Author's preprint version of the article:

Source apportionment of aerosol particles at a European air pollution hot spot using particle number size distributions and chemical composition

Leoni C, Pokorná P, Hovorka J, Masiol M, Topinka J, Zhao Y, Křůmal K, Cliff S, Mikuška P, Hopke PK.

Published in Environmental Pollution 234 (2018) 145e154

DOI: https://doi.org/10.1016/j.envpol.2017.10.097

The original paper is available at:

https://www.sciencedirect.com/science/article/pii/S0269749117331640

Manuscript Details

Manuscript number	ENVPOL_2017_3669
Title	Source apportionment of aerosol particles at a European air pollution hot spot using particle number size distributions and chemical composition
Article type	Research Paper

Abstract

Ostrava in the Moravian-Silesian region (Czech Republic) is a European air pollution hot spot for airborne particulate matter (PM), polycyclic aromatic hydrocarbons (PAHs), and ultrafine particles (UFPs). Air pollution source apportionment is essential for implementation of successful abatement strategies. UFPs or nanoparticles of diameter <100 nm exhibit the highest deposition efficiency in human lungs. To permit apportionment of PM sources at the hot-spot including nanoparticles, Positive Matrix Factorization (PMF) was applied to highly time resolved particle number size distributions (NSD, 14 nm-10 μm) and PM0.09-1.15 chemical composition. Diurnal patterns, meteorological variables, gaseous pollutants, organic markers, and associations between the NSD factors and chemical composition factors were used to identify the pollution sources. The PMF on the NSD reveals two factors in the ultrafine size range: industrial UFPs (28%, number mode diameter - NMD 45 nm), industrial/fresh road traffic nanoparticles (26%, NMD 26 nm); three factors in the accumulation size range: urban background (24%, NMD 93 nm), coal burning (14%, volume mode diameter - VMD 0.5 μm), regional pollution (3%, VMD 0.8 μm) and one factor in the coarse size range: industrial coarse particles/road dust (2%, VMD 5 μm). The PMF analysis of PM0.09-1.15 revealed four factors: SIA/CC/BB (52%), road dust (18%), sinter/steel (16%), iron production (16%). The factors in the ultrafine size range resolved with NSD have a positive correlation with sinter/steel production and iron production factors resolved with chemical composition. Coal combustion factor resolved with NSD has moderate correlation with SIA/CC/BB factor. The organic markers homohopanes correlate with coal combustion and the levoglucosan correlates with urban background. The PMF applications to NSD and chemical composition datasets are complementary. PAHs in PM1 were found to be associated with coal combustion.

Keywords	Industry; Local heating; Nanoparticles; Positive Matrix Factorization; Polycyclic Aromatic Hydrocarbons.

Corresponding Author	Cecilia Leoni
Corresponding Author's Institution	Institute for Environmental Studies, Faculty of Science, Charles University
Order of Authors	Cecilia Leoni, Petra Pokorna, Jan Hovorka, Mauro Masiol, Jan Topinka, Yongjing Zhao, Kamil Krumal, Steven Cliff, Pavel Mikuska, Philip Hopke
Suggested reviewers	Gianluigi de Gennaro, Manuel s Dall'Osto, Adam Reff, Tuan Vu, Roberta Vecchi

Submission Files Included in this PDF

File Name [File Type]

CoverLetter_CL.pdf [Cover Letter]

Highlights.docx [Highlights]

Graphical_abstract_Source_Apportionment_NSD_CC_Leoni2017.pdf [Graphical Abstract]

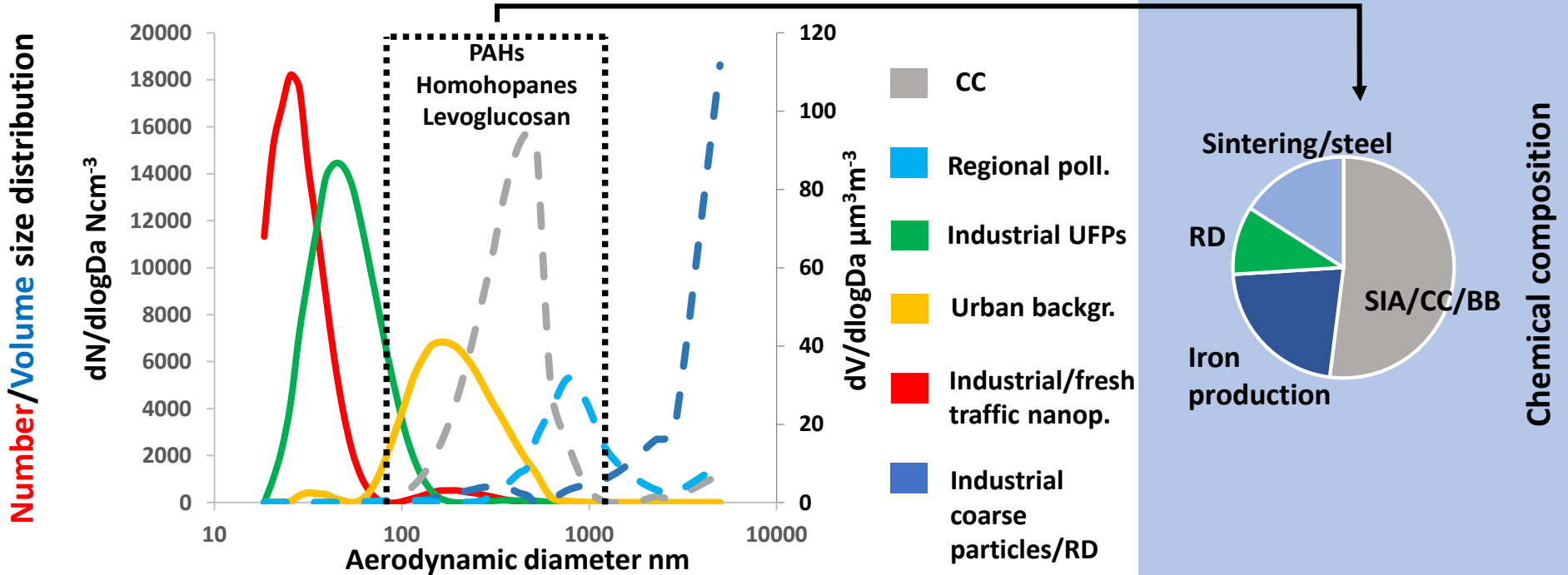
Source_Apportionment_NSD_CC_Leoni2017.docx [Manuscript File]

Supplementary_material_Source_Apportionment_NSD_CC_Leoni2017.docx [Supporting File]

To view all the submission files, including those not included in the PDF, click on the manuscript title on your EVISE Homepage, then click 'Download zip file'.

Highlights

- Industry is the major source of UFPs
- Local combustion sources are the major contributors to $PM_{0.09\text{-}1.15}$
- PAHs in PM_1 are related to coal combustion
- NSD PMF reveals UFPs sources not detectable with chemical composition modeling
- Chemical composition helps to interpret NSD PMF output and identify the sources


Number/Volume size distribution
dN/dlogDa Ncm-3
dV/dlogDa µm3m-3
20000
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
120
100
80
60
40
20
0
10
100
1000
10000
Aerodynamic diameter nm
PAHs
Homohopanes
Levoglucosan
CC
Regional poll.
Industrial UFPs
Urban backgr.
Industrial/fresh traffic nanop.
Industrial coarse particles/RD
Chemical composition
Sintering/steel
RD
SIA/CC/BB
Iron production

Source apportionment of aerosol particles at a European air pollution hot spot using particle number size distributions and chemical composition

Cecilia Leoni[1], Petra Pokorná[1,2], Jan Hovorka[1], Mauro Masiol[3], Jan Topinka[4], Yongjing Zhao[5], Kamil Křůmal[6], Steven Cliff[5], Pavel Mikuška[6], Philip K. Hopke[3]

[1]Institute for Environmental Studies, Faculty of Science, Charles University, Benátská 2, 128 01 Prague 2, Czech Republic.

[2]Laboratory of Aerosol Chemistry and Physics, Institute of Chemical Process Fundamentals CAS, Rozvojová 2/135, 165 02 Prague 6, Czech Republic.

[3]Center for Air Resources Engineering and Science, Clarkson University, Potsdam, NY 13699-5708, USA.

[4]Department of Genetic Toxicology and Nanotoxicology, Institute of Experimental Medicine CAS, Vídeňská 1083, 142 20 Prague 4, Czech Republic.

[5]Air Quality Research Center, University of California, Davis, CA 95616, USA.

[6]Institute of Analytical Chemistry CAS, Veveří 967/97, 602 00 Brno, Czech Republic.

