Source apportionment of polycyclic aromatic hydrocarbons in urban air using positive matrix factorization and spatial distribution analysis

Eunhwa Jang a,b, Mohammed S. Alam a, Roy M. Harrison a,*,1

a Division of Environmental Health & Risk Management, School of Geography, Earth & Environmental Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom
b Busan Metropolitan City Institute of Health and Environment, MandeokDong, BukGu, Busan 616-110, Republic of Korea

HIGHLIGHTS

- PMF has been applied to two large multi-component PAH datasets.
- A complementary spatial analysis of paired sites has been conducted.
- Major PAH sources are traffic (exhaust and unburned fuel), coal burning, wood burning and steel industry.
- Total PAH and benzo(a)pyrene concentrations are apportioned for individual sites.

ABSTRACT

Polycyclic aromatic hydrocarbons (PAH) are currently generating a great deal of interest because of their recognised toxicity, including carcinogenicity. In this study, source apportionment (SA) has been carried out using Positive Matrix Factorisation (PMF) with a dataset of 29 individual PAH (sum of vapour and particulate forms) collected by the UK National Network between 2002 and 2006. Analysis of data from 14 urban sites revealed four major source categories corresponding to unburned petroleum, diesel combustion, wood combustion and coal combustion. When a separate set of sites known to be influenced by local industrial sources was analysed, three source categories were identified corresponding to the unburned petroleum, diesel combustion and coal combustion seen in the full data analysis. When SA data were applied to the individual sites, the estimated apportionment could be explained in terms of local emission characteristics. Unburned petroleum showed the highest contribution to the sum of PAH, averaging 51.9% across the network, but benzo(a)pyrene (BaP) was more influenced by the coal combustion source which contributed 59.5% across the entire network. At the subset of sites with local industrial influence, industry was both the main contributor to the sum of PAH (accounting for 48.4% of PAH mass) and of BaP (67.9% of mass). A spatial analysis was also conducted in which the traffic source was evaluated by the difference between a roadside and a nearby urban background site, the urban source by difference between urban background and a rural site, and the industrial source by difference between a site close to a major steelworks subtracting data from a local urban background site. This showed considerable similarity between the net urban contribution and the road traffic factor, and between the net industrial contribution and the PMF coal factor profile. In both cases the congener
profiles corresponded fairly well to UK national emissions inventory data. When PMF was applied separately to a more recent dataset for particle-bound PAH (2008-10) in three site groupings, it was able to distinguish the domestic coal burning source from the industry-related coal combustion source. For the urban sites, vehicle exhausts contributed the largest amount of particulate PAH and BaP across the whole year, with significant attribution to domestic coal combustion seen in the cold season.

© 2013 The Authors. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Over the past two decades the percentage of the population who live in urban areas has risen to over 50% and this proportion is expected to increase to over 70% by 2050 (WHO-UN-UNHABITAT, 2011). In highly populated areas, anthropogenic activities can cause very high air pollutant concentrations. This includes emissions both from stationary sources such as industry and domestic combustion and mobile sources, especially road traffic. Various organic and inorganic pollutants have been studied as tracers of urban atmospheric pollutants on a regional scale (Marr et al., 2004; Li et al., 2006; Duan et al., 2007; Sharma et al., 2007; Ma et al., 2010). In highly populated areas, anthropogenic activities can be significant as a major source of PAH emissions (Harrison et al., 2005; Marr et al., 2004; Galarneau et al., 2007; Sharma et al., 2007; Vardoulakis et al., 2008; Massei et al., 2007; Sharma et al., 2007; Vardoulakis et al., 2008; Miller et al., 2010). However, far less is known about the contribution of other source categories to the pollution of urban air.

Polycyclic aromatic hydrocarbons (PAH) are an air pollutant of particular concern because of the recognised mutagenic and carcinogenic properties of a number of the individual compounds (termed congeners). These compounds are ubiquitous as they are formed in all incomplete combustion processes. Unlike other persistent organic compounds, they have multiple point and diffuse sources and cannot be controlled by the introduction of substitute chemicals. Road traffic is well recognised as a major source of PAH emissions (Harrison et al., 1996; Galarneau et al., 2007) and many recent studies have focussed upon traffic as a source of PAH in urban areas (Guo, 2003; Motelay-Massei et al., 2007; Sharma et al., 2007; Vardoulakis et al., 2008; Miller et al., 2010). Because of their toxicity, the US Environmental Protection Agency has listed 16 PAHs as priority pollutants and categorised seven of them as carcinogenic chemicals (Wang et al., 2010). In 2005, the European Union set an annual target for benzo(a)pyrene (BaP) in ambient air of 1 ng m$^{-3}$. Some years earlier the government of the United Kingdom adopted an air quality objective for PAH expressed as a concentration of benzo(a)pyrene as an indicator of the PAH mixture of 0.25 ng m$^{-3}$ (Delgado-Saborit et al., 2011).

