Annual variations of carbonaceous $PM_{2.5}$ in Malaysia: Influence by Indonesian peatland fires

Y. Fujii[1,2], S. Tohno[1], N. Amil[3,4], M.T. Latif[3,5], M. Oda[1], J. Matsumoto[6] and A. Mizohata[6]

[1]{Department of Socio-Environmental Energy Science, Kyoto University, Kyoto, Japan}

[2]{Japan Society for the Promotion of Science, Tokyo, Japan}

[3]{School of Environmental and Natural Resource Sciences, Universiti Kebangsaan Malaysia, Bangi, Malaysia}

[4]{School of Industrial Technology, Universiti Sains Malaysia, Penang, Malaysia}

[5]{Institute for Environment and Development, Universiti Kebangsaan Malaysia, Bangi, Malaysia}

[6]{Research Organization for University-Community Collaborations, Osaka Prefecture University, Sakai, Japan}

Correspondence to: Y. Fujii (fujii.yusuke.86n@st.kyoto-u.ac.jp)

Abstract

In this study, we quantified carbonaceous $PM_{2.5}$ in Malaysia through annual observations of $PM_{2.5}$, focusing on organic compounds derived from biomass burning. We determined organic carbon (OC), elemental carbon (EC) and concentrations of solvent-extractable organic compounds (biomarkers derived from biomass burning sources and *n*-alkanes). We observed seasonal variations in the concentrations of pyrolyzed OC (OP), levoglucosan (LG), mannosan (MN), galactosan, syringaldehyde, vanillic acid (VA) and cholesterol. The average concentrations of OP, LG, MN, galactosan, VA and cholesterol were higher during the southwest monsoon season (June–September) than during the northeast monsoon season (December–March), and these differences were statistically significant. Conversely, the syringaldehyde concentration during the southwest monsoon season was lower. The $PM_{2.5}$ OP/OC4 mass ratio allowed distinguishing the seven samples, which have been affected by the

Indonesian peatland fires (IPFs). In addition, we observed significant differences in the concentrations between the IPF and other samples of many chemical species. Thus, the chemical characteristics of $PM_{2.5}$ in Malaysia appeared to be significantly influenced by IPFs during the southwest monsoon season. Furthermore, we evaluated two indicators, the vanillic acid/syringic acid (VA/SA) and LG/MN mass ratios, which have been suggested as indicators of IPFs. The LG/MN mass ratio ranged from 14 to 22 in the IPF samples and from 11 to 31 in the other samples. Thus, the respective variation ranges partially overlapped. Consequently, this ratio did not satisfactorily reflect the effects of IPFs in Malaysia. In contrast, the VA/SA mass ratio may serve as a good indicator, since it significantly differed between the IPF and other samples. However, the OP/OC4 mass ratio provided more remarkable differences than the VA/SA mass ratio, offering an even better indicator. Finally, we extracted biomass burning emissions' sources such as IPF, softwood/hardwood burning and meat cooking through varimax-rotated principal component analysis.

1 Introduction

Peatland is a terrestrial wetland ecosystem where organic matter production exceeds its decomposition, resulting in net accumulation (Page et al., 2006). Indonesia has the third largest peatland area and the largest tropical peatland area in the world (270,000 km^2; Joosten, 2010). Peatland fires occur predominantly in the Sumatra and Kalimantan Islands, Indonesia (Fujii et al., 2014; Page et al., 2002) during the dry season (June–September) mostly due to illegal human activities (Harrison et al., 2009). Because peatland fires are usually underground fires, they are extremely difficult to extinguish. The resulting haze comprises gasses and particulates that are emitted because of biomass burning. It extends beyond Indonesia to the neighbouring countries including Malaysia and Singapore (Betha et al., 2014; Engling et al., 2014; Fujii et al., 2015b; He et al., 2010; See et al., 2006, 2007), limiting visibility and causing health problems to the local population (Emmanuel, 2000; Othman et al., 2014; Pavagadhi, et al., 2013; Sahani et al., 2014). Therefore, Indonesian peatland fires (IPFs) have been recognised as an international problem (Yong and Peh, 2014; Varkkey, 2014).

The main constituent of particulates derived from biomass burning is $PM_{2.5}$ defined as particles having aerodynamic diameters below 2.5 μm, which has been associated with serious health problems (Federal Register, 2006; Schlesinger, 2007). These particulates are primarily composed of organic carbon (OC), which constitutes 50%–60% of the total particle mass (Reid

et al., 2005). At present, there are only four papers concerning the $PM_{2.5}$ chemical speciation resulting from IPFs; these papers are based on surface-recorded source-dominated data (Betha et al., 2013; Fujii et al., 2014, 2015a; See et al., 2007). Organic matter is the main component of $PM_{2.5}$ from IPFs as well as from biomass burning in general (Fujii et al., 2014; See et al., 2007). The primary organic compounds such as cellulose and lignin pyrolysis products have been quantified and potential IPF indicators at the receptor site have been suggested by Fujii et al. (2015a). Additional compounds have been discussed by Betha et al. (2013) (metals) and See et al. (2007) (water-soluble ions, metals and polycyclic aromatic hydrocarbons).

Several studies exist on the chemical characteristics of haze ambient particulates, which have been potentially affected by IPFs in Malaysia and Singapore (e.g., Abas et al., 2004a, b; Betha et al., 2014; Engling et al., 2014; Fang et al., 1999; Fujii et al., 2015b; He et al., 2010; Keywood et al., 2003; Narukawa et al., 1999; Okuda et al., 2002; See et al., 2006; Yang et al., 2013). In most cases, the field observation periods were short. Even when long-term observations have been obtained, however, only typical chemical species such as ions and metals have been analysed. Nevertheless, organic compounds significantly contribute to the IPF aerosols (Fujii et al., 2014). In Malaysia especially, there are no available quantitative data regarding variations of several organic compound concentrations based on long-term observations of $PM_{2.5}$.

The three major sources of air pollution in Malaysia are mobile, stationary and open burning sources including the burning of solid wastes and forest fires (Afroz, et al., 2003). The annual burned biomass in Malaysia has been estimated to be 23 Tg on average (Streets et al., 2003). Therefore, it is necessary to distinguish the effects of IPFs from those of other sources, particularly local biomass burning. Fujii et al. (2015b) reported the total suspended particulate matter (TSP) concentrations in the different carbon fractions (OC1, OC2, OC3, OC4 and pyrolysed OC (OP)) defined by the IMPROVE_A protocol (Chow et al., 2007) in Malaysia during the haze periods affected by IPFs. They proposed the OP/OC4 mass ratio as a useful indicator of transboundary haze pollution from IPFs at receptor sites even in light haze; the ratio during the haze periods were higher (>4) than during the non-haze periods (<2).

In the present study, the carbonaceous $PM_{2.5}$ components are quantitatively characterised using annual $PM_{2.5}$ observations in Malaysia, with special regard to the organic compounds resulting from biomass burning. Furthermore, the OP/OC4 mass ratio is used as an indicator to investigate the effects of IPFs on carbonaceous $PM_{2.5}$ species in this area. In addition, other indicators that potentially record the effects of IPFs are investigated. Finally, possible

carbonaceous $PM_{2.5}$ sources are suggested using varimax-rotated principal component analysis (PCA).

2 Experimental method

2.1 Sampling site and period

The sampling site is the Malaysian Meteorological Department (MMD) located in Petaling Jaya (PJ), Selangor, Malaysia (~100 m above sea level, 3° 06′ 09″ N, 101° 38′ 41″ E). Eighty-one $PM_{2.5}$ samples were collected on the roof of the MMD's main building (eight stories) from August 2011 to July 2012. A detailed description of the sampling site has been provided by Jamhari et al. (2014). In brief, PJ is located in an industrial area (Department of Environment, 2014) ~10 km from Kuala Lumpur. It is predominantly residential and industrial with high-density road traffic.

2.2 Sample collection and analysis

$PM_{2.5}$ samples were continuously collected with a Tisch high-volume air sampler (model TE-3070V-2.5-BL) on a quartz-fibre filter for 24 h at a flow rate of 1.13 m^3 min^{-1}. Before sampling, the quartz-fibre filters were heated to 500 °C for 3 h. After sampling, OC, elemental carbon (EC) and solvent-extractable organic compound (SEOC; biomarkers derived from biomass burning sources and *n*-alkanes) were measured.

The carbonaceous content were quantified using a DRI model 2001 OC/EC carbon analyser, which employs the thermal optical-reflectance method following the IMPROVE_A protocol. As shown in our former report (Fujii et al., 2014), the IMPROVE_A temperature protocol defines temperature plateaus for thermally-derived carbon fractions as follows: 140 °C for OC1, 280 °C for OC2, 480 °C for OC3 and 580 °C for OC4 in helium (He) carrier gas; 580 °C for EC1, 740 °C for EC2 and 840 °C for EC3 in a mixture of 98% He and 2% oxygen (O_2) carrier gas. OC and EC are calculated from the eight carbon fractions as follows:

$$\mathrm{OC} = \mathrm{OC1} + \mathrm{OC2} + \mathrm{OC3} + \mathrm{OC4} + \mathrm{OP}, \quad (1)$$

$$\mathrm{EC} = \mathrm{EC1} + \mathrm{EC2} + \mathrm{EC3} - \mathrm{OP}, \quad (2)$$

where OP is defined as the carbon content measured after the introduction of O_2 until reflectance returns to its initial value at the start of analysis. Blank corrections were performed on the OC and EC data by subtracting the blank filter value from the loaded filter values.