*Please address correspondence to: Cecilia Leoni, Institute for Environmental Studies, Faculty of Science, Charles University, Benátská 2, 128 01 Prague 2, Czech Republic; e-mail: leonic@natur.cuni.cz

Abstract

Ostrava in the Moravian-Silesian region (Czech Republic) is a European air pollution *hot spot* for airborne particulate matter (PM), polycyclic aromatic hydrocarbons (PAHs), and ultrafine particles (UFPs). Air pollution source apportionment is essential for implementation of successful abatement strategies. UFPs or nanoparticles of diameter <100 nm exhibit the highest deposition efficiency in human lungs. To permit the apportionment of PM sources at the hot-spot including nanoparticles, Positive Matrix Factorization (PMF) was applied to highly time resolved particle number size distributions (NSD, 14 nm-10 μm) and $PM_{0.09\text{-}1.15}$ chemical composition. Diurnal patterns, meteorological variables, gaseous pollutants, organic markers, and associations between the NSD factors and chemical composition factors were used to identify the pollution sources. The PMF on the NSD reveals two factors in the ultrafine size range: industrial UFPs (28%, number mode diameter - NMD 45 nm), industrial/fresh road traffic nanoparticles (26%, NMD 26 nm); three factors in the accumulation size range: urban background (24%, NMD 93 nm), coal burning (14%, volume mode diameter - VMD 0.5 μm), regional pollution (3%, VMD 0.8 μm) and one factor in the coarse size range: industrial coarse particles/road dust (2%, VMD 5 μm). The PMF analysis of $PM_{0.09\text{-}1.15}$ revealed four factors: SIA/CC/BB (52%), road dust (18%), sinter/steel (16%), iron production (16%). The factors in the ultrafine size range resolved with NSD have a positive correlation with sinter/steel production and iron production factors resolved with chemical composition. Coal combustion factor resolved with NSD has moderate correlation with SIA/CC/BB factor. The PMF applications to NSD and chemical composition datasets are complementary. The organic markers homohopanes correlate with coal combustion and the levoglucosan correlates with urban background. PAHs in PM_1 were found to be associated with coal combustion.

Main finding:

Industry is the major source of UFPs at the receptor site. Local combustion sources contribute to the accumulation mode particles. PAHs in PM_1 are associated with coal combustion.

Keywords

Industry, Local heating, Nanoparticles, Positive Matrix Factorization, Polycyclic Aromatic Hydrocarbons.

Introduction

During the last 15 years, the 75% of the air quality monitoring stations in European Union (EU) registered a drop in the concentrations of atmospheric particulate matter (PM, EEA Report 2016). However, the concentrations of PM_{10} and $PM_{2.5}$ still exceed the EU limit values in some regions (EEA, 2016). Recently, ultrafine particles (UFPs, diameter <100 nm) have received great attention because they are particularly hazardous for human health: (i) they can reach the alveolar region of lung; (ii) they have high deposition efficiency (Venkatamaraan, 1999); and (iii) they have orders of magnitude higher surface area to mass ratios compared to larger particles (Oberdörster et al. 1994).

The Moravian-Silesian region (Fig.1), in the north-eastern part of the Czech Republic, is an EU *hot spot* for air pollution. Epidemiological studies have demonstrated that the air pollution in Ostrava, the major city of the region, significantly affects the health of the population, with an increased rate of respiratory illnesses compared to other regions of the Czech Republic (Šrám et al., 2013, Topinka et al., 2015). High anthropogenic emissions, due to the steel industry, coke plants, domestic heating, vehicular traffic, and the transport of polluted air masses from Poland, contribute to the worsening of air quality especially in the winter (Mikuška et al., 2015, Pokorná et al., 2015, 2016). The 24-hours PM_{10} limit (50 μg m^{-3}) is frequently exceeded (CHMI, 2017) along with elevated concentrations of polycyclic aromatic hydrocarbons (PAHs) (Mikuška et al., 2015). High concentrations of UFPs (up to 1.4×10^5 cm^{-3}) highly enriched with PAHs (2.9 mg/g) were observed in the winter 2014 (Leoni et al., 2016). PAHs are formed by the incomplete combustion of fossil fuels and wood, and they have carcinogenic and mutagenic properties (Ravindra et al., 2007).

The planning and the application of abatement strategies to improve air quality in Ostrava are only possible when the pollution sources are identified and apportioned. This is challenging in this location, due to the presence of several sources, some of them situated near urban settlements. The source apportionment is possible through the application of Positive Matrix Factorization (PMF) bilinear model (Polissar et al., 2001; Hopke, 2016), where input data are composed of two matrices of temporal variability of aerosol chemical composition and mass. PMF with highly time resolved aerosol data provides information at temporal resolution capable of identifying not only the main PM sources, but also sources that may have too short duration impact to be observed in 24 h integrated samples (e.g. Elsasser et al. 2012; Ancelet et al., 2012, 2014; Pancras et al., 2013; Moreno et al., 2013; Hovorka et al., 2015). Recently, source apportionment studies focus not only on particle mass, but also on NSD (Harrison et al., 2011, Beddows et al., 2015, Masiol et al., 2017, Sowlat et al., 2016). The analysis of the

mass chemical composition data distinguish sources contributing mainly to particle mass, while the analysis of the particle NSD identifies sources contributing principally to particle number, enabling the source apportionment down to nanoparticles (Beddows et al., 2015; Masiol et al., 2017).

Recent Ostrava pollution source apportionment studies, based on size segregated aerosol chemical composition (Pokorná et al. 2015, 2016) revealed coal combustion (CC), raw iron production, steel production and traffic being sources of $PM_{0.34-1.15}$ and road dust source of coarse particles ($PM_{1.15-10}$). The use of specific molecular markers (Mikuška et al., 2015) revealed combustion of wood and coal, vehicular emissions and industrial production of coke and iron the main $PM_{2.5}$ sources. Despite these studies, detailed pollution source identification down to nanoparticles was not performed. Additionally, the identification of the CC sources is difficult: coke plant, home heating, and transports of polluted air masses from the near Polish industrial areas may have very similar chemical profiles (Pokorná et al., 2015).

The aim of this study is to give further insights on air pollution sources in Ostrava, including in the nanoparticle pollution. Gaseous pollutants, organic markers and meteorological variables are used to help the source identification. The factors resolved with NSD are compared with mass chemical composition factors and associations are disclosed. Lastly, since PAHs are in high concentration in Ostrava and they are strongly harmful to human health, their association with PMF-revealed sources is discussed.

Materials and methods

Experimental

An intensive sampling campaign was performed from 4th February to 7th March 2014 in the residential district of Ostrava – Radvanice and Bartovice located in the southeastern part of the city. A large metallurgy complex is located 1.5 km southwest of the sampling station (Fig. 1). The instruments were placed in an air-conditioned container. Five-minute integrated particle NSD were measured with a Scanning Mobility Particle Sizer (14 – 730 nm, SMPS-3936L25, TSI Inc.) and an Aerodynamic Particle Sizer (0.523 – 10 μm, APS-3321, TSI Inc.). Size segregated PM was collected with a Davis Rotating-drum Uniform-size-cut Monitor – 8DRUM (DELTA Group UC-Davis), from the 10th to 28th of February and is used to provide 2-h resolved PM compositions. Particles were collected on Mylar substrates lightly greased with Apiezon™. The 8DRUM collects particles in 8 size ranges from 0.09 μm to 10 μm. The

five smallest size range samples, from 0.09 µm to 1.15 µm, were analyzed for 24 elements (Na, Mg, Al, Si, P, S, Cl, K, Ca, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Ga, As, Se, Br, Rb, Pb) using synchrotron X-ray fluorescence (S-XRF) by the Air Quality Research Center, University of California. For the modeling, data from the five size ranges were merged in $PM_{0.09\text{-}1.15}$ to overlap the size range of the NSD. The 2-hours mass concentrations were measured with beta attenuation monitor (Ortec 590A).

Five-minute PM_{10} mass concentration was measured with a beta attenuation monitor (FH 62 I-R, Thermo ESM Andersen); wind speed – WS and wind direction – WD (WindSonic, Mc Gill); temperature – T (Comet 200-80/E); precipitation by disdrometer (Laser Precipitation Monitor, Thies); CO, NO_2, NO, O_3, SO_2, CH_4 and non-methane hydrocarbons with automatic monitors (Horiba-360 series). Hourly concentrations of organic and elemental carbon – OC/EC $PM_{2.5}$ (Sunset) were measured in semi continuous regime (45 minutes collection and 15-minute analysis) using the NIOSH protocol (Birch and Cary, 1996), but only total carbon (TC) was eventually available for the data analysis due to technical problems with the analyzer. 24-hours levoglucosan, 16 US-EPA PAHs, 22R-17α(H),21β(H)-homohopane and 22S-17α(H),21β(H)-homohopane were determined in PM_1 collected on quartz filters using a high volume sampler (DHA-80, Digitel, 30 m^3 h^{-1}). Details of the sample preparation and the GC-MS analysis are described elsewhere (Mikuška et al., 2015).