There is consequently a strong imperative to devise efficient strategies to reduce pollution of the urban atmosphere by PAH. A key aspect of such a requirement is the need for source apportionment in order to identify those sources most responsible for measured concentrations of PAH in urban air. Diagnostic ratios have been widely used particularly to distinguish the petrogenic (originated from petroleum) and pyrogenic (derived from combustion) PAH over many years (Grimmer et al., 1983; Rogge et al., 1993; Khalili et al., 1995; Ravindra et al., 2008). This technique uses the ratios between two PAH congeners as a form of source marker, but in a situation where multiple sources contribute to airborne concentrations, diagnostic ratios are of very limited value (Galarneau, 2008). Additionally, vapour-particulate phase partitioning can influence measured ratios as these are typically measured in the particulate phase.

Principal Component Analysis (PCA) has been very widely used as a statistical factor analysis method capable of separating

### Table 1

<table>
<thead>
<tr>
<th>Site information and details of datasets.</th>
<th>Source apportionment</th>
<th>Spatial analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total PAH</td>
<td>HMW PAH</td>
</tr>
<tr>
<td></td>
<td>urban</td>
<td>industry</td>
</tr>
<tr>
<td><strong>Andersen sampler dataset</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ballymena Ballykeel</td>
<td>–6.25829</td>
<td>54.85862</td>
</tr>
<tr>
<td>Birmingham</td>
<td>–1.83058</td>
<td>52.51172</td>
</tr>
<tr>
<td>Cardiff Lakeside</td>
<td>–3.16934</td>
<td>51.51241</td>
</tr>
<tr>
<td>Derry Brandywell</td>
<td>–7.33213</td>
<td>54.99234</td>
</tr>
<tr>
<td>Edinburgh St Leonards</td>
<td>–3.18219</td>
<td>55.94559</td>
</tr>
<tr>
<td>Glasgow Centre</td>
<td>–4.25516</td>
<td>55.85773</td>
</tr>
<tr>
<td>Harwell</td>
<td>–1.32528</td>
<td>51.57108</td>
</tr>
<tr>
<td>Hove</td>
<td>–0.18298</td>
<td>50.83659</td>
</tr>
<tr>
<td>Kinlochleven</td>
<td>–4.96418</td>
<td>56.71445</td>
</tr>
<tr>
<td>Leeds</td>
<td>–1.57862</td>
<td>53.76611</td>
</tr>
<tr>
<td>Lisburn Dumurr High School</td>
<td>–6.01402</td>
<td>54.53791</td>
</tr>
<tr>
<td>Liverpool Speke</td>
<td>–2.84433</td>
<td>53.34633</td>
</tr>
<tr>
<td>London Ashdown House</td>
<td>–0.13840</td>
<td>51.49680</td>
</tr>
<tr>
<td>London Brent</td>
<td>–0.27622</td>
<td>51.58977</td>
</tr>
<tr>
<td>London Crystal Palace Parade</td>
<td>–0.07553</td>
<td>51.42468</td>
</tr>
<tr>
<td>London Marylebone</td>
<td>–0.15461</td>
<td>51.52253</td>
</tr>
<tr>
<td>Manchester Law Courts</td>
<td>–2.25198</td>
<td>53.48080</td>
</tr>
<tr>
<td>Newcastle Centre</td>
<td>–1.61053</td>
<td>54.97825</td>
</tr>
<tr>
<td>Newport</td>
<td>–2.97728</td>
<td>51.60120</td>
</tr>
<tr>
<td>Swansea Cwm Level Park</td>
<td>–3.93945</td>
<td>51.64584</td>
</tr>
<tr>
<td>Bolsover</td>
<td>–1.29708</td>
<td>53.25637</td>
</tr>
<tr>
<td>Holyhead</td>
<td>–4.37590</td>
<td>58.18430</td>
</tr>
<tr>
<td>Middlesbrough</td>
<td>–1.22087</td>
<td>54.56930</td>
</tr>
<tr>
<td>Port Talbot Margam</td>
<td>–3.77082</td>
<td>51.58395</td>
</tr>
<tr>
<td>Royston</td>
<td>–1.43945</td>
<td>53.60028</td>
</tr>
<tr>
<td>Scunthorpe Low Santon</td>
<td>–0.59724</td>
<td>53.59583</td>
</tr>
<tr>
<td>Scunthorpe Town</td>
<td>–0.63681</td>
<td>53.58634</td>
</tr>
<tr>
<td>South Hiendley</td>
<td>–1.40084</td>
<td>53.61194</td>
</tr>
</tbody>
</table>
chemical constituents of the atmosphere according to their source (Hopke et al., 2005; Viana et al., 2008). We have previously analysed UK PAH data using PCA and have shown that the method is able to separate a number of contributory sources (Mari et al., 2010). However, the drawbacks of PCA including the common issue of negative solutions are well known and Positive Matrix Factorisation (PMF) has been widely adopted as a preferable technique for source apportionment of atmospheric aerosol constituents. This is an advanced factor analysis tool which has no negative constraints and is able to quantify the factor contribution directly without a subsequent use of multiple regression analysis. The method has been very widely applied to multi-element datasets (e.g. Viana et al., 2008) and to particle number datasets (Harrison et al., 2011). It has also been applied to PAH datasets (Dvorská et al., 2012), although most PAH apportionment studies have focussed on the 16 USEPA priority PAHs and few have examined temporal and spatial patterns in PAH.

