the effects of combining speciation data under multiple time resolutions.

2. Material and methods

2.1. Data collection

Field data was collected at the Tucheng Photochemical Assessment Monitoring Station (PAMS; 23°43′03″N, 120°12′10″E) of the Taiwan Environmental Protection Administration (TEPA). The Tucheng district is located in southwest Taipei Basin, Taiwan. The PAMS site is situated on top of a 5.5 m tall high school building; 1.9 km from Tucheng Industrial Park, and 3.6 km from Shulin Industrial Park. Two major highways are located to the south and northwest of the monitoring site (Fig. 1).

$PM_{2.5}$ and VOCs were sampled at the monitoring site. $PM_{2.5}$ samples were collected on 37 mm Teflon filters using Harvard Impactors operated at a flow rate of 10 L/min on a 12-h cycle (7.2 m^3 per sample) from 17 May to 5 June 2012. $PM_{2.5}$ elemental analysis was performed by an energy-dispersive X-ray fluorescence spectrometer (ED-XRF, Epsilon 5, PANalytical), which has a 600 W power X-ray tube, following the protocol adopted from the US EPA Chemical Speciation Network (RTI International, 2009). Ionic analysis was performed by Ion Chromatography (IC, DX-100, Dionex) with a conductivity detector for water-soluble ionic species extracted by deionized water (Chen et al., 2010; Cheng et al., 2009; Lung et al., 2004). A total of 16 elements (Mg, Al, Si, S, K, Ca, Ti, V, Cr, Mn, Fe, Ni, Cu, Zn, Sb, and Pb) and 10 ions (Na^+, NH_4^+, K^+, Mg^{2+}, Ca^{2+}, Cl^-, NO_2^-, NO_3^-, PO_4^{3-}, and SO_4^{2-}) were analyzed. VOC monitoring data were obtained from PAMS. They were monitored by a commercial auto-GC system (TM-TD1/Clarus500GC, PerkinElmer) consisting of a cryogen free sorbent trap, dual columns, and dual flame ionization detectors (FID). Ambient air is drawn into the trap at a flow rate of 20 mL/min for 45 min (900 cm^3 per measurement; Tsai et al., 2008; Yang et al., 2005). Hourly data were available for 56 VOC species, including 29 alkanes, 10 alkenes, 1 alkyne, and 16 aromatics. Method detection limits (MDLs) were determined from the standard deviations of the blank values or certified reference materials.

2.2. Receptor modeling

Implementation of the source apportionment model is detailed in our previous study (Liao et al., 2013). A brief summary of the model applied to this study is as follows: the PMF model was based on the mass balance equation for multiple time resolution data (Zhou et al., 2004):

$$x_{sj} = \frac{1}{t_{s2}-t_{s1}+1} \sum_{p=1}^{P} \left[\left(\sum_{i=t_{s1}}^{t_{s2}} g_{ip} \right) f_{pj} \right] + e_{sj} \qquad (1)$$

where x_{sj} represents the concentration of species j measured in sample s (i.e. hourly VOCs or 12-h $PM_{2.5}$ species), and t_{s1} and t_{s2} are start and end times, respectively, as determined by the number of time units. In this study, the time unit was set to 1 h based on the VOC sampling interval. In Eq. (1), g_{ip} is the contribution of source p during the time units for sample s, f_{pj} is the mass fraction of species j in source profile p, and e_{sj}
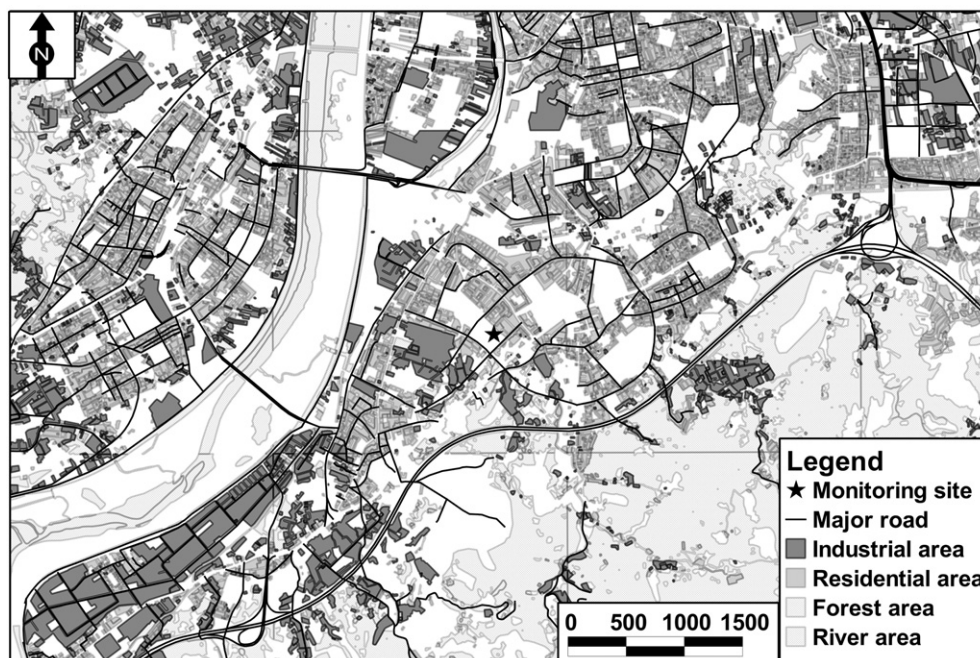

Fig. 1. Map of the study region, including location of the monitoring site and potential environmental pollution sources.