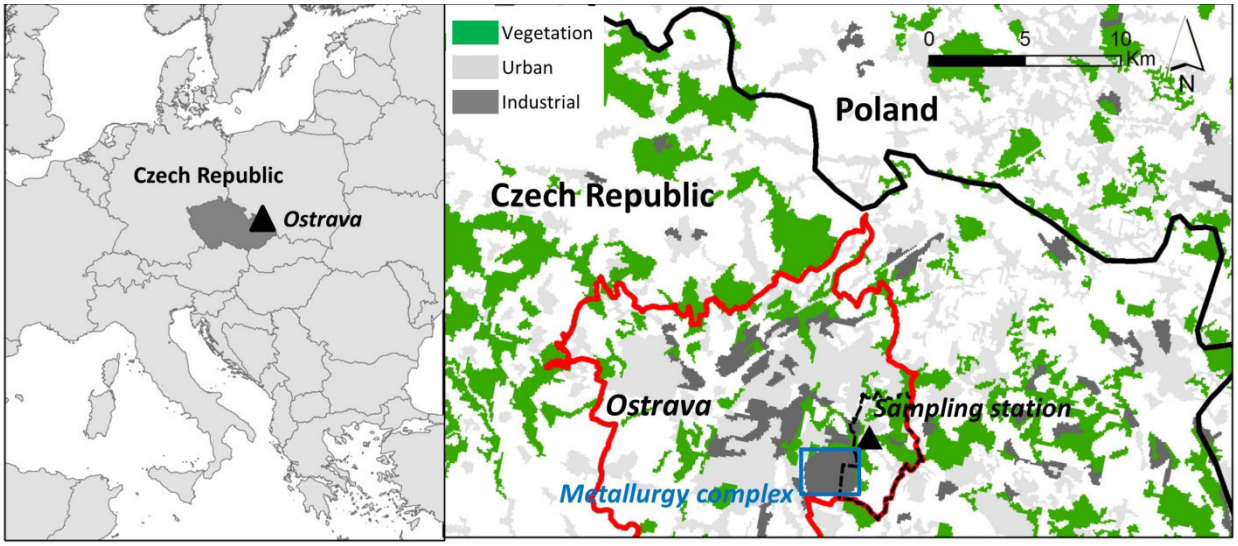


Figure 1. Left: Czech Republic and Ostrava. Right: Sampling station (triangle), close to the metallurgy complex. Background map: Corine Land Cover 2012 version v.18.5.1.

PMF modeling

PMF (USEPA version 5.0) was applied to NSD data (14 nm – 10 µm) and to $PM_{0.09\text{-}1.15}$ chemical composition separately.

For the NSD PMF we tested time resolution from 5 minutes to two-hour and determined 15 minute resolution was optimal as these data are the best compromise to maintain a high time resolution but avoid unwanted noise. Additionally, the input data were handled by merging three consecutive bins, in order to (i) reduce the noise of the raw data, (ii) to decrease the variables number, (iii) to reduce the number of zeroes found in the coarse SMPS and APS bins (Masiol et al., 2016). Since the size segregation of SMPS is based on particle electrical mobility while the APS and 8DRUM impactor use particle aerodynamic properties, the mobility diameter was converted into aerodynamic using a standard ambient aerosol density (1.5 g cm^{-3}, Hinds, 1999). The final matrix had 2905 rows (samples) and 42 columns (bins/variables, Table S2). The uncertainties were calculated according to Vu et al., 2015. The total variable was calculated summing all the bins. PMF was run several times using: (i) different uncertainty input matrices, different C_3 value (Vu et al., 2015) in order to obtain the highest S/N ratio and the Q_{true} closest to $Q_{theoretical}$; (ii) different extra modeling uncertainty; (iii) different number of factors.

2-hours integration time $PM_{0.09-1.15}$ chemical composition was used as input for PMF. The five size ranges of the chemical composition were modeled first separately, however the PMF output of the $PM_{0.09-1.15}$ chemical composition was adequately informative. The analytical uncertainties of TC were multiplied by 4. The data matrix was composed of 2-hours $PM_{0.09-1.15}$ mass concentrations calculated as a sum of the strip weights by beta gauge balance, corresponding elemental composition for 24 elements and $PM_{2.5}$ TC. The final matrix had 217 rows (samples) and 26 columns (species/elements). Mg, P, V, Ni, Ga and Rb variables were set as bad; Na, Al, Se and TC and PM as total variable were classified as weak.

Polar plots and daily pattern were obtained using the R Openair Package (Carslaw and Ropkins, 2012).

Results and discussion

Campaign overview

The average PM_{10} concentration during the sampling campaign was 47±29 μg m^{-3}, with peaks up to 286 μg m^{-3}, registered with south-west WD. $PM_{2.5}$ mass concentration, recalculated from merged SMPS and APS number concentration, was on average 37±20.4 μg m^{-3}. The TC content in $PM_{2.5}$ was on average 16.3±14.2 μg m^{-3}. The gaseous pollutant campaign averages for NO_2, NO, SO_2, and CO were 28.3 μg m^{-3}, 12.1 μg m^{-3}, 4.4 ppb, 0.7 ppm, respectively (Fig. S2).

Nanoparticles (14-30 nm), Aitken mode particles (30-80 nm), accumulation mode particles (80 nm – 1 μm) and coarse particles (1-10 μm) concentrations were respectively 3.5 $\times 10^3$, 6.4×10^3, 6×10^3, and 3.7 particles cm^{-3} (Fig. S2) Winds from the third quadrant of the wind-rose (Fig S4) were the most frequent (180°- 270°deg. south-west direction, 32% of the campaign) and had the highest WS (average: 2.1 $m.s^{-1}$). The northeast wind was the second most frequent (10% of the campaign), with an average WS of 1.6 $m.s^{-1}$. For a detailed description of the UFPs number concentration and size distribution dynamics refer to Leoni et al., 2016.

Sources from NSD data

The model was run several times in order to find the most physically meaningful result and the best diagnostics. Four variables showed high scale residuals in the preliminary runs and they were marked as weak and their uncertainty tripled (18.5 nm, 20.6 nm, 0.47 μm and 5 μm). For 18.5 nm, 20.6 nm, the high number of residuals is likely due to the nature of the nanoparticles characterized by fast changes in NSD. The model was run with different factor number (4 to 9); the most stable solution was found for a 6-factor solution. Extra modeling uncertainty (8.5%) was included to encompass errors not considered in the uncertainty assessment. All run converged, the scales residuals were normally distributed, no unmapped factors were detected with bootstrap error estimation, no swaps were observed with displacement error analysis, indicating that the solution has no data error and it is well defined (Table S1). Solution having <6 factors returned unresolved sources, a high number of scaled residuals >3, and Q_{true} higher than $Q_{theoretical}$. When more than 6 factors were selected, a spurious factor appeared. This factor does not have a physical meaning: it lacks any distinct diurnal pattern, it is uncorrelated to other auxiliary variables, and it exhibits very small contribution to particle number concentration (PNC). In addition, solutions with >6 factors returned Q_{true} values well below the $Q_{theoretical}$.