PAH are known to be chemically reactive in the atmosphere and since the congeners react at different rates, the profiles of compounds emitted do not remain unchanged in the atmosphere. Consequently, Katsoyiannis et al. (2011) concluded that it was not feasible to obtain meaningful source apportionment information by receptor modelling techniques as congener profiles had changed between source and receptor. However, our earlier

![PMF factor profiles and temporal variation at urban sites (Andersen dataset).](image-url)
work with PCA (Mari et al., 2010) had reassured us that meaningful information remained within the congener profiles and consequently this study has used PMF to study the source apportionment of PAH as measured at urban sites across the UK. The UK network sampled particulate and vapour forms using Andersen “pesticide” samplers until 2006. Subsequently, from 2008, Digitel samplers were employed collecting only the particulate fraction.

2. Methodology

2.1. Data selection

(a) All details of datasets are given in Table 1. The Andersen sampler quarterly PAH datasets between 2002 and 2006 were downloaded from the UK Department of Environment, Food and Rural Affairs (DEFRA) website for PMF analysis. Total PAH concentration (vapour + particulate) was used to minimise the influence from ageing, photochemical degradation and partitioning (Harrison et al., 1996; Kim et al., 2009; Manoli et al., 2004); Datasets of urban (14) and urban industry (5) sites with 29 variables were separately prepared with the assumption that they have different source profiles. The Andersen sampler datasets were collected over 14-day periods, during which substantial decomposition of the particle-associated PAH can occur (Brown and Brown, 2013).

(b) For spatial analysis, three sampling sites with similar longitude and latitude were considered. Digitel sampler monthly PAHs concentrations from 2008 to 2009 were compared. The difference between the urban background (London Brent) and rural background (Harwell) was used to characterise local emissions in London. Additionally, the concentration difference between an urban traffic site (London Crystal Palace Parade) and an urban background site (London Brent) was taken to represent road vehicle emissions. The Digitel samples are collected over 24-h periods, during which only modest losses of particulate components occur — about 5% for benzo(a)pyrene (Brown and Brown, 2013).

A comparison between an urban industrial site with a local steelworks (Port Talbot Margam) and an urban background site (Swansea Cwm Level Park) was used to evaluate the steel industry signature of PAH.

(c) PAH atmospheric emission estimates between 2002 and 2006 were collected from the National Atmospheric Emission Inventory (NAEI) website and used to construct congener emission profiles for specific source types.

---

**Fig. 2.** PMF factor profiles and temporal variation at industrial urban sites (Andersen dataset).
Three Digitel sampler monthly datasets (domestic/industry/urban) between Dec 2008 and Nov 2010 were prepared. The known domestic combustion dataset included three Northern Ireland sites (Ballymena Ballykeel, Derry Brandywell and Lisburn Dunmurry). These parts of Northern Ireland lacked natural gas fuel, and hence coal and heating oil are widely used for domestic heating. Seven industry sites were combined to prepare a representative dataset. Thirdly, fourteen urban background sites including two strongly traffi

2.2. Data pretreatment and PMF analysis

The input data file consisted of receptor concentration (C) and uncertainty (U) matrices. U for each variable was calculated using C and the method detection limit (MDL). If C was above the MDL, U was assumed to 0.1C + MDL/3, and if vice versa C and U were replaced by MDL/2 and 0.2C + MDL/3, respectively (Tauler et al., 2009). For missing values, geometric mean concentration ($G_{\text{mean}}$) and $4 \times G_{\text{mean}}$ are often used for the C and U, respectively (Kim and Hopke, 2007); however, there was no missing value in this study. Overall, through the pretreatment process of input datasets, PMF allowed each sample and variable to be weighted individually, which is markedly different from PCA in which all are equally weighted (Park et al., 2011).

PMF decomposes receptors by three matrices- factor contribution (G), factor profile (F) and residual (E). 

$$X_{ij} = \sum G_{ik} F_{kj} + E_{ij}$$

where $X_{ij}$ is the $j$th measured species concentration in the $i$th sample, $G_{ik}$ is the factor contribution of $k$ source to $i$th sample, $F_{kj}$ is the fraction of $j$ species in the $k$ source, and $E_{ij}$ is the residuals matrix.

The solution in PMF is calculated through the weighted least square fit to search for proper $e_{ij}$ by minimizing the sum of the normalised $Q$ (Hopke, 2000; Larsen and Baker, 2003).