SEOC obtained from the quartz-fibre filters were quantified by gas chromatography mass spectrometry (GC/MS). Biomarker organic compound speciation was accomplished following the procedures reported previously (Fujii et al., 2015a, b). To quantify *n*-alkanes, aliquots from the quartz-fibre filter were spiked with internal standards of eicosane-d_{42} and triacontane-d_{62} before extraction. Each spiked filter was extracted by ultrasonic agitation for 2 × 20 min periods using 8 mL hexane (Kanto Chemical, purity >96.0%). The combined extracts were filtered through a polytetrafluoroethylene syringe filter (pore size 0.45 μm), dried completely under a gentle stream of nitrogen gas and re-dissolved to 0.1 mL in hexane. Before the GC/MS analysis, ~1.05 μg of tetracosane-d_{50} dissolved in 50 μL of hexane was added as a second internal standard. The *n*-alkanes values were reported in carbon numbers, ranging from 22 to 33 (C_{22}–C_{33}). The extract samples were analysed on a Shimadzu GC/MS system (GCMS-QP2010-Plus, Shimadzu) equipped with a 30 m HP-5MS column (0.25 μm film thickness, 0.25 mm ID). The carrier gas was helium (purity >99.9%) at a pressure of 73.0 kPa (37.2 cm s^{-1} at 100 °C). The GC oven temperature program was as follows: isothermal at 100 °C for 5 min, 100–300 °C at 10 °C min^{-1} and then 300 °C for 20 min. The injection port and transfer line were maintained at 300 °C. The data for quantitative analysis were acquired in the electron impact mode (70 eV). The mass spectrometer was operated under the selected ion-monitoring scanning mode, and the monitored ions for the quantification of *n*-alkanes were 85 m/z. The monitored ions corresponding to the internal standards were 66 m/z. The recovery ratios for known amounts of *n*-alkane standards (1 μg addition) on the quartz-fibre filters range from 73 to 110% (mean ± standard deviation: 94 ± 6.3%). Blank corrections were performed on the biomarker and *n*-alkane data by subtracting the blank filter value from the loaded filter values.

2.3 Source apportionment method

Varimax-rotated PCA was used to identify the possible carbonaceous $PM_{2.5}$ sources at PJ. The following two datasets were considered: (i) PJ_A data, which includes 25 variables (all quantified compounds) and 81 samples (all samples), and (ii) PJ_S data, which includes 25 variables and 65 samples (excluded are the samples acquired on September 2011 and June 2012,

which are influenced by IPFs as shown in the Section 3). PCA results with these datasets are expected to show IPF effects on other sources. It has been suggested that the minimum number of samples (n) for factor analysis should satisfy the following condition (Henry et al., 1984; Karar and Gupta, 2007):

$$n > 30 + \frac{V + 3}{2}, \quad (3)$$

where V represents the number of variables. Both datasets satisfy this condition.

Varimax-rotated PCA followed the procedure proposed by Karar and Gupta (2007) and was accomplished with the R-software (http://www.R-project.org). The eigenvalues correspond to the number of factors, which was selected to ensure that the cumulative variance contribution rate is greater than 80%.

3 Results and discussion

3.1 Air quality and monthly hotspot data

Figure 1 presents the daily variability of the Malaysian Air Pollutant Index (MAPI) and visibility during the sampling periods. The MAPI data were obtained from the Department of Environment Ministry of Natural Resources and Environment website (http:// apims.doe.gov.my/apims/hourly2.php). Hourly visibility data (7:00–17:00) provided by the MMD were used to produce the daily variation in visibility after removing the hourly data corresponding to periods of rainfall. The MAPI values of 0–50, 51–100, 101–200, 201–300 and >300 correspond to good, moderate, unhealthy, very unhealthy and hazardous air quality conditions (Department of Environment, 2014; Fujii et al., 2015b). Good MAPI levels dominate the sampling periods except August 2011, September 2011 and June 2012. On the other hand, moderate air quality is observed in August 2011, September 2011 and June 2012. The two MAPI values for 15 and 16 June 2012 indicate unhealthy air quality conditions. The average visibility during these two sampling periods (Fig. 1) was below 2.7 km, corresponding to extremely low visibility compared with other intervals.

Figure 2 presents the monthly hotspot counts in the Sumatra Island detected by the NOAA-18 satellite (Indofire). During the southwest monsoon season on September 2011 and June 2012, hotspots exceeded 3,000 on several occasions. The hotspot counts in September 2011 and June

2012 mainly derived from the South Sumatra (60% of the hotspot counts) and the Riau (42%) provinces, respectively. The sampling sites are dominantly downwind regions in the Sumatra Island during the southwest monsoon season. Thus, some samples have probably been affected by IPFs. The three-day backward air trajectories for the sampling periods (Fig. S1) support this conclusion.

3.2 $PM_{2.5}$ chemical characteristics and seasonal variations

3.2.1 OC and EC

The annual concentrations of OC and EC are 7.0 ± 5.4 and 3.1 ± 1.1 μgC m^{-3}, respectively. The OC and EC concentrations' statistical results for each monsoon season appear in Table 1. The average OC concentration during the southwest monsoon season (June–September) is higher than that during other seasons. In particular, an extremely high OC concentration (>25 μg m^{-3}) is observed on 12 September 2011 and on 15 and 16 June 2012. There is no statistically significant difference in the EC concentration between the southwest and northeast (December–March) monsoon seasons according to the two-sided Wilcoxon rank sum test (p-value: p >0.05) with R-software. In Bangi (~30 km southeast of the sampling site), the OC concentration was 11 ± 3.2 μgC m^{-3} in September 2013 (Fujii et al., 2015c), in good agreement with the present results for the southwest monsoon season. The OC/EC mass ratios during the southwest monsoon, post-monsoon (October–November), northeast monsoon and pre-monsoon (April–May) season range among 1.2–6.5, 1.4–2.4, 0.99–3.0 and 1.2–2.3, respectively. A high OC/EC mass ratio value (>4) is found only for some samples collected on September 2011 and June 2012. These values have probably been affected by biomass burning, because aerosols emitted from biomass burning usually present higher OC/EC mass ratios (Cong et al., 2015).

The daily variations of the OC fractions' mass concentrations during the sampling periods are presented in Fig. 3. The annual concentrations of OC1, OC2, OC3, OC4 and OP are 0.51 ± 0.80, 1.9 ± 1.1, 2.3 ± 1.4, 1.2 ± 0.36 and 1.1 ± 2.2 μg m^{-3}, respectively. Statistically significant differences among the OP concentrations during the southwest and northeast monsoon seasons are observed according to the two-sided Wilcoxon rank sum test (p <0.001). In particular, high OP concentrations are clearly observed in September 2011 and June 2012, in addition to the higher OC/EC mass ratios described above. Fujii et al. (2015b) supported that the enhanced OP concentrations in TSP, which are observed in Malaysia during the haze periods, are affected by

the IPFs. The enhanced OP concentrations in $PM_{2.5}$ during the southwest monsoon season, which are observed in the present study, are also probably affected by IPFs from the Sumatra Island. The increased number of hotspots recorded (Fig. 2) and backward air trajectories (Fig. S1) further support this conclusion.

3.2.2 Biomarkers

Ten biomarkers are identified in this study, which have been suggested as indicators of biomass burning processes such as wood burning and meat cooking. The annual concentrations of levoglucosan (LG), mannosan (MN), galactosan, *p*-hydroxybenzoic acid, vanillic acid (VA) and syringic acid (SA) are 86 ± 95, 4.8 ± 5.7, 1.2 ± 1.6, 1.1 ± 1.3, 0.19 ± 0.28 and 0.25 ± 0.28 ng m^{-3}, respectively; notably, they exhibit great variability. The annual concentrations of vanillin, syringaldehyde, dehydroabietic acid and cholesterol are 1.2 ± 0.80, 0.51 ± 0.42, 1.3 ± 1.0 and 1.3 ± 0.72 ng m^{-3}, respectively. The biomarker statistical results for each monsoon season are listed in Table 1.

LG is a specific indicator for cellulose burning emissions and generally formed during cellulose pyrolysis at temperatures above 300 °C (Fujii et al., 2015b; Lin et al., 2010; Shafizadeh, 1984; Simoneit et al., 1999). The MN and galactosan are derived from hemicellulose pyrolysis products; they can also be used as tracers of biomass burning besides LG (e.g., Engling et al., 2014; Fujii et al., 2014, 2015b; Zhu et al., 2015). Statistically significant differences are observed among the concentrations of LG, MN and galactosan obtained during the southwest and northeast monsoon seasons on the basis of the two-sided Wilcoxon rank sum test (p <0.001); high concentrations of these compounds are mostly observed during the southwest monsoon season (especially September 2011 and June 2012; Fig. S2). In Singapore, Engling et al. (2014) suggested that the enhanced concentrations of these compounds during the haze periods were due to the IPFs during the southwest monsoon season. Thus, the presently observed enhanced concentrations of these compounds may also be attributed to the IPFs.