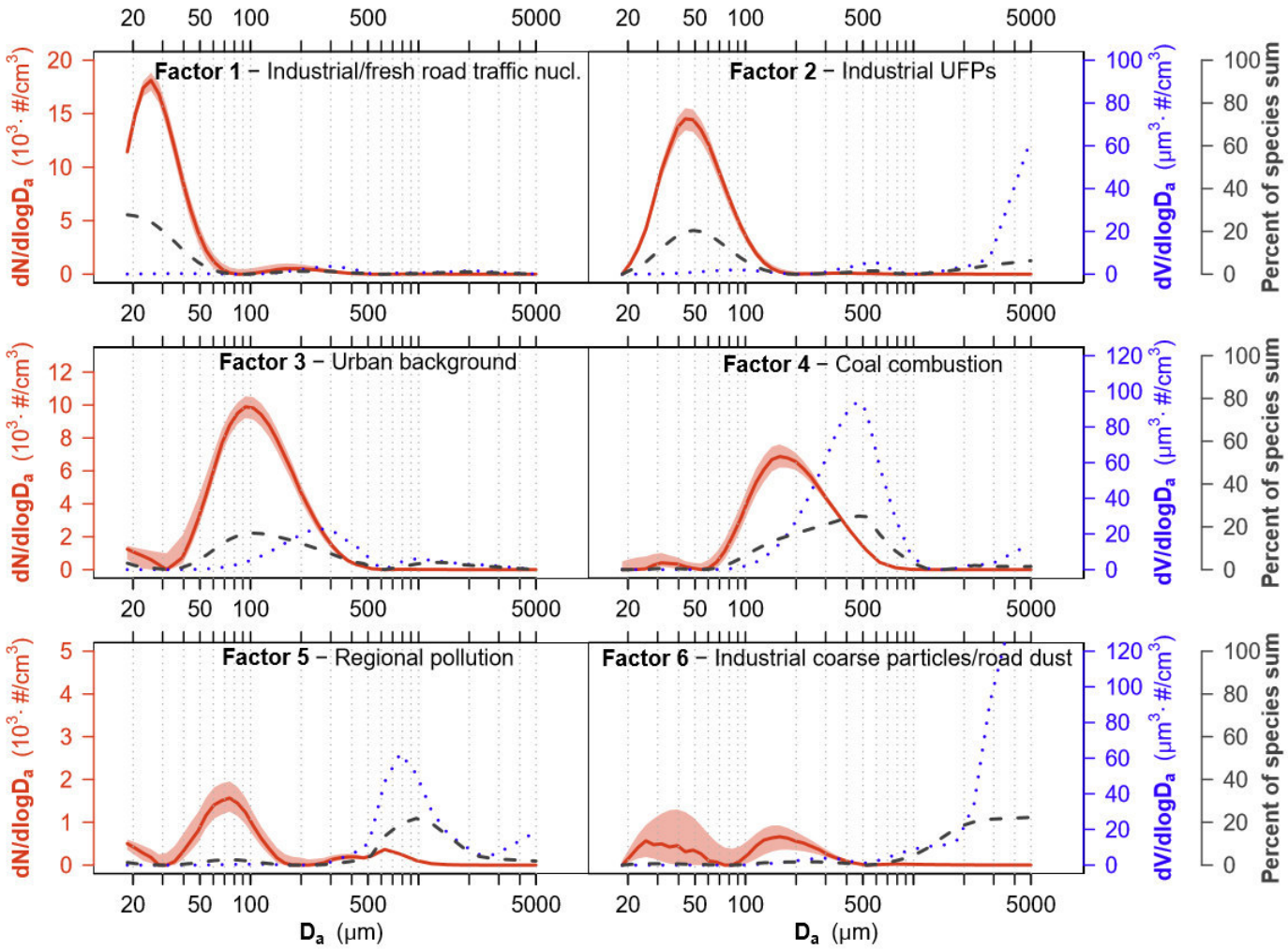


Figure 2. NSD PMF factors. NSD (red line), volume size distribution (blue dotted line) and explained variation (black dashed line). The volume size distribution was re-calculated from the NSD assuming spherical particles.

Factor 1. This factor includes most of the particles in the nucleation range; it exhibits a sharp mode at 26 nm (Fig. 2), with the second largest contribution to PNC (26%). The average concentration of this factor is 4.5×10^3 particles cm^{-3}, but it shows peaks up to 3.4×10^4 particles cm^{-3}. The contribution to the volumetric concentration is low (2%). Photochemical nucleation is excluded because we do not observe new particle formation events with clear particle growth (*banana* shape) during the campaign, and we do not observe higher concentration at noon, when the solar radiation has the highest intensity (Leoni et al., 2017). Also, the high pollution level during the campaign is not a favorable condition for nucleation events due to high coagulation rate caused by high number of accumulation mode particles. For winds blowing from south-west, nanoparticles concentration increases with the WS (Fig.

3). The factor shows positive Spearman correlation with WS, NO_2, SO_2 (Fig. 3). The polar plots show the same area of origin for factor 1 and for the gaseous pollutants SO_2 and CO (Fig.S3), on the south west of the sampling station, where the metallurgy complex is located (Fig.1). However, the daily pattern of factor 1 shows two peaks corresponding to rush hours (8 a.m. and 6 p.m.), suggesting road traffic emissions (Fig. 3). Peaks of factor 1 nanoparticles are also observed at the same time with peaks of NO (Fig. S5). These peaks are attributable to fresh road traffic emissions because NO is a pollutant primarily emitted by road traffic. The nanoparticles peaks, simultaneously with NO peaks, are observed on the days 5th, 21st, 24th, 26th and 27th, with a low WS (<1 ms^{-1}). The possible source of traffic nanoparticles can be the district's main roundabout, located at 130 m north from sampling location, or the street located 60 m in south-west direction. However, a statistically significant difference between weekdays and weekends is not observed (Nonparametric Kruskal Wallis test, p value >0.05). Industrial emissions and a minor contribution of fresh road traffic emissions can represent the sources of factor 1 nanoparticles. The metallurgy industry has been identified as a source of nanoparticles (Riffault et al., 2016, Leoni et al., 2016). Emissions of stationary combustion sources such as industrial stack plumes can contain high nanoparticles concentrations (Dall'Osto et al., 2008; Marris et al., 2012; Leoni et al., 2016). For example, Marris et al. (2012) observed particles with size 10-30 nm downwind of a large metallurgy facility in northern France.

Other studies reported multi-source factors in the ultrafine size range (Gu et al., 2011; Masiol et al., 2016). Nanoparticles can originate from various sources, and it is sometimes not possible to attribute to this factor only to one source. Even with varying the number of factors in the PMF modeling, this factor appeared in all the solutions, demonstrating its robustness and the lack of artifacts in the PMF solution. The problem in resolving the industrial and the freshly emitted road traffic can be due to the strong emission source represented by industry, close to the receptor. Analogous results are observed other studies having similar circumstances with very large local emissions dominating the PM concentrations (c.f., Owoade et al., 2015).

Factor 2. It is the major contributor to PNC (28%) and a minor contributor to volume concentration (6%), having a prominent mode at 45 nm (Fig. 2). The average concentration of factor 2 is 4.5×10^3 particles cm^{-3} with peaks up to 4.2×10^4 particles cm^{-3}. As for factor 1, the concentration increases with the WS and with south western WD (Fig. 3), where the large metallurgy complex is situated (Fig. 1). Additionally, factor 2 polar plot is very similar to the SO_2 and CO polar plots (Fig. S3). The factor shows positive correlation with SO_2, CO, NO_2

Author's preprint version of the article:

Source apportionment of aerosol particles at a European air pollution hot spot using particle number size distributions and chemical composition

Leoni C, Pokorná P, Hovorka J, Masiol M, Topinka J, Zhao Y, Křůmal K, Cliff S, Mikuška P, Hopke PK.

Published in Environmental Pollution 234 (2018) 145e154

DOI: https://doi.org/10.1016/j.envpol.2017.10.097

The original paper is available at:

https://www.sciencedirect.com/science/article/pii/S0269749117331640

(Fig. 3), gaseous pollutants that are emitted by iron and coke production. The factor does not show a clear daily pattern (Fig. 3), and there is no statistically significant difference in the concentration between the weekdays and the weekends (Nonparametric Kruskal Wallis test, p value >0.05). The auxiliary variables indicate this is the industrial source. Indeed, metallurgy facilities can emit high concentrations of ultrafine particles with this size range (Riffault et al., 2016). In the same location, high concentrations of UFPs were measured with airborne measurement, pointing at steelwork chimney as major source (Leoni et al., 2016). Cheng et al., 2008, measured ultrafine particle size distribution in an iron foundry close to the iron production sector and observed high concentrations of UFPs with a mode at 46 nm. Weitkamp et al., 2005, measured the industrial emissions from a steelwork and observed particles in with 45 nm size, when downwind of the facility.