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

Table 2

<table>
<thead>
<tr>
<th>PAHs mass contribution ($\Sigma_{\text{total}}$PAH) change between 2002–2004 and 2005–2006 in urban sites.</th>
<th>2002 to 2004</th>
<th>2005 to 2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diesel emissions</td>
<td>32.1 %</td>
<td>11.8 %</td>
</tr>
<tr>
<td>Unburned petroleum</td>
<td>43.3%</td>
<td>46.0%</td>
</tr>
<tr>
<td>Coal combustion</td>
<td>8.8%</td>
<td>23.4%</td>
</tr>
<tr>
<td>Wood combustion</td>
<td>15.8%</td>
<td>18.8%</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>
where $e_{ij}$ is the scaled residual, and $\sigma_{ij}$ is the estimated uncertainty.

The robust mode is the default mode in PMF 3.0, which downweights the expected outliers with the value of outlier distance greater than four ($\left( \frac{e_{ij}}{\sigma_{ij}} \right)^2 > 4$ in eq (2)), in order to prevent outliers affecting the results (Reff et al., 2007; Sofowote et al., 2011). Unusual extreme events in the time series and samples with extremely high residuals were excluded with caution to minimize distortion in comparison between site-to-site quantified concentration levels (Norris et al., 2008).

Poorly categorised variables were excluded from among 29 PAHs congeners: acenaphthylene (Acy), acenaphthene (Ace), fluorene (Flu), phenanthrene (Phe), anthracene (Ant), 2-methyl

![PMF factor profiles](image)

![UK emission estimates](image)

**Fig. 5.** Comparison of profiles between PMF urban factors and UK emission inventory estimates (2002–2006).
phenanthrene (2 MPhe), 2-methyl anthrace (2 MA), 1-methyl anthracene (1 MA), 1-methyl phenanthrene (1 MPhe), 9-methyl anthracene (9 MA), 4,5-methylene phenanthrene (4,5 MPhe), fluoranthene (FluA), pyrene (Pyr), retene (Ret), benzo(c)phenanthrene (BcPhe), benz(a)anthracene (BaA), chrysene (Chr), cyclopenta(c,d)pyrene (CcdP), benzo(b)napah(2,1-d)thiophene (BN21T), 5-methyl chrysene (5MChr), benzo(a)pyrene (BaP), benzo(e)pyrene (BeP), indeno(123,cd)pyrene (IcdP), benzo(ghi)perylen (BghiPe), anthanthrene (Anth), dibenzo(al)pyrene (DalP), dibenzo(ae)pyrene (DaeP), dibenzo(ai)pyrene (DaiP) and dibenzo(ah)pyrene (DahP) remained. For the particle-bound PAH sampled by the Digitel sampler, fifteen variables- BaA, Chr, CcdP, BN21T, Benzo(b+j)fluoranthene (BbjF), Benzo(k)fluoranthene (BkF), BaP, BeP, Perylene (Per), IcdP, Dibenzo(ah,ac)anthracene (Dah,acA), BghiPe, Anth, DaiP and Coronene (Cor) from the domestic (73), industry (167) and urban (334) datasets were considered. Extra uncertainty (0–25%) and the number of factors (3–9) were explored to find the optimal PMF solutions.

The least squares approach can produce multiple solutions depending on the initial starting point because multiple F and G matrices with the same minimum Q value can exist (Sofowote et al., 2011). So rotation by changing Fpeak values between –1.4 and 0.3 was conducted to find an ultimate possible solution, and a bootstrapping technique based on arbitrary selection of n samples from datasets was tested to estimate a statistically stable solution.

2.3 Calculation of contribution

PMF consists of two factor functional matrices (G_{ik} and F_{kj}) and one residual matrix (E_{ij}). In our study, to quantify the distribution of each sample varying factors, we assumed that every factor has the same proportion of residuals, and calculated the factor contribution for each sample.

\[ X_{ij} = (G_{i1}F_{1j} + E_{ij}/n) + (G_{i2}F_{2j} + E_{ij}/n) + \cdots + (G_{in}F_{nj} + E_{ij}/n) \]

(3)

Fig. 6. Calculated net concentration of traffic and urban emissions (2008–2009).
Our input dataset for PMF analysis was a combination of several sites. However, the number of samples for each site was not exactly the same, because some outliers were excluded. Spatial patterns in PAH were also examined by calculation of an average concentration for each site.

2.4. Sensitivity test

The PMF method depends upon the congener profile of PAH from a given source type remaining unchanged during atmospheric transport. In practice, PAH are quite reactive in the atmosphere, leading to changes in congener profiles. This was investigated through a sensitivity study in which additional uncertainty was added to concentrations of each PAH. These added uncertainties were scaled from 5% for the least reactive compounds (dibenzopyrenes and anthanthrene) to 300% for the most reactive (anthracene), based upon rate coefficients for reaction with hydroxyl, where available. The effect upon source profiles extracted by PMF was very small.