In a previous report, $PM_{2.5}$ lignin unit-originating compounds in samples collected at IPF source were quantified (Fujii et al., 2015a). Lignin is an aromatic polymer consisting of phenylpropane units linked through many ether and C–C linkages. Its aromatic structure varies depending on the species; softwood lignins exclusively contain guaiacyl (G) types, hardwood lignins include both G and syringyl (S) types, whereas herbaceous plants include G, S and *p*-hydroxyphenyl

(H) types (Fujii et al., 2015a, b). The composition of these aromatic nuclei within the lignin pyrolysis products resulting from biomass burning may be useful in identifying the biomass type (Fujii et al., 2015a; Simoneit et al., 1993). In the present study, vanillin and VA (compounds derived from G units), syringaldehyde and SA (compounds derived from S units) as well as and *p*-hydroxybenzoic acid (compounds derived from H units or the secondary decomposition of G and S units) (Fujii et al., 2015b) have been quantified. There are significant differences between the concentrations of syringaldehyde and VA derived from lignin pyrolysis during the southwest and northeast monsoon seasons on the basis of the two-sided Wilcoxon rank sum test (p <0.001), corresponding to seasonal variations. The average VA concentration during the southwest monsoon season is 5.3 times greater than that during the northeast monsoon season. In contrast, the average concentration of syringaldehyde during the northeast monsoon season is 2.6 times greater than that during the southwest monsoon season. This may be due to the transboundary pollution by prevailing winds from the Chinese region including Thailand and Vietnam during the northeast monsoon season (Fig. S1; Khan et al., 2015).

Dehydroabietic acid and cholesterol are quantified as indicators of softwood burning and meat cooking, respectively (Fujii et al., 2015b; Lin et al., 2010). The two-sided Wilcoxon rank sum test indicates that the difference between the cholesterol concentration during the southwest and northeast monsoon seasons is statistically significant (p <0.001). The dehydroabietic acid and cholesterol concentrations recorded in the interval between June and July 2014 in Bangi, which is located ~30 km southeast of the sampling site, range between 2.6–8.7 and 1.5–5.7 ng m^{-3}, respectively (Fujii et al., 2015b). The PJ industrial area's concentrations of these compounds are lower than those in the Bangi suburban area owing to the decreased impact of softwood burning and meat cooking in PJ.

3.2.3 *N*-alkanes

The total annual concentration of *n*-alkanes is 79 ± 63 ng m^{-3}. The total *n*-alkanes concentration during the southwest monsoon, post-monsoon, northeast monsoon and pre-monsoon season is 110 ± 93, 57 ± 20, 67 ± 18 and 55 ± 41 ng m^{-3}, respectively. The highest concentration is observed during the southwest monsoon season. Figure 4 illustrates the molecular distribution of *n*-alkanes during the southwest and northeast monsoon seasons. There are no significant differences among the concentrations of C_{22}–C_{26}, C_{29}, C_{30} and C_{32} in the two seasons (p >0.05). High concentrations of >C_{24} are mainly observed in September 2011 and June 2012 when many

hotspots are detected in the Sumatra Island (Fig. 2). Fujii et al. (2015a) suggested that IPFs increase the C_{27}, C_{28} and C_{29} concentrations in $PM_{2.5}$ at the receptor site relative to other sources such as vehicle and biomass burning. Thus, the enhanced *n*-alkanes concentrations in $PM_{2.5}$ during the southwest monsoon season may be mainly attributed to IPFs.

The carbon number maximum (C_{max}) in *n*-alkanes during the southwest and northeast monsoon seasons is C_{27} (in 83% of the samples) and C_{26} (89%), respectively (Fig. 5). Reported C_{max} values range from 27 to 33, characteristic of biogenic sources (higher plant-wax), whereas lower C_{max} values may indicate major petrogenic input (Abas et al., 2004a; Gogou et al., 1996; He et al., 2010). The C_{max} during the southwest monsoon season (C_{27}) suggests primarily biogenic sources and is in perfect agreement with the measured value for the IPF source (Fujii et al., 2015b).

The carbon preference index (CPI) has been widely used to roughly estimate the effects of anthropogenic or biogenic sources (e.g., Bray and Evans, 1961; Chen et al., 2014; He et al., 2010; Yamamoto et al., 2013). The CPI values are calculated by the following equation based on the suggestion by Bray and Evans (1961).

$$\mathrm{CPI} = 0.5 \times \left(\frac{C_{25} + C_{27} + C_{29} + C_{31}}{C_{26} + C_{28} + C_{30} + C_{32}} + \frac{C_{25} + C_{27} + C_{29} + C_{31}}{C_{24} + C_{26} + C_{28} + C_{30}}\right) \quad (4)$$

The CPI values are generally high (CPI > 5) when there is no serious input from fossil fuel hydrocarbons (CPI = 1) (Yamamoto et al., 2013, and references therein). The CPI values during the southwest and northeast monsoon seasons are 1.3 ± 0.12 and 1.0 ± 0.14, respectively; these values are close to one for both seasons, indicating an anthropogenic *n*-alkane source. Thus, the CPI value is not susceptible to IPF influence, since the CPI value at IPF source is 1.6 ± 0.13 (Fujii et al., 2015a), which is not high. Consequently, the CPI cannot be used to identify IPFs sources at a receptor site.

3.3 Indonesian peatland fire effect

The hotspot data and backward air trajectories suggest that IPFs strongly modify many chemical species concentrations mostly during the southwest monsoon season. However, IPFs do not always occur during the southwest monsoon season. Therefore, significant differences in some chemical species concentrations among samples affected by IPF and others should be observed. To distinguish IPF samples from other samples obtained during the southwest monsoon season,

the OP/OC4 mass ratio is used, which is a useful indicator for IPF (Fujii et al., 2015b). The ratio value is >4 for seven samples (11–13 September 2011 and 14–17 June 2012); these samples are regarded as the IPF samples. The OP/OC4 mass ratio for the IPF and other samples is 7.4 ± 3.4 and 0.44 ± 0.49, respectively, exhibiting significant differences among them according to the two-sided Wilcoxon rank sum test (p <0.001). Figure 6 shows the p-values used to determine the statistical significance in a hypothesis test of the differences between the IPF and other samples for all the quantified species. Significant differences (p <0.001) are recorded for many chemical species. Thus, the chemical characteristics of $PM_{2.5}$ in Malaysia are significantly influenced by IPFs.

Furthermore, the VA/SA and LG/MN mass ratios in the IPF source are investigated as potential indicators, as suggested in previous studies (Fujii et al., 2014, 2015a). The VA/SA mass ratio for IPF and other samples is 1.7 ± 0.36 and 0.59 ± 0.27, respectively, providing a good indicator (p <0.001). Although the VA/SA mass ratio at the IPF source is 1.1 ± 0.16 (Fujii et al., 2015a), the ratios for IPF samples are higher. Opsahl and Benner (1998) reported photochemical reactivity of VA and SA in the Mississippi River water. They demonstrated that the early degradation of SA in the water is mostly due to its higher photochemical reactivity compared with VA. Even though there are no reports of such degradations in air, SA is considered to be less stable than VA in air as well as in water, which leads to an increased VA/SA ratio after long-range transportation. On the other hand, the LG/MN mass ratio for the IPF and other samples ranges from 14 to 22 and 11 to 31, respectively (Fig. S3). Therefore, the LG/MN mass ratio is inappropriate to extract the effects of IPF in Malaysia, because its value's ranges in the IPF and other samples partially overlap.

The daily variability of the C_{27} and LG concentration as well as the VA/SA and OP/OC4 mass ratios are presented in Fig. 7; similar trends are observed in all cases. However, the concentrations of LG, MN and galacotsan (Fig. S2) increase abruptly on 10 August 2011, although this sample is not categorised as an IPF sample. We hypothesised that this increase results from local biomass burning, since LG emissions are produced by several different biomass burning sources (Oros and Simoneit, 2001a,b; Oros et al., 2006). Therefore, LG levels are not directly indicative of the IPF contribution in Malaysia; instead, C_{27} may be a useful indicator (Fig. 7). Although the VA/SA mass ratio can be used as an IPF indicator, as we mentioned before, the OP/OC4 mass ratio highlights the differences between the IPF and other samples better than the VA/SA mass ratio (Fig. 7).