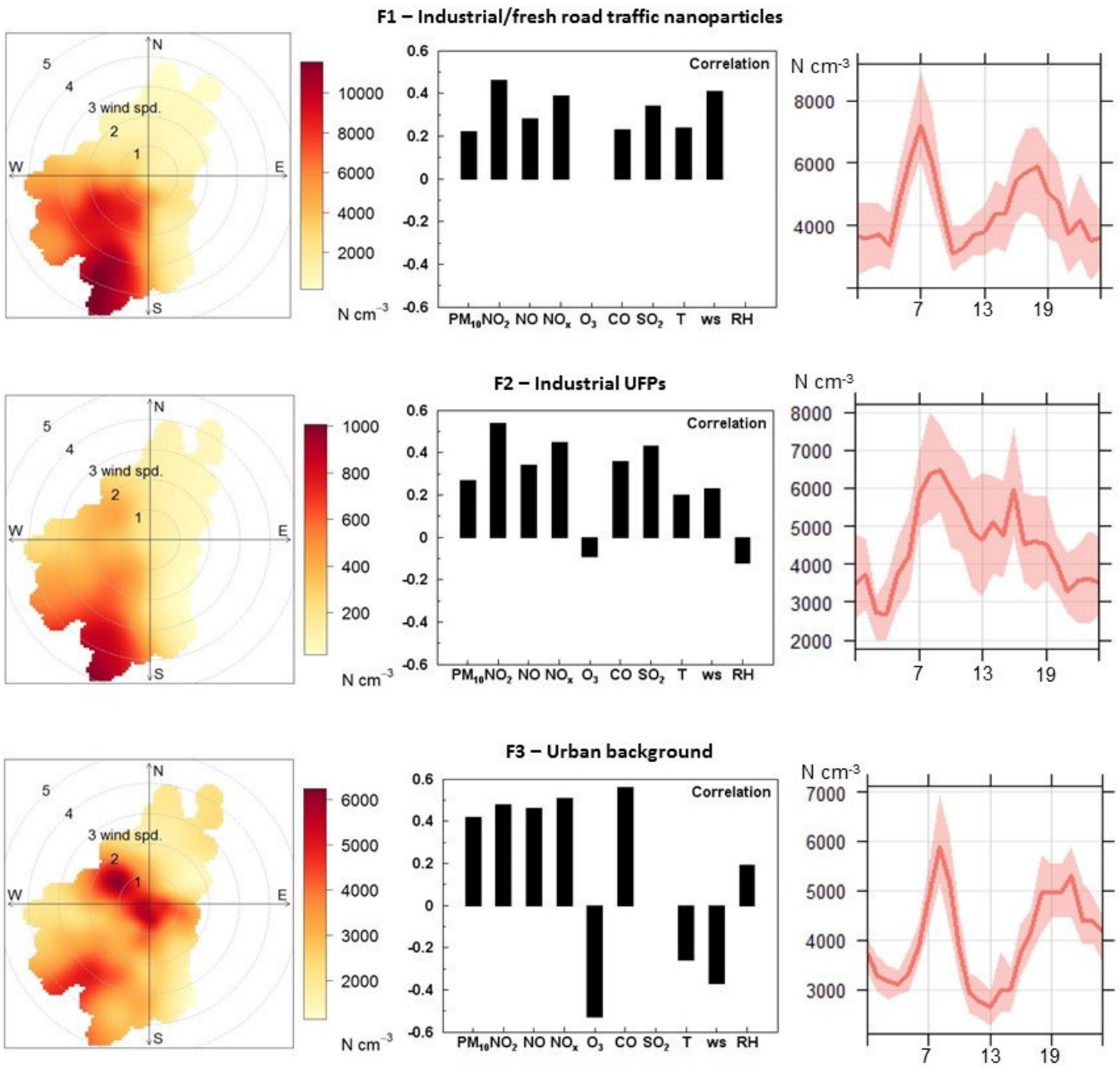


Figure 3. Continued

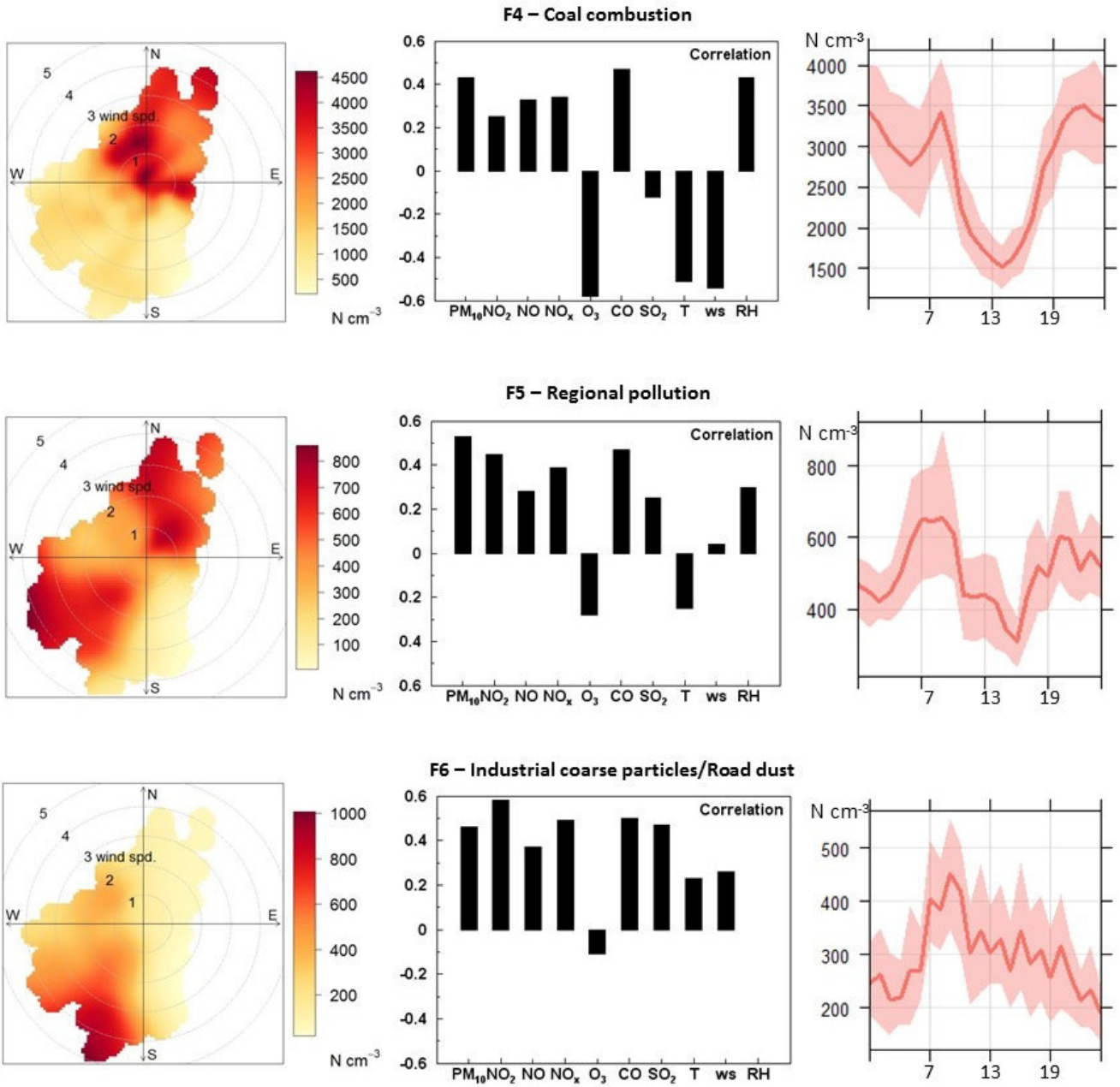


Figure 3. Left: Polar plot with factors concentrations (15-minutes vector averaged WS and WD). Middle: histogram of Spearman correlation coefficients for auxiliary variables (only statistical significant coefficients $\rho<0.05$ are shown). Right: daily pattern of PMF factors, the shaded area represents the 95th% confidence interval.

Factor 3. Includes accumulation mode particles. The size distribution has a predominant mode, with number mode diameter (NMD) at 93 nm and a volume mode diameter (VMD) at 246 nm (Fig. 2). The average concentration was 4×10^3 particles cm^{-3} with peaks up to 2.2×10^4 particles cm^{-3}. This factor is an important contributor to PNC (24%), and a moderate

contributor to the volume concentration (14%). This factor represents the urban background composed by particles originating from combustion sources, such as biomass burning and road traffic. This factor has a positive correlation with NO_x, CO and PM_{10}, and the daily variability reflects the traffic pattern, with a dominant morning peak at 7 a.m., but also we observe a broad peak in the evening from 6 p.m. to 10 p.m. (Fig. 3). The same variability is observed in the weekends. The highest concentrations are observed with low WS (<1 m/s), and with north eastern direction, where the major roundabout of the district is situated (Fig 4). Additionally, lower contribution is observed with south western WD and WS >2 m/s, where single family houses are located (Fig. 3). Factor 3 correlated well (r^2=0.62) with 24-hours levoglucosan concentrations (Fig. S6). Levoglucosan is a tracer for biomass burning (BB). It is not detectable in smoke from coal burning (Simoneit et al., 1999). Particle emissions from BB are dominated by an accumulation mode peaking at 100-150 nm, and occasionally a nucleation mode (Janhall et al., 2010). A small contribution of nanoparticle can be observed in Fig. 3, with a small mode in the ultrafine size range. Similar factor with NMD at 80 nm and at 22 nm was observed by dall'Osto et al., 2012 and attributed to urban background. Also, Vu et al., 2016 observed a similar factor with NMD at 93 nm, attributed to a mixture of carbonaceous particles originated by BB, and solid particles originated from traffic emissions. An urban background factor with the same daily pattern as in this study, consistent with both road traffic and with building heating, was observed by Beddows et al., 2015 and by Masiol et al., 2017. A recent study at the same location (Mikuška et al., 2015) found that levoglucosan is the most abundant compound in the $PM_{2.5}$ organic content, and that BB has a large contribution to the organic fraction of the atmospheric aerosol particles collected in Ostrava.

Factor 4. Includes particles in the accumulation size range, with NMD at 160 nm and VMD at 0.5 μm (Fig. 2). This factor has the highest contribution for the volume concentration (42%), and a moderate contribution to the number (14%). The number concentration average was 2.7×10^3 particles cm^{-3}, with peaks up to 1.3×10^4 particles cm^{-3}. Factor 4 correlates (r^2 = 0.62) with R-17α(H),21β(H)-homohopane, indicating that these particles likely originate from coal combustion (CC) emissions (Fig. S6). The homohopanes are organic compounds present in the smoke of CC but also in the lubricating oils of gasoline and diesel engines. The homohopane index (S/(S+R)) is the ratio between the concentration of R and S isomer of the 17α(H),21β(H)-homohopane and it is used to distinguish CC from traffic emissions and to sort coal according to its maturity. Higher concentrations of R isomer than S isomer, indicates dominant CC emissions, while equal isomer concentration indicates traffic emission being

dominant (Křůmal et al., 2013). In a previous study, Mikuška et al., 2015, found that the concentration of R isomer of 17α(H),21β(H)-homohopane were higher than the concentration of S isomer, indicating that hopanes in Ostrava in winter originated predominantly from CC. In this study, the homohopane index ranged between 0.2 and 0.38 indicating emissions of semi-bituminous and bituminous CC (Oros and Simoneit, 2000).

Even if in the area is impacted by coke production emissions, this factor does not seem to be related to the industrial source. In fact, this factor is related to low WS and north-northeasterly winds (Fig. 3). The time series shows a typical local heating pattern, with a peak in the morning around 8 a.m. and a peak in the evening starting from 7 p.m., with a maximum at 10 p.m. and lasting all night (Fig. 3). This pattern can be also linked to the stagnant atmospheric conditions during winter nights, which trap the emissions in lower altitude. The time series is the mirror image of the air temperature. The positive correlation with NO_x, PM_{10} (Fig. 3) and TC (Fig. S7) indicates combustion sources as well. The highest concentrations of these particles, simultaneously correlated with homohopanes, were registered on the days 12th and 22nd to 25th, with calm winds and low temperature. For example, on February 12 the temperature did not go over 2°C. This factor likely depicts local sources, with a contribution from the north-northwest where an urban settlement with single family houses in located.