3. Results and discussion


3.1.1. Urban sites

A four factor model gave the best fit in the PMF analysis. Mass contributions ($\sum_{\text{totalPAH}}$) were 21.4, 50.5, 15.0 and 13.1% for factors 1, 2, 3 and 4 respectively. Factor 1 (shown in Fig. 1a) showed two PAH congeners (2 MA and BN21T) to stand out with 87.7 and 51.6% contributions respectively from this factor. Additionally, over 20% of Phe, FluA, Pyr and Ret were explained by this factor. Westerholm and Li (1994) have reported that 2 MA is one of the most abundant PAH compounds in the emissions from diesel fuel vehicles. Other workers have suggested that thia-arenes including BN21T are strong diesel markers (McCary et al., 1996). Other studies have used BN21T and Pyr as diesel indicators to differentiate between diesel and gasoline sources (Alsberg et al., 1989; Larsen and Baker, 2003). This factor showed a slightly elevated contribution in the cold season (see Fig. 1a) and a sudden decline from 2005. This decline is consistent with the trend in the sulphur content of diesel fuels as reported also by Lim et al. (2005). We therefore assign this factor to emissions from diesel vehicles but note that its disappearance is a reflection of the changing composition of diesel and not of a sharp reduction in PAH emissions from diesel engines.

Factor 2 (Fig. 1b) showed a significant contribution of Phe and its methyl derivatives (MPhes). It reflects a substantial influence of lower molecular weight PAH. The factor accounts for 93.3% of 1 MA along with high loadings of 2 MPhe, 1 MPhe and 4.5 MPhe, which showed over a 70% contribution. Also, 59.1% of Phe was explained by this factor. Several studies have attributed high loadings of Phe and MPhes to either coal combustion or unburned petroleum from vehicles (Rogge et al., 1993; Gogou et al., 1996; Kavouras et al., 2001; Zuo et al., 2007). The fact that the contribution is higher in the warm season (Fig. 1b) and the association with low molecular weight PAH leads us to the conclusion that this factor arises from unburned petroleum-based fuels mainly by evaporation (Tuominen et al., 1988; Coleman et al., 1997; Meijer et al., 2008; Park et al., 2011). Earlier work has attributed such compounds both to gasoline emissions (Meijer et al., 2008; Park et al., 2011) or to emissions from diesel fuels (Marr et al., 2004).

Factor 3 (Fig. 1c) was more influenced by higher molecular weight PAH with it accounting for over 80% of DaLP, DaeP and Anth. Over 50% of other medium and high molecular weight PAH were accounted for including four-ring compounds (BpPhe, BaA, 5MChr, Chr), five-ring PAHs (Bp, Bpe, CcdP) and six-ring PAHs (BghiPe, IcdP). Compounds of this kind have previously been associated with emissions from industry (Lin et al., 2011) and from coke oven emissions (Daisey et al., 1986; Yang et al., 2002). Yang et al. (2002) has reported that emissions of PAHs from the steel and iron industry can be produced through coke manufacturing, sintering,
iron and steel making, casting and cooling processes and Ciaparra et al. (2009) used PCA to separate two atmospheric profiles from steel making processes, one with low molecular weight PAHs (Flu, Phe, Ant, FluA and Pyr) explained by coke making and the other with high molecular weight PAHs (BaP, IcdP, DahP, DacP and BghiP) arising from the sintering process. Slightly overlapped clusters between urban and industry sites in PCA across the UK may be partially explained by the influence of industrial activities upon the urban PAH (Brown and Brown, 2012). Consequently, we attribute factor 3 to coal utilisation deriving in part from metallurgical industries.

Factor 4 seen in Fig. 1d was characterised by a 84.6% contribution to Acy with substantial contributions to BghiPe (25.6%), Ant (23.5%) and Ret (23.1%). This factor showed markedly higher emissions in the colder months and no obvious long-term temporal trend. Mcdonald et al. (2000) reported a range of PAHs including Acy, Ant, BaP and BeP as being emitted from wood combustion, and retene is widely suggested as a wood combustion indicator (Khalili et al., 1995; Wang et al., 2009; Dvorská et al., 2012). Since wood combustion is mainly a winter phenomenon during the cold season for residential heating, we feel that this factor well fits the behaviour expected from wood burning.

3.1.2. Industrial urban sites

Five of the urban sites were identified as having a specific local industrial contribution. Port Talbot, Scunthorpe and Middlesbrough have local integrated steelworks. Holyhead was until recently a location used for aluminium smelting and Bolsover had a works for

<table>
<thead>
<tr>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
<th>Factor 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digitel sampler (industry)</td>
<td>Seasonal variation</td>
<td>Digitel sampler (domestic)</td>
<td></td>
</tr>
<tr>
<td><img src="image1" alt="Factor 1" /></td>
<td><img src="image2" alt="Seasonal variation" /></td>
<td><img src="image3" alt="Factor 3" /></td>
<td></td>
</tr>
<tr>
<td><img src="image4" alt="Factor 2" /></td>
<td><img src="image5" alt="Factor 2" /></td>
<td><img src="image6" alt="Factor 1" /></td>
<td></td>
</tr>
<tr>
<td><img src="image7" alt="Factor 3" /></td>
<td><img src="image8" alt="Factor 3" /></td>
<td><img src="image9" alt="Factor 4" /></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 8. PMF factor profiles of particle-associated PAH from industrial sites (left panel) and domestic combustion sites (right panel).
converting coal into processed solid fuels akin to the coking process. When PMF was applied to the pooled data from these five sites, it identified three factors accounting for 34.1, 48.4 and 17.5% respectively of PAH mass (\(\Sigma_{\text{PAH}}\)). Factor 1 seen in Fig. 2a had a similar source profile to factor 2 from the urban sites. The concentration profile was similar and also the factor accounted for 86.6% of MA as well substantial contributions to 1 MPhe (71.7%) and 2 MPhe (66.0%). The factor also accounted for a substantial proportion (53.6%) of retene but some studies have shown this to arise from sources other than wood burning (Khalili et al., 1995; Shen et al., 2012). For the reasons outlined above, and its similarity to Factor 2 in the urban data results, this factor was attributed to unburned petroleum.