3.4 Carbonaceous $PM_{2.5}$ contributions

The possible sources of carbonaceous $PM_{2.5}$ are investigated through varimax-rotated PCA of the PJ_A and PJ_S datasets. Over 80% of the cumulative variance in the PJ_A and PJ_S datasets is explained by three and five factors, respectively (Table 2). For the PJ_A data (Table 2a), the total variance explained by the three factors is 80%. Factor A1, which explains 60% of the variance, is heavily loaded (loading factor: >0.65) with OC, LG, MN, galactosan, *p*-hydroxybenzoic acid, VA and C_{25}–C_{33}, which direct towards an IPF source. Factor A2, which corresponds to 12% of the variance, is heavily loaded with C_{22}–C_{24}, suggesting a petrogenic source (Abas et al., 2004a; Gogou et al., 1996; He et al., 2010). Factor A3, which explains 8.0% of the variance in the data set, is heavily loaded with SA and dehydroabietic acid, indicating mixed (softwood and hardwood) biomass burning sources. For the PJ_S dataset (Table 2b), the total variance explained by five factors is 82%. Factor S1 explains 43% of the data's variance and is heavily loaded with C_{27}–C_{33}, which suggests tire wear emission (Rogge et al., 1993). Factor S2 explains 19% of the variance and is heavily loaded with LG, MN, galactosan, VA and SA, which correspond to biomass burning source. Factor S3, which explains 11% of the variance, is heavily loaded with C_{22}–C_{26}, which indicate a petrogenic source, similar to factor A2. Although heavy loading with only syringaldehyde is found in factor S4 (5.0% of the variance), its source could not be identified. Finally, factor S5 explains 4.5% of the variance and is heavily loaded with EC and cholesterol, which are produced when cooking meat.

The differences of the factor loadings between PJ_A and PJ_S data are observed. For the PCA result of PJ_A dataset, the factors such as tire wear (factor S1) and cooking (factor S5) as shown in Table 2b are not extracted due to strong influence of IPFs. Although a petrogenic source is identified from both results, C_{25} and C_{26} are not heavily loaded for PJ_A dataset. This is also considered to be due to strong influence of IPFs.

Wahid et al. (2013) reported varimax-rotated PCA results on the distribution of inorganic ions within fine-mode aerosols (<1.5 μm) at Kuala Lumpur, which is close to the present study's sampling site (~10 km). They extracted three principal components from this analysis: (1) motor vehicles, (2) soil and earth's crust and (3) sea spray. Jamhari et al. (2014) applied varimax-rotated PCA on polycyclic aromatic hydrocarbon data in PM_{10} at Kuala Lumpur. They extracted two factors, which were attributed to (1) natural gas emission and coal combustion and (2)

vehicles and gasoline emissions. In the present study, only biomass burning could be identified as a factor through comparison with these previous analyses. Factors such as soil, sea spray and coal combustion could not be identified, because the key inorganic compounds produced from these sources were not determined.

4 Conclusions

Annual $PM_{2.5}$ observations in Malaysia have been conducted to quantitatively characterise carbonaceous $PM_{2.5}$, especially focusing on organic compounds derived from biomass burning for the first time. The main conclusions are summarised as follows:

Concentrations of OP, LG, MN, galactosan, syringaldehyde, VA and cholesterol exhibit seasonal variability. The average concentrations of OP, LG, MN, galactosan, VA and cholesterol during the southwest monsoon season are higher than those during the northeast monsoon season, and the differences are statistically significant. In contrast, the syringaldehyde concentration during the southwest monsoon season is lower.

Seven IPF samples are distinguished on the basis of the $PM_{2.5}$ OP/OC4 mass ratio. In addition, significant differences are observed for the concentrations of many chemical species between the IPF and other samples. Thus, the $PM_{2.5}$ chemical characteristics in Malaysia are clearly influenced by IPFs during the southwest monsoon season. Furthermore, two previously suggested indicators of IPF sources have been evaluated, the VA/SA and LG/MN mass ratio. The LG/MN mass ratio ranges from 14 to 22 in the IPF samples and from 11 to 31 in the other samples. The two ratio distributions partial overlap. Thus, the LG/MN mass ratio is not considered appropriate for extracting the effects of IPFs in Malaysia. In contrast, significant differences among the VA/SA mass ratios in the IPF and other samples suggest that it may serve as a good indicator. However, the OP/OC4 mass ratio differentiates the IPF samples better than VA/SA mass ratio. Consequently, the OP/OC4 mass ratio is proposed as a better indicator than the VA/SA mass ratio. Finally, varimax-rotated PCA enabled to discriminate biomass burning components such as IPFs, softwood/hardwood burning and meat cooking.

Acknowledgements

This study was supported by JSPS Kakenhi Grant Number (15H02589, 15J08153).

References

Abas, M.R., Oros, D.R., and Simoneit, B.R.T.: Biomass burning as the main source of organic aerosol particulate matter in Malaysia during haze episodes, Chemosphere, 55, 1089–1095, 2004a.

Abas, M.R.B., Rahman, N.A., Omar, N.Y.M.J., Maah, M.J., Samah, A.A., Oros, D.R., Otto, A., and Simoneit, B.R.T.: Organic composition of aerosol particulate matter during a haze episode in Kuala Lumpur, Malaysia, Atmos. Environ., 38, 4223–4241, 2004b.

Afroz, R., Hassan, M.N., and Ibrahim, N.A.: Review of air pollution and health impacts in Malaysia, Environ. Res., 92, 71–77, 2003.

Betha, R., Pradani, M., Lestari, P., Joshi, U.M., Reid, J.S., and Balasubramanian, R.: Chemical speciation of trace metals emitted from Indonesian peat fires for health risk assessment, Atmos. Res., 122, 571–578, 2013.

Betha, R., Behera, S.N., and Balasubramanian, R.: 2013 Southeast Asian Smoke Haze: Fractionation of Particulate-Bound Elements and Associated Health Risk, Environ. Sci. Technol., 48, 4327–4335, 2014.

Bray, E.E. and Evans, E.D.: Distribution of *n*-paraffins as a clue to recognition of source beds, Geochim. Cosmochim. Acta, 22, 2–15, 1961.

Chen, Y., Cao, J., Zhao, J., Xu, H., Arimoto, R., Wang, G., Han, Y., Shen, Z., and Li, G.: *n*-Alkanes and polycyclic aromatic hydrocarbons in total suspended particulates from the southeastern Tibetan Plateau: Concentrations, seasonal variations, and sources, Sic. Total Environ., 470–471, 9–18, 2014.

Chow, J.C., Watson, J.G., Chen, L.-W.,A., Chang, M.C.O., Robinson, N.F., Trimble, D., and Kohl, S.: The IMPROVE_A Temperature Protocol for Thermal/Optical Carbon Analysis: Maintaining Consistency with a Long-Term Database, J. Air & Waste Manage. Assoc., 57, 1014–1023, 2007.

Cong, Z., Kang, S., Kawamura, K., Liu, B., Wan, X., Wang, Z., Gao, S., and Fu, P.: Carbonaceous aerosols on the south edge of the Tibetan Plateau: concentrations, seasonality and sources, Atmos. Chem. Phys., 15, 1573–1584, 2015.

Department of Environment, Malaysia: Malaysia Environmental Quality Report 2013, Department of Environment, Ministry of Natural Resources and Environment, Malaysia, 2014.

Emmanuel, S.C.: Impact to lung health of haze from forest fires: the Singapore experience, Respirology, 5, 175–182, 2000.

Engling, G., He, J., Betha, R., and Balasubramanian, R.: Assessing the regional impact of indonesian biomass burning emissions based on organic molecular tracers and chemical mass balance modeling, Atmos. Chem. Phys., 14, 8043–8054, 2014.

Fang, M., Zheng, M., Wang, F., To, K.L., Jaafar, A.B., and Tong, S.L.: The solvent-extractable organic compounds in the Indonesia biomass burning aerosols – Characterization studies, Atmos. Environ., 33, 783–795, 1999.

Federal Register: National Ambient Air Quality Standards for Particulate Matter: Final Rule, In: 40 CFR Parts 50, 53, and 58, vol.62, US. EPA, Office of Air and Radiation, Office of Air Quality Planning and Standards, Research Triangle Park, NC, 2006.

Fujii, Y., Iriana, W., Oda, M., Puriwigati, A., Tohno, S., Lestari, P., Mizohata, A., and Huboyo, H.S.: Characteristics of carbonaceous aerosols emitted from peatland fire in Riau, Sumatra, Indonesia, Atmos. Environ., 87, 164–169, 2014.

Fujii, Y., Kawamoto, H., Tohno, S., Oda, M., Iriana, W., and Lestari, P.: Characteristics of carbonaceous aerosols emitted from peatland fire in Riau, Sumatra, Indonesia (2): Identification of organic compounds, Atmos. Environ., 110, 1–7, 2015a.

Fujii, Y., Mahmud, M., Oda, M., Tohno, S., and Mizohata, A.: A key indicator of transboundary particulate matter pollution derived from Indonesian peatland fires in Malaysia, Aerosol Air Qual. Res., 2015b, in press.

Fujii, Y., Mahmud, M., Tohno, S., Okuda, T., and Mizohata, A.: Characteristics of PM2.5 in Bangi, Selangor, Malaysia during the southwest monsoon season: Case study, Aerosol Air Qual. Res., 2015c, accepted.