Factor 5. This factor contributes to 3% of the number and 15% of the volume size distribution. It shows a VMD at 0.8 μm and a small mode with NMD at 75 nm (Fig. 2), an average concentration of 504 particles cm^{-3}. Higher concentrations, up to 3.8×10^3 particles cm^{-3}, occur with both south-western and north-eastern WDs, which are the two most frequent WDs, and the concentration increases with WS (Fig. 3). This factor shows a diurnal pattern with a broad peak in the morning, from 7 to 11 a.m., a minimum in the afternoon, and a further increase in the evening from 7 p.m. It does not show correlation with organic compounds, but it has a moderate correlation with PM_{10}, NO_2 and CO, (Fig. 3). The source of this factor is more difficult to attribute. Factor 5 concentration increases with the WS and it rises with both south-western and north-eastern WD. This factor can be attributed to regional pollution sources within Moravia-Silesia region (Fig. S11), but also pollution from the nearby Polish Silesia region cannot be excluded (Pokorna et al., 2015, Mikuska et al., 2015). Even if this factor presents the lower contributor to the number concentration, it makes appreciable contribution to the volume and for this reason was included in the solution. Also, it appeared in all the PMF solutions with 5 or more factors. Its exclusion produced a poor fit of the variables 0.8, 0.9, 1.2 and 1.4 μm, with high scaled residuals and ratio $Q_{true}/Q_{theoretical}$ >1.

Similar factor with peak between 0.7 and 3 μm was found by Gu et al., 2011 in Ausburgh, Germany, attributed to long range PM pollution.

Factor 6. It is characterized by super-micrometer mode particles (Fig. 2). This factor makes an appreciable contribution to volume concentration (20%), and the smallest contribution to the number concentration (2%). It shows an average concentration of 296 particles cm^{-3} and peaks up to 3.4×10^3 particles cm^{-3} with increasing WS and south-south western direction. Factor 6 does not have a clear daily pattern. The polar plots of factor 1, 2 and 6 are similar (Fig. 3), pointing on industrial source. This factor can be associated with the unloading and stocking of iron ores and coal for the ironmaking processes, and to the preparatory blending of raw materials. Coarse particles with VMD mode at 8 μm are observed when downwind industrial emissions by Mbengue et al. 2014, in Dunkirk (France). Also, Dall'Osto et al. 2008, found in Port Talbot (UK) the vicinity of large steelwork coarse particles with mode centered at 6 μm.

Higher WS can promote soil and road dust re-suspension (Hinds, 1999). Peaks of factor 6 coarse particles appear simultaneously with peaks of the road dust factor in the mass concentration (Fig. S13), for example on February 19, 20, 25, 28. Gu et al., 2011 observed a peak in the NSD between 2 and 8.8 μm associated with natural crustal dust and resuspended road dust. Masiol et al., 2016, observed a peak in the volume concentration at 5 μm attributed to resuspension of crustal dust.

Sources from chemical composition data and factors associations

To estimate the PMF optimal number of sources, 3 to 7 factor solutions were analysed. The Q values, the resulting source profiles, and the scaled residuals were studied. The optimal factor number was 4 (Fig. S9, S10).

Factor 1. This factor was associated with high concentrations of S, As, Se, Br, K and TC. According to the chemical profile, the factor represented a mix of secondary inorganic aerosol (SIA), CC, and BB. Due to the lack of BB tracer as soluble-potassium (K^+) or 2-hour levoglucosan concentration, the model was not able to resolve the biomass burning factor. The $Q/Q_{expected}>2$ for total-potassium (K) could indicate the missing BB factor. Nevertheless, K should be used with caution as a BB tracer, although it is largely emitted by BB, it also derives from other sources, such as crustal dust, sea-salt, coal usage, waste incinerators and others (Duan et al., 2004; Wang et al. 2007; Caseiro et al. 2009; Mkoma et al. 2013). The time series shows two broad peaks, one in the morning and one in the evening, with a maximum at 10 p.m. reflecting the heating habits (Fig. S8). This factor dominated during the

measurement campaign because it accounted for the 52% of $PM_{0.09–1.15}$ mass. The highest concentrations were registered on the 20th, 21st, 22nd and 24th days, when the WS was predominantly low (< 2 ms^{-1}). Higher concentration is also observed with higher WS and north/north-east WD (Fig. S12).

Factor 2. The second factor road dust from re-suspension and abrasion was represented by high contributions of Al, Si, Ca, Ti, Cu and TC (Han et al., 2005; Cheung et al., 2012), and contributed 22% to the $PM_{0.09–1.15}$ mass. Higher concentrations are observed with WSs between 1 and 3 ms^{-1} from east/north-east where a roundabout of the city district is situated (Fig. S12).

Factor 3. This factor was characterized by Na, Cl, Ca, Zn and TC and it was assigned to the sinter/steel production (Hleis et al., 2013) with a contribution of 16% to the $PM_{0.09–1.15}$ mass. According to the previous studies by Pokorná et al., 2015, 2016 this factor combines two industrial sources. The steel production source of very fine particles ($PM_{0.09-0.26}$) and the sinter production – hot and cold phase source of fine ($PM_{0.15-1.15}$) and coarse particles ($PM_{1.15-10}$), respectively. The high concentration with average WS of 3 m/s and contribution from south-west, points on local industrial source (Fig. S12). In winter with south-west wind episodes, this factor contributes significantly to the PM. This finding is not in agreement with the study by Vossler et al., 2016 due to the choice of the measurement season (10/17 – 6/12 does not response to the cold season) and consequent factor misinterpretation (factor mixed Cl ascribed to the combustion processes as transportation, biomass burning domestic waste, and/or coke; and road dust and road salt).

Factor 4 was ascribed to raw iron production because its profile shows high concentration of metals especially Mn, Fe and Co (Querol et al., 2007; Zhou et al., 2004; Cohen et al., 2010). The highest concentrations were registered when the average WS was 3 m/s. The polar plot matches the factor 3 polar plot (Fig. S12), pointing to local industrial source. This factor contributed 10% to the $PM_{0.09–1.15}$ mass.

Comparison of PMF results on PND and mass

In order to compare the different factors obtained by the separate modeling, the PMF NSD output data were averaged to 2-hour time resolution after the modeling. Although the two datasets reflect in one case the physical and in the second case the chemical characteristics of the atmospheric particles, several factors show associations. NSD PMF factors industrial coarse particles, industrial UFPs, industrial/fresh traffic emitted nanoparticles have a positive correlation with sinter/steel production and iron production factors (Table 1). The CC factor

412 resolved with NSD has moderate correlation with the SIA/CC/BB factor and moderate
413 inverse correlation with iron and sintering/steel production.

	SIA/BB/CC	Iron production	Sinter/steel prod.	RD
Industrial/fresh traffic nanop.	-0.43	**0.65**	**0.35**	0.18
Industrial UFPs	-0.36	**0.4**	**0.37**	0.16
Urban Background		-0.17	0.19	-0.25
Coal Combustion	**0.45**	**-0.51**	**-0.35**	0.03
Regional pollution		0.28		0.10
Industry coarse part./RD	-0.21	**0.51**	**0.34**	0.22

414 *Table 1. Spearman correlation coefficients ρ between PMF factors. Only statistically*
415 *significant correlation coefficients ($p<0.05$) are shown. Coefficients with moderate*
416 *($0.35<r<0.65$) and strong ($p>65$) correlation are in bold.*

417 ## Sources of PAHs

418 PAHs are formed by the incomplete combustion of fossil fuels and wood and are associated
419 with mobile, domestic and industrial emissions. Domestic emissions are predominantly
420 associated with burning of coal, wood and waste. Metallurgy industry emits PAHs mainly
421 from coke production (Ravindra et al., 2007). PAHs source identification is often challenging
422 because they are produced by several sources and depending on the atmospheric conditions,
423 they can partition between gaseous or particulate phase. In Czech Republic, especially in the
424 urbanized areas of the Moravian-Silesian region and of the Silesia Province in Poland
425 ambient concentrations of PM-bound PAH are the highest in Europe (Rogula-Kozłowska et
426 al., 2013; CHMI, 2014; Mikuška et al., 2015; Topinka et al., 2015; Rogula-Kozłowska, 2015;
427 EEA 2016). A study conducted in the same location in winter 2012 reported PAHs
428 concentration in $PM_{2.5}$ during a smog and a post-smog episode: 1223 ng m^{-3} and 224 ng m^{-3},
429 respectively (Mikuška et al., 2015). PAHs can be associated with particles with different size
430 ranges (Topinka et al., 2013). Modes in the mass size distribution for PAHs are typically
431 within the ultrafine and accumulation mode (0.1-1 μm) (Keyte et al., 2013). In a previous
432 study performed by Topinka et al. (2015) in the same locality, PAHs were measured in the
433 size segregated PM with 24 hour time resolution and were found in all the PM size fractions,
434 mostly in the upper accumulation mode (0.5–1 μm). Nevertheless, the scanning electron
435 microscopy suggested that the accumulation mode fraction was mostly composed of UFPs
436 aggregates. This finding was confirmed by our previous study: high concentrations of UFPs
437 highly enriched with PAHs were measured when downwind of the large steelworks (Leoni at
438 al., 2016).