Factor 2 accounted for 48.4% of PAH mass as well as a substantial proportion of most of the PAH congeners except the methylated PAHs. Anth had the highest factor contribution of 83.8% followed by significant amounts of BaA (73.0%) and derivatives of pyrene (BaP, IcdP and DaeP, DalP) (about 60%). This factor showed similarity to factor 3 in the urban dataset and reflected the emissions concluded as arising in part from the steel industry, reflecting the profiles reported by Ciaparra et al. (2009). Interestingly, the low molecular weight PAH congeners were more important in factor 2 from this dataset than in factor 3 from the urban dataset, perhaps reflecting a greater degree of atmospheric ageing of the PAH in the former-urban dataset rather than in these data which were collected much closer to source.

Factor 3 (Fig. 2c) showed a similar profile to the diesel emission factor (factor 1) from the urban sites. It accounted for high proportions of 2 MA (92.8%) and BN21T (82.8%). As noted above, Alsberg et al. (1989), Westerholm and Li (1994) and McCurry et al. (1996) attributed a factor of this composition to diesel as opposed to gasoline combustion. It showed no strong seasonal variation and a substantial decrease from 2005 onwards as with the urban diesel factor.

### 3.2. Analysis of Digitel (particulate-only) dataset: Dec 2008–Nov 2010

#### 3.2.1. Source profiles by subtraction

Fig. 3 shows pie charts of the mass distribution of the four source types in the two parts of the study (three source types in the industrial site sub-set). When these sources are estimated for each individual sampling site (shown in Fig. 4), the importance of the coal combustion contribution both to \(\Sigma_{\text{PAH}}\) (Fig. 4a) and to benzo(a)pyrene (Fig. 4b) is clearly seen. The most notable feature is the large contribution of coal burning to benzo(a)pyrene concentrations right across the range of sites. However, it seemed that the coal combustion profile could not be separated into a domestic combustion and an industrial emission profile, because one site known to be influenced by domestic solid fossil fuel combustion (Lisburn Dunmurry) had no tendency to produce another factor in PMF. In terms of total PAH mass concentration, the two traffic-related factors of diesel emissions and unburned petroleum typically dominate at all sites in the main urban dataset which is consistent with the conclusions of Mari et al. (2010). It is notable that the urban site with the highest BaP concentration except Lisburn Dunmurry is Kinlochleven which was until recently influenced by a local aluminium smelting plant.

On average, 13.1% of \(\Sigma_{\text{PAH}}\) was attributed to the wood combustion source in the main urban dataset. It was also the second most important contributor to BaP concentrations (24.2%). The change in prevalence of urban factor 1 attributed to diesel emissions is temporally consistent with the mandatory introduction of ultra low sulphur fuels in Europe. Under EU Directive 98/70/EC the introduction of ultra low sulphur fuels (less than 50 ppm) was required by 2005 (Brannigan et al., 2009). It is unclear whether this change in fuel quality was accompanied by a reduction in emissions of all PAH or whether the diesel combustion factor became combined with one of the other factors identified. Table 2 shows the
apparent change in PAH mass contribution between the years 2002–2004 and 2005 to 2006 when the datasets for these years were analysed separately. The reduction in the diesel emission profile is not reflected in any change in the unburned petroleum factor but there are apparent increases in the coal and wood combustion source contributions. Factor 1 shows some similarity to both of these and possibly more so than to factor 2 and there may be some artifactual reflection of diesel emissions in these two factors, but there is no straightforward way of resolving this issue.

3.2.2. Comparison with UK emissions inventory
The UK National Atmospheric Emissions Inventory includes the 16 USEPA priority PAHs. Eleven of these are common with the compounds included the PMF analysis and Fig. 5 shows a comparison of the UK emissions estimates for diesel vehicles, the iron and steel industry and domestic wood combustion with factors 1 to 4 from the urban dataset analysis. Perhaps the greatest similarity is between factor 3 and the iron and steel works profile seen in Fig. 5, but the emissions inventory data for domestic wood combustion compares very closely to that for factor 4 if Acy is excluded from the comparison. This could be a measurement issue as Acy is very volatile and potentially lost in the sampling and also highly reactive. Also, if Acy is excluded from consideration, then a composite of factor 1 and factor 2 bears considerable similarity to the diesel profile from the inventory in Fig. 5.