Gogou, A., Stratigakis, N., Kanakidou, M., and Stephanou, E.G.: Organic aerosols in Eastern Mediterranean: components source reconciliation by using molecular markers and atmospheric back trajectories, Org. Geochem., 25, 79–96, 1996.

Harrison, M.E., Page, S.E., and Limin, S.H.: The global impact of Indonesian forest fires, Biologist, 56, 156–163, 2009.

He, J., Zielinska, B., and Balasubramanian, R.: Composition of semi-volatile organic compounds in the urban atmosphere of Singapore: Influence of biomass burning, Atmos. Chem. Phys., 10, 11401–11413, 2010.

Henry, R.C., Lewis, C.W., Hopke, P.K., and Williamson, H.J.: Review of receptor model fundamentals, Atmos. Environ., 18, 1507–1515, 1984.

Indofire. [online] [Accessed 17 July 2013]. Available: http://www.indofire.org/indofire/hotspot.

Jamhari, A.A., Sahani, M., Latif, T.M., Chan, K.M., Tan, H.S., Khan, M.F., and Tahir, N.M.: Concentration and source identification of polycyclic aromatic hydrocarbons (PAHs) in PM_{10} of urban, industrial and semi-urban areas in Malaysia, Atmos. Environ., 86, 16–27, 2014.

Joosten, H.: The Global Peatland CO_2 picture, Peatland Status and Drainage Associated Emissions in all Countries of the World, Wetlands International, Ede, The Netherlands, 2010.

Karar, K. and Gupta, A.K.: Source apportionment of PM_{10} at residential and industrial sites of an urban region of Kolkata, India, Atmos. Res., 84, 30–41, 2007.

Keywood, M.D., Ayers, G.P., Gras, J.L., Boers, R., and Leong, C.P.: Haze in the Klang Valley of Malaysia, Atmos. Chem. Phys., 3, 591–605, 2003.

Khan, M.F., Latif, M.T., Lim, C.H., Amil, N., Jaafar, S.A., Dominick, D., Nadzir, M.S.M., Sahani, M., and Tahir, N.M.: Seasonal effect and source apportionment of polycyclic aromatic hydrocarbons in $PM_{2.5}$, Atmos. Environ., 106, 178–190, 2015.

Lin, L., Lee, M.L., and Eatough, D.J.: Review of recent advances in detection of organic markers in fine particulate matter and their use for source, J. Air & Waste Manage., 60, 3–25, 2010.

Narukawa, M., Kawamura, K., Takeuchi, N., and Nakajima, T.: Distribution of dicarboxylic acids and carbon isotopic compositions in aerosols from 1997 Indonesian forest fires, Geophys. Res. Lett., 26, 3101–3104, 1999.

Okuda, T., Kumata, H., Zakaria, M.P., Naraoka, H., Ishiwatari, R., and Takada, H.: Source identification of Malaysian atmospheric polycyclic aromatic hydrocarbons neaby forest fires using molecular and isotopic compositions, Atmos. Environ., 36, 611–618, 2002.

Opsahl, S. and Benner, R.: Photochemical reacivity of dissolved lignin in river and ocean waters, Limnol. Oceanogr., 43, 1297–1304, 1998.

Oros, D.R. and Simoneit, B.R.T.: Identification and emission factors of molecular tracers in organic aerosols from biomass burning Part 1. Temperate climate conifers, Appl. Geochem., 16, 1513–1544, 2001a.

Oros, D.R. and Simoneit, B.R.T.: Identification and emission factors of molecular tracers in organic aerosols from biomass burning Part 2. Deciduous trees, Appl. Geochem., 16, 1545–1565, 2001b.

Oros, D.R., Abas, M.R.B., Omar, N.Y.M.J., Rahman, N.A., and Simoneit, B.R.T.: Identification and emission factors of molecular tracers in organic aerosols from biomass burning Part 3. Grasses, Appl. Geochem., 21, 919–940, 2006.

Othman, J., Sahani, M., Mahmud, M., and Ahmad, M.K.S.: Transboundary smoke haze pollution in Malaysia: Inpatient health impacts and economic valuation, Environ. Pollut., 189, 194–201, 2014.

Page, S.E., Siegert, F., Rieley, J.O., Boehm, H.-D.V., Jaya, A., and Limin, S.: The amount of carbon released from peat and forest fires in Indonesia during 1997, Nature, 420, 61–65, 2002.

Page, S.E., Rieley, J.O., and Wüst, R.: Chapter 7, Lowland tropical peatlands of Southeast Asia. In: Martini, I.P., Martinez Cortizas, A., Chesworth, W. (Eds.), Developments in Earth Surface Processes, Peatlands: Evolution and Records of Environmental and Climate Changes, 9, Elsevier, 145–172, 2006.

Pavagadhi, S., Betha, R., Venkatesan, S., Balasubramanian, R., and Hande, M.P.: Physicochemical and toxicological characteristics of urban aerosols during a recent Indonesian biomass burning episode, Environ. Sci. Pollut. Res., 20, 2569–2578, 2013.

Reid, J.S., Koppmann, R., Eck, T.F., and Eleuterio, D.P.: A review of biomass burning emissions part II: intensive physical properties of biomass burning particles, Atmos. Chem. Phys., 5, 799–825, 2005.

Rogge, W.F., Hildemann, L.M., Mazurek, M.A., Cass, G.R., and Simoneit, B.R.T.: Sources of fine organic aerosol. 3. Road dust, tire debris, and organometallic brake lining dust: roads as sources and sinks, Environ. Sci. Technol., 27, 1892–1904, 1993.

Sahani, M., Zainon, N.A., Mahiyuddin, W.R.W., Latif, M.T., Hod, R., Khan, M.F., Tahir, N.M., and Chan, C.-C.: A case-crossover analysis of forest fire haze events and mortality in Malaysia, Atmos. Environ., 96, 257–265, 2014.

Schlesinger, R.: The health impact of common inorganic components of fine particulate matter ($PM_{2.5}$) in ambient air: Critical review, Inhal. Toxicol., 19, 811–832, 2007.

See, S.W., Balasubramanian, R., and Wang, W.: A study of the physical, chemical, and optical properties of ambient aerosol particles in Southeast Asia during hazy and nonhazy days, J. Geophys. Res., 111, D10S08, doi:10.1029/2005JD006180, 2006.

See, S.W., Balasubramanian, R., Rianawati, E., Karthikeyan, S., and Streets, D.G.: Characterization and source apportionment of particulate matter ≤ 2.5 μm in Sumatra, Indonesia, during a recent peat fire episode, Environ. Sci. Technol., 41, 3488−3494, 2007.

Shafizadeh, F.: The chemistry of pyrolysis and combustion, In Chemistry of Solid Wood, Advances in Chemistry Series; R. Rowell, Ed.; American Chemical Society: Washington, DC, 207, 489–529, 1984.

Simoneit, B.R.T., Rogge, W.F., Mazurek, M.A., Standley, L.J., Hildemann, L.M., and Cass, G.R.: Lignin pyrolysis products, lignans, and resin acids as specific tracers of plant classes in emissions from biomass combustion, Environ. Sci. Technol., 27, 2533–2541, 1993.

Simoneit, B.R.T., Schauer, J., Nolte, C., Oros, D., Elias, V., Fraser, M., Rogges, W., and Cass, G.: Levoglucosan, a tracer for cellulose in biomass burning and atmospheric particles, Atmos. Environ., 33, 173–182, 1999.

Streets, D.G., Bond, T.C., Carmichael, G.R., Fernandes, S.D., Fu, Q., He, D., Klimont, Z., Nelson, S.M., Tsai, N.Y., Wang, M.Q., Woo, J.-H., and Yarber, K.F.: An inventory of gaseous and primary aerosol emissions in Asia in the year 2000, J. Geophys. Res., 108, D21, 8809, doi:10.1029/2002.JD003093, 2003.

Varkkey, H.: Regional cooperation, partronage and the ASEAN Agreement on transboundary haze pollution, Int. Environ. Agreements, 14, 65–81, 2014.

Wahid, N.B.A., Latif, M.T., and Suratman, S.: Composition and source apportionment of surfactants in atmospheric aerosols of urban and semi-urban areas in Malaysia, Chemosphere, 91, 1508–1516, 2013.

Yamamoto, S., Kawamura, K., Seki, O., Kariya, T., and Lee, M.: Influence of aerosol source regions and transport pathway on δD of terrestrial biomarkers in atmospheric aerosols from the East China Sea, Geochim. Cosmochim. Acta, 106, 164–176, 2013.

Yang, L., Nguyen, D.M., Jia, S., Reid, J.S., and Yu, L.E.: Impacts of biomass burning smoke on the distributions and concentrations of C_2–C_5 dicarboxylic acids and dicarboxylates in a tropical urban environment, Atmos. Environ., 78, 211–218, 2013.

Yong, D.L. and Peh, K. S.-H.: South-east Aisa's forest fires: blazing the policy trail, Oryx, doi:10.1017/S003060531400088X, 2014.