In this study, PAH concentrations measured in PM_1 ranged from 23 to 197 ng m^{-3}, with an average of 66 ng m^{-3}. PAHs showed positive correlation with R-homopane (r^2=0.88) and levoglucosan (r^2=0.67). Correlation is also found between total PAHs and the CC factor resolved with NSD (Fig. 4), with r^2=0.68. Higher PAH concentrations are observed with low WS (1-2 m s^{-1}) and north WD, where the urban settlement with single family houses is situated (Fig. S12).

In summary, according to the previous and present studies, three sources are contributing to PAH pollution at the receptor site, based on WD: industrial emissions with southwesterly wind, BB and CC with north-northeasterly wind. The size of the particles enriched with PAHs is different: in the case of industrial emissions, PAHs enrich the UFPs; in the case of the combustion emissions, PAHs enrich the accumulation fraction. Higher time resolution analysis can give further insight on the sources of PAHs at this location. Highly resolved PAH analysis and comparison with the WS and direction provides information on the variable impact of industrial/combustion emissions, and consequently elucidates the major PAHs source.

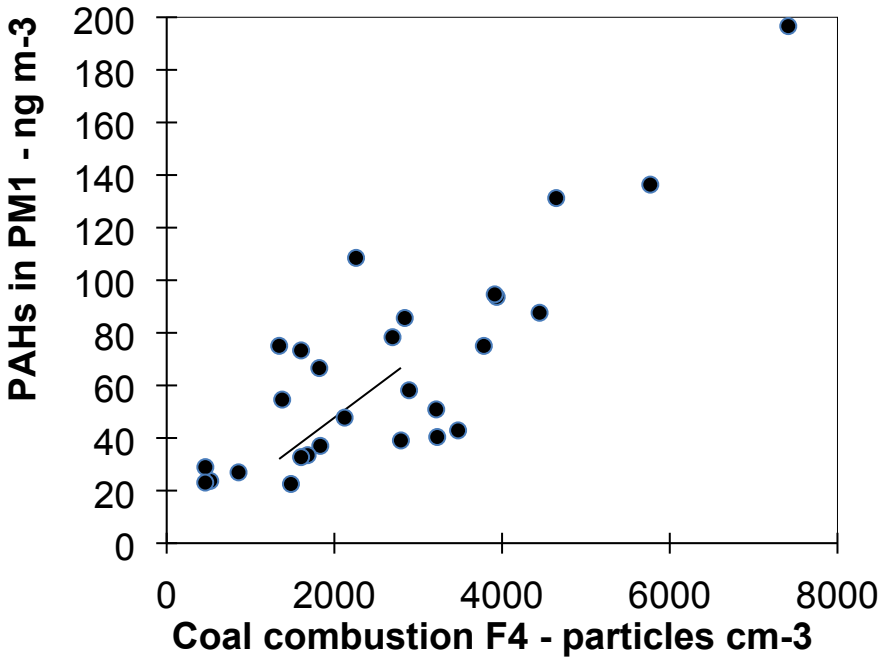


Figure 4. Linear regression of total PAHs (ng m^{-3}) measured in PM_1 and the factor CC (particles cm^{-3}) resolved with NSD.

Conclusions

The aim of this study was to identify the major sources of atmospheric particles in an EU air pollution *hot spot* down to nanoparticles and to give further insight on PAHs sources. Here

460 we find that the industrial source is the major contributor to the PNC at the receptor. The
461 largest contribution to PNC is made by UFPs or nanoparticles (54%). The UFPs
462 concentration increases with WS and south-westerly WD. Additionally, in the ultrafine size
463 range, a minor source is fresh road traffic emissions. This study was not able to make a clear
464 separation and quantification of nanoparticles originated from industry or road traffic. Also, it
465 was not possible to assign a specific industrial process to the different factors in the ultrafine
466 size range (for example coke production, iron production or sintering) due to the complexity
467 of the industrial emissions, the mixture of stationary and fugitive emissions, the continuous or
468 batch processes, the close proximity of processes in the facility, the vicinity of the facility
469 itself to the receptor site. The major sources of particle volume concentration at the receptor
470 are CC and urban background, composed by particles originating from biomass burning and
471 road traffic (56%). The contribution of combustion sources is observed with low WS and
472 north-north easterly WD.
473 Data analysis based on both NSD and chemical composition, elucidated many components in
474 common, as well as other factors which were unique to each method. From this study, we can
475 conclude that the two approaches are complementary. Due to the very small mass of the
476 UFPs, it is often not possible to include them in PMF modeling, of PM sampled with
477 impaction techniques. In this study, the two factors in the ultrafine size range were partially
478 below the lowest cutoff of the 8DRUM impactor (90 nm). On the other hand, the source
479 attribution to the NSD factors alone can be challenging due to the lack of chemical
480 information. For the NSD source identification, the auxiliary variables and the meteorology
481 data are important. Nevertheless, multiple sources can emit gaseous pollutants, for example
482 CO, NO, NO_2, can be originated by combustion, vehicular traffic, industry. A more specific
483 tracer of CC is SO_2, which is not produced by traffic emission, but distinguishing between
484 industrial and coal combustion for heating purposes is often difficult. The organic tracers
485 (levoglucosan and homohopanes) played a crucial role in the sources identification in this
486 study, because they allowed apportioning the sources without ambiguity.
487 Regarding PAHs, the results from this study and from the previous studies were analyzed.
488 Three main sources are contributing to PAH pollution in Ostrava: industrial emissions,
489 combustion of biomass and coal. These sources contribute to different particle size and
490 depend on WS and WD. With southwestern WD, industrial sources are dominant, and with
491 north-north eastern WD the combustion sources play a major role.
492
493

Acknowledgment

This work was supported by the Czech Science Foundation of the Czech Republic under the project P503/12/G147. The research fellowship of one of the authors (P.P.) at the CARES, Clarkson University was funded by the Fulbright Scholar Program. This research used resources of Beamline 10.3.1 of the Advanced Light Source, which is a DOE Office of Science User Facility under contract no. DE-AC02-05CH11231.

Conflicts of interest: none

References

Ancelet T., Davy PK., Mitschell T., Trompetter WJ., Markwitz A. and Weatherburn DC. 2012. Source of particulate matter pollution in a small New Zealand city. Environmental Science & Technology 46 4767-4774. https://doi.org/10.5094/APR.2014.066

Ancelet., T., Davy, PK., Trompetter, WJ., Markwitz, A., WEatherburn, DC., 2014. Sources and transport of particulate matter on an hourly time-scale during the winter in a New Zealand urban valley. Urban Climate 10, 644-655. https://doi.org/10.1016/j.uclim.2014.06.003

Beddows, D.C.S., Harrison, RM., Green, DC, Fuller, GW., 2015. Receptor modeling of both particle composition and size distribution. Atmospheric Chemistry and Physics 15, 10107-10125. https://doi.org/10.5194/acp-15-10107-2015

Birch M.E.Cary R. A., 1996. Elemental carbon-based method for occupational monitoring of particulate diesel exhaust: methodology and exposure issues. Analyst, 121, 1183-1190. https://doi.org/10.1039/AN9962101183

Carslaw D.C., Ropinks. K., 2012. Openair – an R package for air quality data analysis. Environmental Modelling & Software, 27 28, 52-61.

Caseiro A., Bauer H., Schmidl C., Pio C.A., Puxbaum H. 2009. Wood burning impact on PM_{10} in three Austrian regions. Atmospheric Environment 43: 2186–2195. https://doi.org/10.1016/j.atmosenv.2009.01.012

521 Cheng, Y.H., Chao, Y.C., Wu, C.H., Tsai, C.J., Uang, S.N., Shih, T.S., 2008.
522 Measurement of ultrafine particles concentrations and size distribution in an iron foundry.
523 Journal of Hazardous Materials 158, 124-130.
524 https://doi.org/10.1016/j.jhazmat.2008.01.036

525 Cheung K., Schafer M., Schauer J.J., Sioutas C. 2012. Historical trends in the mass and
526 chemical species concentrations of coarse particulate matter in the Los Angeles Basin and
527 relation to sources and air quality regulations. Journal of Air and Waste Management
528 Association, 62, 541–56. http://dx.doi.org/10.1080/10962247.2012.661382

529 Cohen, D. D., Crawford, J., Stelcer, E., Bac, V.T., 2010. Characterisation and source
530 apportionment of fine particulate matter at Hanoi 2001 to 2008. Atmospheric
531 Environment 44, 230 – 328. https://doi.org/10.1016/j.atmosenv.2009.10.037

532 Czech Hydro-meteorological Institute (CHMI), 2017. website, annual graphic overview.
533 http://portal.chmi.cz/files/portal/docs/uoco/isko/grafroc/14groc/gr14e/Obsah_GB.html