3.2.3. Net concentration method
The difference in concentration data from paired sites was analysed. Firstly, an urban traffic profile was derived by subtracting data from the urban background site at London Brent from the...
urban traffic site at London Crystal Palace Parade. The results appear in Fig. 6a as the net traffic profile. Fig. 6b shows the net urban profile derived by subtracting the rural Harwell site from the urban background site at London Brent. These are shown in Fig. 6c as percentages, i.e. (traffic-urban)/traffic × 100% and (urban-rural)/urban × 100% respectively. Some compounds show small negative concentrations but these are barely significant. Data are plotted separately for summer and winter to make some allowance for chemical reactivity with considerably greater differences seen in the net urban profile than the net traffic profile which is consistent with the longer atmospheric residence times of compounds measured at the rural Harwell site. Alternatively, this apparent high seasonality seen in the net urban concentration may reflect the seasonal variation of local emissions at urban background sites. It should also be noted that these data are derived from Digitel samplers which sample only the particulate phase component and that the change of phase from particulate to vapour between winter and summer will be very significant for some of the compounds. A comparison of the net traffic and net urban profiles with chemicals mostly consisting of above MDL values in Fig. 6c shows a considerable similarity between the two. The largest difference is for D(ah,ac)A, lower in net traffic than net urban, reflecting the very low traffic emissions of these compounds. Factor 4 of the industrial site dataset (Fig. 8a) and factor 3 of the Northern Ireland site dataset (Fig. 8e) are characterised by BN21T, recognised as a diesel emission marker in the previous total PAH dataset. The fact that all UK highway vehicle fuel was required to be “sulphur free” (<10 ppm) by 1 January 2009 (Jones et al., 2012) requires an alternative explanation. Sulphur levels of domestic fuel oil (kerosene) were not required to meet sulphur-free standards,

3.2.4. PMF factor profiles: Digitel/particulate-only dataset

PMF was applied to fifteen congeners of higher molecular weight expected to be present predominantly in the condensed phase.

3.2.4.1. Sites with known source influences. Fig. 8 compares PMF factors derived from the industry sites (left panel) with those from the Northern Ireland (domestic fuel sources) sites (right panel).

Fig. 11. Seasonal and spatial distribution of Andersen sampler HMW PAH in the urban dataset (diesel emissions (D.E.); traffic exhaust (T.E.); domestic coal combustion (D.C.); metallurgical industry (M.I.)).
and the seasonal variation in the domestic dataset led us to attribute this factor to domestic oil combustion (Ward et al., 2012).

Factor 2 of both the industrial dataset (Fig. 8b) and Northern Ireland (Fig. 8f) dataset showed a high contribution to the known vehicle exhaust markers. Many studies have reported that BghiPe, IcdP and Cor are strong identifiers of gasoline engine exhaust aerosols (Marr et al., 2006; Andreou and Rapsomanikis, 2009; Amador-Munoz et al., 2010; Li et al., 2011). A tunnel study in which degradation effects were negligible showed particle-associated BghiPe and Cor to be markers of gasoline engine exhaust, and Chr is related more to diesel engine exhaust (Rajput and Lakhani, 2010). BkF, which is considered a more stable substance than BeP (Park et al., 2002) was seen significantly in the traffic source profile (Rajput and Lakhani, 2010; Li et al., 2011). Riddle et al. (2007) have noted that Cor originally suggested as a gasoline marker can also be emitted from heavy duty diesel vehicles at low speed (Riddle et al., 2007). Consequently, this source was assigned to traffic exhaust.

Factor 3 of the industry dataset (Fig. 8c) and factor 1 of the domestic dataset (Fig. 8g) are strongly associated with CcdP and Anth recognised as coal combustion identifiers. A high contribution to BaA, CcdP and BaP from bituminous coal combustion was reported in Greece (Andreou and Rapsomanikis, 2009). A large seasonal variation between summer and winter in both datasets is consistent with this source as due to coal combustion for domestic heating. A significant contribution of domestic coal burning to particle-bound PAH including BaP was seen at the three sites in Northern Ireland known to be heavily influenced by domestic combustion (Figure 9c and d), and a lesser contribution was seen at the industrial sites (Fig. 9a and b).

One profile (Factor 1) in the industrial dataset (Fig. 8d) was similar to the domestic coal burning factor except for Per and DaiP. The source of Per has been controversial because of its different trend from other PAH isomers generally produced through combustion processes (Reddy et al., 2002; Lima et al., 2002). However, factor analysis (PCA) with aerosol samples by Kavouras et al. (2001) extracted one single component factor (Per), and the fact that PAH in fine particles are mostly generated through high temperature processes may support an attribution to pyrolysis sources. A small seasonal variation is consistent with emissions from industry. A relatively higher contribution to BaP was apparent at the two sites in Scunthorpe (Fig. 9b), known to be influenced by steel industry emissions.