Zhu, C., Kawamura, K. and Kunwar, B.: Effect of biomass burning over the western North Pacific Rim: wintertime maxima of anhydrosugars in ambient aerosols from Okinawa, Atmos. Chem. Phys., 15, 1959–1973, 2015.

1 Table 1. Statistical results of chemical species concentrations. Av = Average. Sd = Standard deviation.

	Southwest monsoon (June–September)		Post-monsoon (October–November)		Northeast monsoon (December–March)		Pre-monsoon (April–May)	
Compounds	Av ± Sd	Range	Av ± Sd	Range	Av ± Sd	Range	Av ± Sd	Range
OC and EC [μg m^{-3}]								
OC	10 ± 7.8	3.6–36	5.6 ± 2.4	2.5–11	5.2 ± 1.4	2.7–8.2	4.2 ± 1.4	2.8–7.3
EC	3.0 ± 0.95	1.0–5.6	3.2 ± 1.3	1.1–5.9	3.4 ± 1.1	1.6–6.1	2.6 ± 1.2	1.4–4.5
Biomarkers [ng m^{-3}]								
levoglucosan	160 ± 130	32–490	64 ± 39	19–130	40 ± 14	17–64	49 ± 21	23–86
mannosan	8.4 ± 8.2	1.5–30	3.4 ± 2.6	0.95–9.1	2.6 ± 1.2	0.84–5.3	2.5 ± 1.2	1.2–5.3
galactosan	2.3 ± 2.3	0.38–8.3	0.86 ± 0.72	0.29–2.8	0.60 ± 0.35	0.13–1.3	0.62 ± 0.34	0.33–1.5
p-hydroxybenzoic acid	1.9 ± 1.9	0.18–7.5	0.79 ± 0.67	0.036–2.2	0.64 ± 0.30	0.20–1.2	0.50 ± 0.25	0.24–1.0
vanillin	1.6 ± 1.1	0.54–5.5	1.2 ± 0.66	0.45–2.2	1.0 ± 0.38	0.21–1.7	0.96 ± 0.42	0.30–1.7
syringaldehyde	0.29 ± 0.22	0.085–1.0	0.59 ± 0.22	0.26–1.2	0.77 ± 0.54	0.074–2.2	0.36 ± 0.22	0.093–0.77
vanillic acid	0.39 ± 0.39	0.074–1.9	0.11 ± 0.070	0.031–0.22	0.073 ± 0.057	0.013–0.26	0.066 ± 0.027	0.034–0.12
syringic acid	0.35 ± 0.41	0.075–2.4	0.26 ± 0.21	0.058–0.59	0.17 ± 0.13	0.029–0.64	0.16 ± 0.084	0.049–0.28
dehydroabietic acid	1.7 ± 1.1	0.10–5.4	1.1 ± 0.69	0.31–2.4	1.1 ± 1.1	0.14–4.6	0.67 ± 0.24	0.16–0.98
cholesterol	1.8 ± 0.82	0.50–3.7	1.2 ± 0.51	0.57–2.0	0.98 ± 0.51	0.026–2.0	1.3 ± 0.56	0.51–2.0
n-alkanes [ng m^{-3}]								
docosane	3.2 ± 0.82	1.8–5.0	2.9 ± 0.61	2.0–4.0	3.0 ± 0.53	1.9–4.2	4.0 ± 4.8	2.1–19
tricosane	3.6 ± 1.2	2.0–7.2	3.2 ± 0.91	2.0–4.8	3.2 ± 0.65	1.8–4.4	5.0 ± 7.6	2.1–29
tetracosane	5.8 ± 3.2	2.5–19	5.7 ± 1.7	3.3–8.7	6.1 ± 2.3	2.9–15	6.3 ± 8.5	2.7–33
pentacosane	8.9 ± 6.7	3.5–34	5.7 ± 2.3	3.1–11	6.0 ± 1.6	3.7–9.2	5.8 ± 5.5	3.2–23
hexacosane	13 ± 9.8	4.3–49	8.6 ± 3.7	3.6–18	9.7 ± 2.8	5.0–16	7.1 ± 5.3	3.5–23

heptacosane	16 ± 14	4.7–64	7.2 ± 2.6	3.6–12	8.2 ± 2.4	3.7–14	5.8 ± 3.4	3.3–16
octacosane	12 ± 12	2.6–54	4.3 ± 1.8	1.7–7.9	5.9 ± 3.0	2.3–17	3.6 ± 1.7	2.3–8.2
nonacosane	13 ± 13	3.0–55	4.9 ± 2.1	1.5–8.7	6.3 ± 2.2	3.3–13	4.5 ± 1.4	2.6–7.8
triacontane	7.9 ± 7.8	2.0–36	3.8 ± 2.0	1.6–9.0	5.2 ± 2.7	2.0–16	3.3 ± 1.7	1.7–8.3
hentriacontane	14 ± 14	2.8–59	4.8 ± 1.9	1.8–8.4	5.7 ± 2.0	3.3–11	4.3 ± 1.2	2.9–6.9
dotriacontane	6.7 ± 5.5	1.6–27	3.4 ± 0.72	2.4–4.5	4.6 ± 1.3	2.8–7.8	3.1 ± 0.88	1.8–4.4
tritriacontane	6.8 ± 7.1	1.2–33	2.5 ± 0.97	1.1–4.2	2.8 ± 0.92	1.2–5.0	2.1 ± 0.72	1.5–3.8

1 Table 2a. Factor loadings from varimax-rotated PCA of PJ_A data. A1–A3 indicate factors.

	A1	A2	A3
OC	**0.97**	0.10	0.16
EC	0.29	0.37	0.51
levoglucosan	**0.81**	-0.05	0.17
mannosan	**0.89**	0.00	0.11
galactosan	**0.90**	0.02	0.08
p-hydroxybenzoic acid	**0.94**	0.04	0.22
vanillin	0.61	0.15	0.25
syringaldehyde	-0.17	0.12	0.40
vanillic acid	**0.65**	-0.10	0.55
syringic acid	0.28	-0.11	**0.81**
dehydroabietic acid	0.15	-0.01	**0.86**
cholesterol	0.36	0.14	0.39
C_{22}	0.03	**0.95**	0.05
C_{23}	0.07	**0.95**	0.05
C_{24}	0.30	**0.92**	0.06
C_{25}	**0.81**	0.54	0.14
C_{26}	**0.86**	0.43	0.13
C_{27}	**0.95**	0.23	0.13
C_{28}	**0.96**	0.18	0.07
C_{29}	**0.97**	0.13	0.12
C_{30}	**0.92**	0.25	0.05
C_{31}	**0.97**	0.10	0.13
C_{32}	**0.93**	0.15	0.11
C_{33}	**0.97**	0.10	0.13
% variance	60	12	8.0
% cumulative	60	72	80

2

1 Table 2b. Factor loadings from varimax-rotated PCA of PJ_S data. S1–S5 indicate factors.

	S1	S2	S3	S4	S5	2
OC	0.47	0.47	0.10	0.08	0.57	
EC	0.39	0.20	0.25	0.26	**0.65**	
levoglucosan	0.09	**0.71**	-0.03	-0.52	0.19	
mannosan	0.19	**0.84**	0.02	-0.26	0.28	
galactosan	0.17	**0.83**	0.06	-0.09	0.41	
p-hydroxybenzoic acid	0.26	0.62	0.08	0.23	0.42	
vanillin	0.22	0.32	0.07	0.05	0.61	
syringaldehyde	0.24	0.13	0.01	**0.74**	0.07	
vanillic acid	-0.12	**0.81**	-0.04	0.22	-0.01	
syringic acid	0.02	**0.81**	0.00	0.37	0.26	
dehydroabietic acid	0.18	0.44	0.04	0.12	0.60	
cholesterol	0.01	0.17	0.15	-0.21	**0.77**	
C_{22}	0.05	-0.02	**0.97**	-0.04	0.05	
C_{23}	0.05	0.00	**0.97**	-0.04	0.04	
C_{24}	0.28	-0.03	**0.94**	0.04	-0.01	
C_{25}	0.33	0.10	**0.85**	0.05	0.35	
C_{26}	0.61	0.05	**0.68**	0.14	0.24	
C_{27}	**0.67**	0.08	0.53	0.10	0.35	
C_{28}	**0.86**	0.06	0.27	-0.01	0.01	
C_{29}	**0.89**	0.14	0.18	0.08	0.29	
C_{30}	**0.84**	0.03	0.33	0.04	-0.12	
C_{31}	**0.77**	0.24	0.07	0.10	0.47	
C_{32}	**0.88**	-0.04	0.02	0.10	0.16	
C_{33}	**0.72**	0.28	-0.03	0.14	0.49	
% variance	43	19	11	5.0	4.5	
% cumulative	43	62	72	77	82	

Figure Captions

Figure 1. Daily variability of the MAPI and visibility during the sampling periods.

Figure 2. Monthly hotspot counts in the Sumatra Island.

Figure 3. Daily variation of the OC fractions' mass concentrations during the sampling periods.

Figure 4. Box-whisker plots of molecular distributions of *n*-alkanes during the (a) southwest and (b) northeast monsoon seasons. The horizontal lines in the box represent the 25th, 50th, and 75th percentiles. The whiskers represent the 10th and 90th percentiles.

Figure 5. Number fraction of C_{max} in the $PM_{2.5}$ samples for each monsoon season.

Figure 6. P-values to determine significance in the two-sided Wilcoxon rank sum test between the IPF and other samples.

Figure 7. Daily variability of the C_{27} and LG concentration as well as the VA/SA and OP/OC4 mass ratios during the sampling periods.

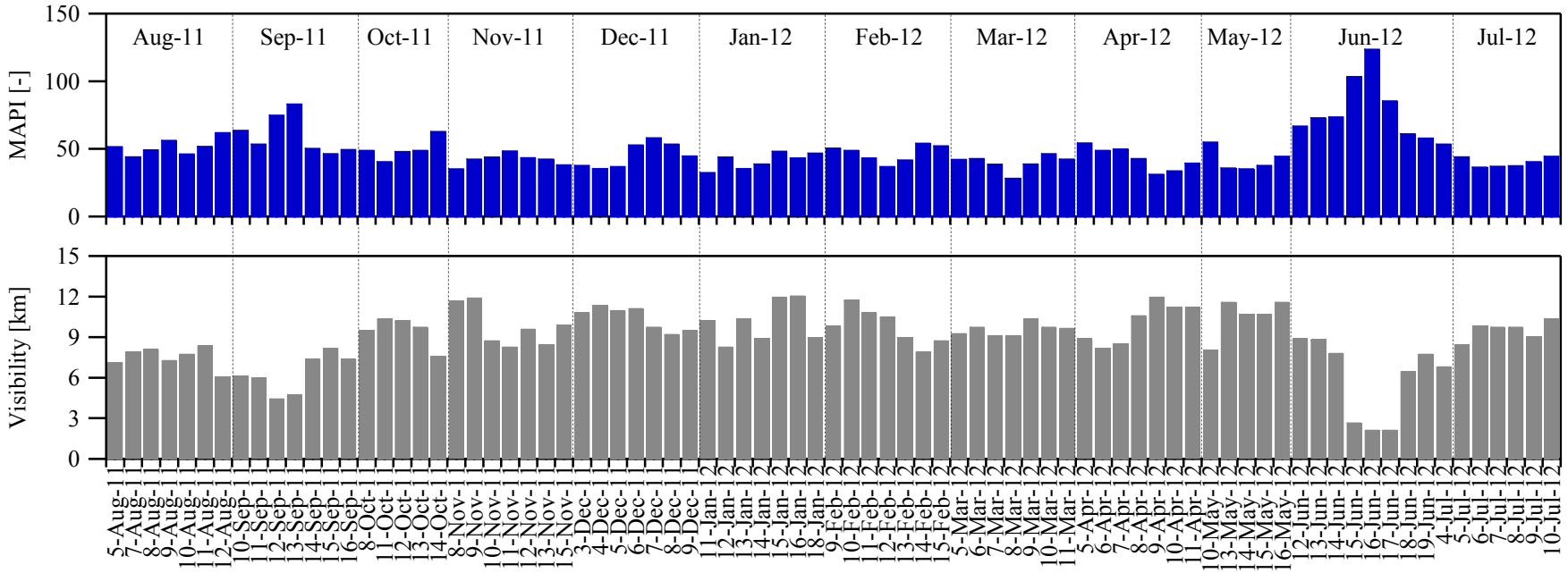

Figure 1. Daily variability of the MAPI and visibility during the sampling periods.

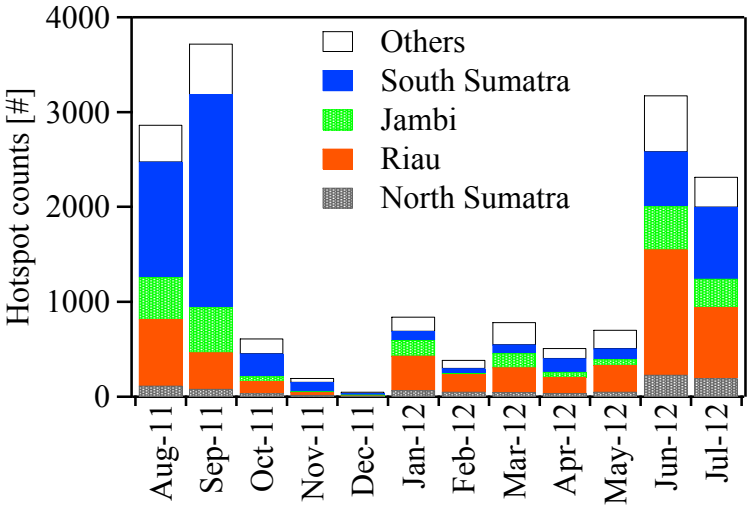

Figure 2. Monthly hotspot counts in the Sumatra Island.

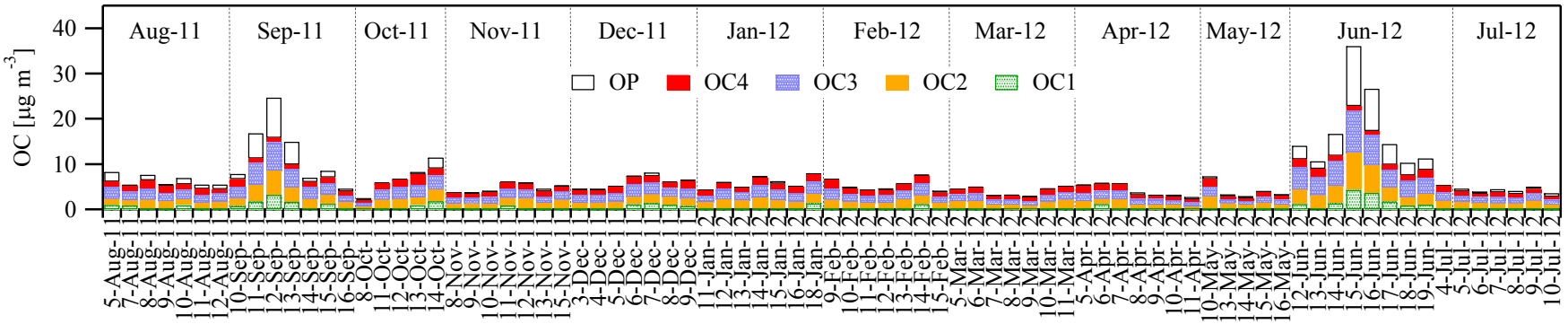

Figure 3. Daily variation of the OC fractions' mass concentrations during the sampling periods.

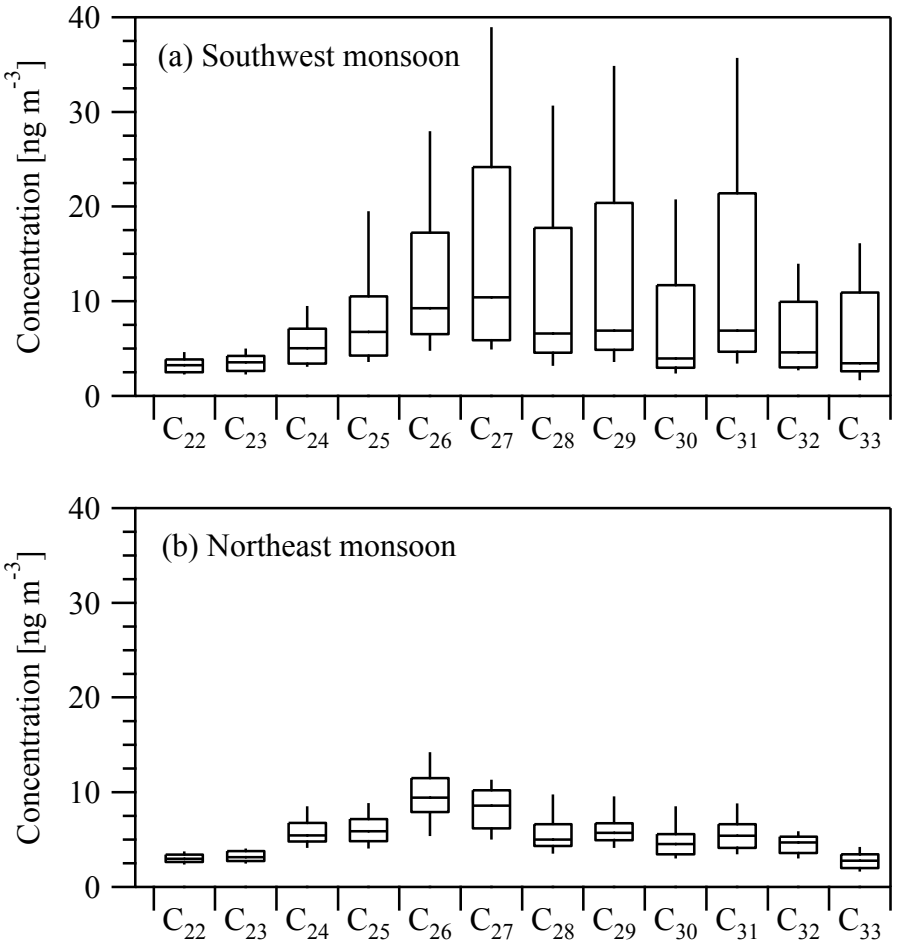

Figure 4. Box-whisker plots of molecular distributions of *n*-alkanes during the (a) southwest and (b) northeast monsoon seasons. The horizontal lines in the box represent the 25th, 50th, and 75th percentiles. The whiskers represent the 10th and 90th percentiles.