3.3. Comparison of Andersen high molecular weight with Digitel data

Fig. 10 shows the results of comparison of the application of PMF to the two datasets, using only high molecular weight species predominantly in the particle phase. A comparison of Factor 1 of the Andersen sampler dataset (Fig. 10a) and factor 4 of the Digitel sampler dataset (Fig. 10e) showed both factors to be significantly characterised by BN21T understood as diesel marker. However, the profile of other congeners differs substantially between the Andersen dataset (2002–2006) and the Digitel dataset (Dec 2008–Nov 2010). Our interpretation is that this reflects a change in the predominant source of BN21T from diesel fuel to domestic oil combustion following the transition to sulphur-free diesel.

Factor 2 (Fig. 10b) of the Andersen HMW PAH was characterised by high contributions (>40%) of Chr, BeP, IcdP and BghiPe. Factor 1...
(Fig. 10f) of the Digitel PAH dataset also showed a similar profile and BbF, BkF, Dah,aC and Cor additionally contributed to this factor. These congeners have been reported as traffic emission markers in many studies. The importance of BghiP as a traffic marker is confirmed by its strong presence in the traffic profile in Fig. 6. In the seasonal and spatial distribution, this source contributed significantly to the particle-associated PAH in the cities, especially in London (Crystal Palace and Marylebone Road) throughout the entire year (Fig. 11, Fig. 12), and a predominant contribution in the warm season was apparent across the UK, especially in the post-sulphur reduction dataset (Fig. 12).

Factor 3 of the Andersen sampler dataset (Fig. 10c) and Digitel sampler dataset (Fig. 10g) was strongly associated with CcDp and Anh with over 70% contributions. Noticeable seasonality and a high similarity to the domestic combustion dataset (Fig. 8g) lead us to assign this factor to domestic coal combustion in the Digitel sampler dataset (Fig. 10g). A lesser seasonality in the Andersen sampler dataset (Fig. 10c) seems to be related to the relatively higher influence from coal combustion in industry. The change in Kinlochleven from little seasonality in the Andersen dataset (Fig. 11) to stronger seasonality in the later (Digitel) dataset (Fig. 12) supports the concept of two coal combustion sources.

Factor 4 in the Andersen sampler dataset (Fig. 10d) and Factor 2 from the Digitel sampler (Fig. 10h) dataset were characterised by limited variables such as DaIP and Anh. Significant DaIP was explained by industrial activities in the previous industry dataset.

Different photochemical degradation rates and vapour/particle partitioning between individual PAHs may cause changes in source profiles, when expressed as concentrations or abundances. However, it has been reported that specific PAH factor patterns were preserved between two datasets having different oxidants levels (Dvorská et al., 2012). In addition, a high similarity of source profiles between an Andersen high molecular weight dataset and a Digitel particulate dataset using different sampling durations, fitted by its strong presence in the traffic profiles (Fig. 6). In the second source apportionment of both total PAH and of benzo(a)pyrene is from the coal combustion in industry. The change in Kinlochleven from little seasonality in the Andersen dataset (Fig. 11) to stronger seasonality in the later (Digitel) dataset (Fig. 12) supports the concept of two coal combustion sources.

4. Conclusions

Despite initial doubts over the application of PMF to PAH data arising from the atmospheric reactivity of PAH, this study has shown a remarkable ability of PMF to separate out the contributions of four different major sources to total PAH concentrations within UK urban areas. These sources in the general urban data are seen to be a diesel emission source, unburned petroleum, coal combustion and wood combustion. For the urban sites, the largest contribution to the sum of PAH mass (\(\Sigma_{\text{total}}\)PAH and \(\Sigma_{\text{pa}}\)PAH) is from the traffic-related sources as was concluded previously from a PCA analysis (Mari et al., 2010). However, perhaps most importantly, the work shows that the most significant contribution to the highly carcinogenic benzo(a)pyrene is from the coal combustion source at the urban and industrial sites. Only small differences are detected between the PAH profile emissions from domestic coal combustion, and those from use of coal in the steel industry, but these sources have been separated in the Digitel dataset.

Brown and Brown (2012), using a novel combined diagnostic ratio-PCA approach, were able to group UK sites in terms of solid fuel use, industry and traffic, and those of an urban and rural nature. The results of our study are broadly consistent with those of Brown and Brown (2012), but go one step further in providing quantitative source apportionment of both total PAH and of benzo(a)pyrene.

Substantial progress has been made in reducing airborne concentrations of PAH in the United Kingdom (Smith and Harrison, 1996) but clearly if exposures to carcinogenic PAH are to be reduced further, detailed attention needs to be given to the coal burning emissions particularly as abatement measures applied to road traffic are having a major beneficial impact upon emissions of both hydrocarbon vapour (i.e. catalytic converters) and particulate matter (diesel particle traps).

Acknowledgements

We are grateful to the support from UK Natural Environment Research Council (grant number NE/F016581/1) and a Chevening Scholarship (KR/D/3/7/002/001/133, SEO/2267/1110) to Eunhwa Jang. The underlying research materials for this study, i.e. data used, can be accessed via email requesting information to EXJ101@bham.ac.uk.

References