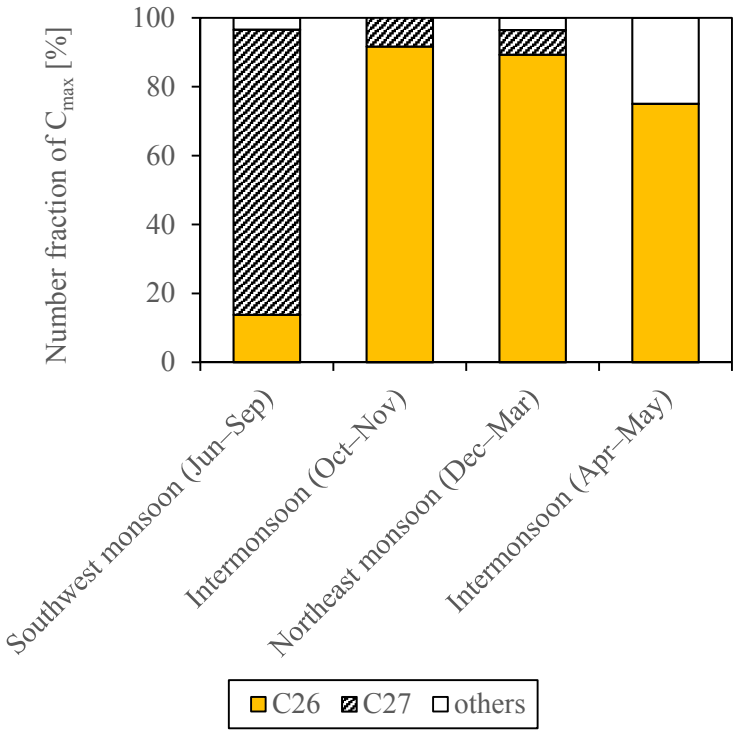

Figure 5. Number fraction of C_{max} in the $PM_{2.5}$ samples for each monsoon season.

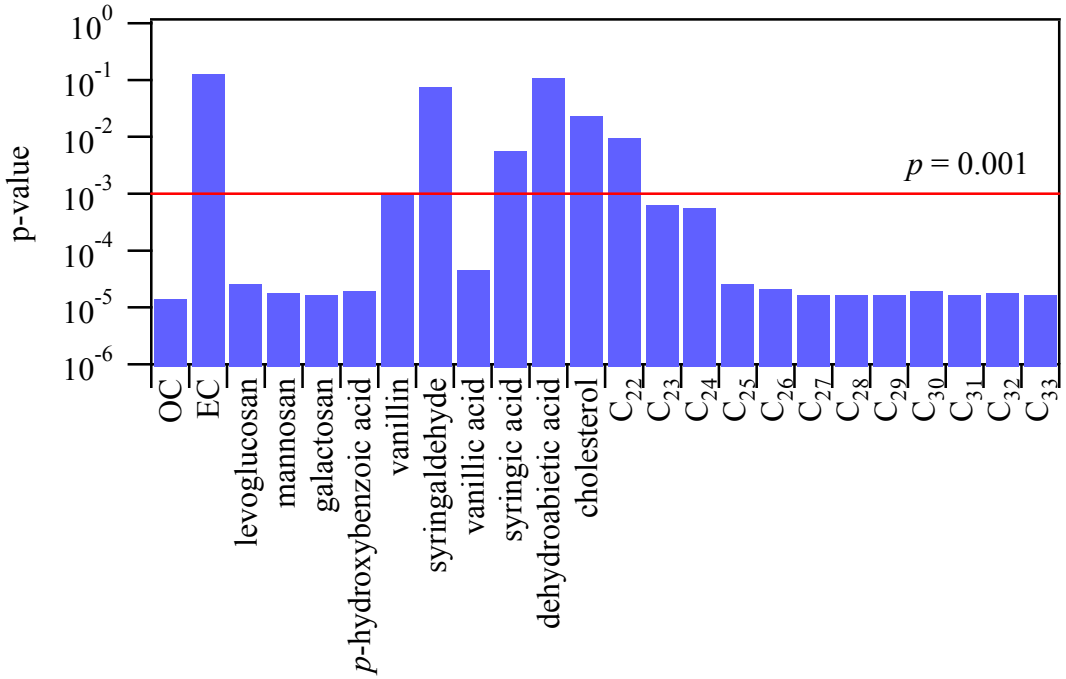

Figure 6. P-values to determine significance in the two-sided Wilcoxon rank sum test between the IPF and other samples.

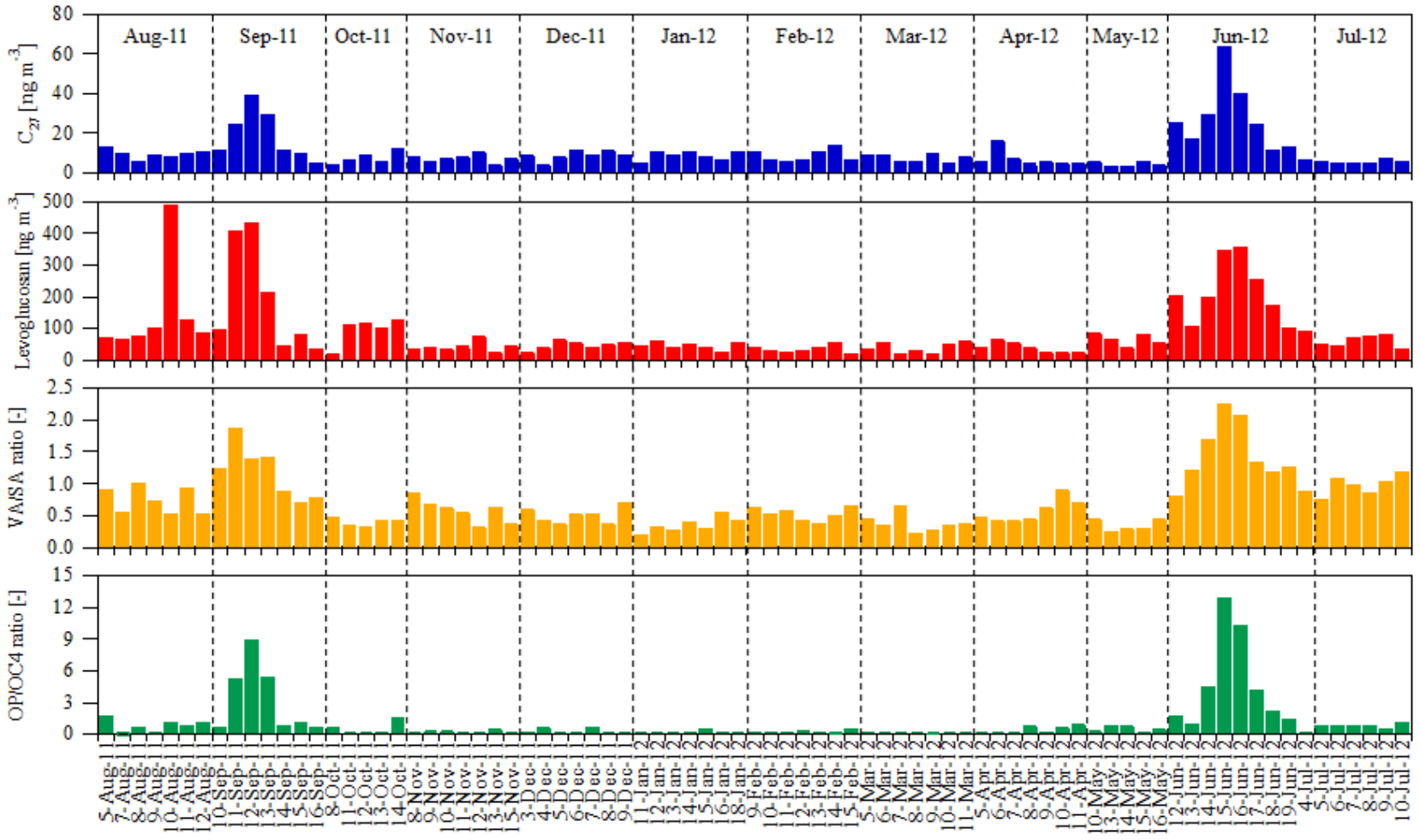

Figure 7. Daily variability of the C_{27} and LG concentration as well as the VA/SA and OP/OC4 mass ratios during the sampling periods.