534 Czech Hydro-meteorological Institute (CHMI), 2014. Graphic Yearbook.
535 http://portal.chmi.cz/files/portal/docs/uoco/isko/grafroc/14groc/gr14e/Obsah_GB.html

536 Dall'Osto, M., Booth, M.J., Smith, W., Fisherb, R., Harrison, R.M., 2008. A study of the
537 size distributions and the chemical characterization of airborne particles in the vicinity of
538 a large integrated steelworks. Aerosol Science and Technology. 42, 981–991.
539 https://doi.org/10.1080/02786820802339587

540 Dall'Osto, M., Beddows, D. C. S., Pey, J., Rodriguez, S., Alastuey, A., Harrison, R. M.
541 and Querol, X., 2012. Urban aerosol size distributions over the Mediterranean city of
542 Barcelona, NE Spain. Atmospheric Chemistry and Physics, 12, 10693-10707.
543 https://doi.org/10.5194/acp-12-10693-2012

544 Duan M., Zhuang G., Li X., Tao H., Zhuang Y., 2004. The characteristics of
545 carbonaceous species and their sources in $PM_{2.5}$ in Beijing. Atmospheric Environment 38:
546 3443–3452. https://doi.org/10.1016/j.atmosenv.2004.02.052

547 EEA, European Environment Agency, 2016. EEA Report No 28/2016.

548 Elsasser, M., Crippa, M., Orasche, J., DeCarlo, P.F., Oster, M., Pitz, M., Cyry, J.,
549 Gustafson, T.L., Petterson, J.B.C., Schnelle-Kreis, J., Prevot, A.S.H., Zummermann, R.,
550 2012. Organic molecular markers and signature from wood combustion particles in winter
551 ambient aerosols: aerosol mass spectrometer (AMS) and high time-resolved GC-MS
552 measurements in Augsburg, Germany. Atmospheric Chemistry and Physics 12, 6113-
553 6128. https://doi.org/10.5194/acp-12-6113-2012

554 Gu J., Pitz M, Schnelle-Kreis J, Diemer J, Reller A. 2011. Source apportionment of
555 ambient particles: comparison of positive matrix factorization analysis applied to particle
556 size distribution and chemical composition data. Atmospheric Environment 45 1849-
557 1857. https://doi.org/10.1016/j.atmosenv.2011.01.009

558 Han J.S., Moon K.J., Ryu S.Y., Kim Y.J., Perry K.D. 2005. Source estimation of
559 anthropogenic aerosols collected by a DRUM sampler during spring 2002 at Gosan,
560 Korea. Atmospheric Environment. 39, 3113–25.
561 https://doi.org/10.1016/j.atmosenv.2005.01.047

562 Harrison R. M., Beddows D. C. S., Dall'Osto M., 2011. PMF analysis of wide range
563 particle size spectra collected on a major highway. Environmental Science and
564 Technology 45, 5522-5528. https://doi.org/10.1021/es2006622

565 Hinds W. C. 1999. Aerosol technology, Properties, Behavior, and Measurement of
566 Airborne Particles, Wiley (second edition), U.S.A.

567 Hleis, D., Fernández-Olmo, I., Ledoux, F., Kfoury, A., Courcot, L., Desmonts, T.,
568 Courcot, D., 2013. Chemical profile identification of fugitive and confined particle
569 emissions from integrated iron and steelmaking plant. Journal of Hazardous Materials
570 250-251, 246-255. https://doi.org/10.1016/j.jhazmat.2013.01.080

571 Hopke, P.K., 2016. Review of receptor modelling methods for source apportionment.
572 Journal of Air & Waste Management Association. 66, 237-259.
573 http://dx.doi.org/10.1080/10962247.2016.1140693

574 Hovorka J., Pokorná P., Hopke P.K., Křůmal K., Mikuška P. and Píšová M. 2015. Wood
575 combustion, a dominant source of winter aerosol in residential district in proximity to a

large automobile factory in Central Europe. Atmospheric Environment 113, 98-107. https://doi.org/10.1016/j.atmosenv.2015.04.068

Janhall S., Andreae M. O., and Poschl U., 2010. Biomass burning aerosol emissions from vegetation fires: particle number and mass emission factors and size distributions. Atmospheric Chemistry and Physics, 10, 1427–1439. https://doi.org/10.5194/acp-10-1427-2010, 2010.

Keyte I.J., Harrison R.J., Lammel G., 2013. Chemical reactivity and long-range transport potential of polycyclic aromatic hydrocarbons – a review. Chemical Society Reviews 42, 9333-9391. https://doi.org/10.1039/C3CS60147A

Křůmal K., Mikuška P., Večeřa Z., 2013. Polycyclic aromatic hydrocarbons and hopanes in PM_1 aerosols in urban areas. Atmospheric Environment 67, 27-37. https://doi.org/10.1016/j.atmosenv.2012.10.033

Leoni C., Hovorka J., Docekalova V., Cajthaml T., Marvanova S. 2016. Source impact determination using airborne and ground measurements of industrial plumes. Environmental Science & Technology 50, 9881-9888. https://doi.org/10.1021/acs.est.6b02304

Marris, H.; Deboudt, K.; Augustin, P.; Flament, P.; Blond, F.; Fiani, E.; Fourmentin, M.; Delbarre, H. 2012. Fast changes in chemical composition and size distribution of fine particles during the near-field transport of industrial plumes, Science of the Total Environment, 427-428, 126-138. https://doi.org/10.1016/j.scitotenv.2012.03.068

Masiol M., Vu T. V., Beddows D. C. S., Harrison R. M., 2016. Source apportionment of wide range particle size spectra and black carbon collected at the airport of Venice (Italy), Atmospheric Environment 139, 56-74. https://doi.org/10.1016/j.atmosenv.2016.05.018

Masiol, M., Hopke, P.K., Felton, H.D., Frank, B.P., Rattigan, O.V., Wurth, M.J. and LaDuke, G.H., 2017. Source apportionment of $PM_{2.5}$ chemically speciated mass and particle number concentrations in New York City. Atmospheric Environment 148, 215-229. https://doi.org/10.5094/APR.2013.018

603 Mbengue S., Alleman L.Y., Flament P., 2014. Size-distributed metallic elements in
604 submicronic and ultrafine atmospheric particles from urban and industrial areas in
605 northern France. Atmospheric Research 135-136, 35-47.
606 http://dx.doi.org/10.1016/j.atmosres.2013.08.010

607 Mikuška P., Křůmal K., Večeřa Z., 2015. Characterization of organic compounds in the
608 $PM_{2.5}$ aerosol in winter in an industrial urban area. Atmospheric Environment 105, 97-
609 108. https://doi.org/10.1016/j.atmosenv.2015.01.028

610 Mkoma S.L., Kawamura K., Fu P.Q. 2013. Contributions of biomass/biofuel burning to
611 organic aerosols and particulate matter in Tanzania, East Africa, based on analyses of
612 ionic species, organic and elemental carbon, levoglucosan and mannosan. Atmospheric
613 Chemistry and Physics 13, 10325–10338. https://doi.org/10.5194/acp-13-10325-2013

614 Moreno, T., Kojima., T., Amato, A., Lucarelli, F., de la Rosa, J., Calzolai, G., Nava., S.,
615 Chiari, M., Alastuey, A., Querol, X., Gibbons, W., 2013. Daily and hourly chemical
616 impact of springtime transboundary aerosols on Japanese air quality. Atmospheric
617 Chemistry and Physics 13, 1411-1424. https://doi.org/10.5194/acp-13-1411-2013, 2013

618 Oros D.R., Simoneit B.R.T., 2000. Identification and emissions rates of molecular tracers
619 in coal smoke particulate matter. Fuel 79, 515-536. https://doi.org/10.1016/S0016-
620 2361(99)00153-2

621 Owoade K.O., Hopke P.K., Olise F.S., Oundele L.T., Fawole O.G, Olaniyi B.H., Jegede
622 O.O., Ayoola M.A., Bashiru M.I., 2015. Chemical composition and source identification
623 of particulate matter ($PM_{2.5}$ and $PM_{2.5-10}$) from a scrap iron and steel smelting industry
624 along the Ife-Ibadan highway, Nigeria. Atmospheric Pollution Research 6, 107-199.
625 https://doi.org/10.5094/APR.2015.013

626 Paatero, P., 1997. Least squares formulation of robust non-negative factor analysis.
627 Chemometrics and Intelligent Laboratory System 37, 23-35.
628 https://doi.org/10.1016/S0169-7439(96)00044-5

629 Pancras JP., Landis MS., Norris GA., Vedantham R. and Dvonch JT. 2013. Source
630 apportionment of ambient fine particulate matter in Dearborn, Michigan, using hourly

631 resolved PM chemical composition data. Science of the Total Environment 448, 2-13.
632 https://doi.org/10.1016/j.scitotenv.2012.11.083

633 Pokorná, P.; Hovorka, J.; Hopke, P.K. 2016. Elemental composition and source
634 identification of very fine aerosol particles in a European air pollution hot-spot.
635 Atmospheric Pollution Research, 7, 671-679. https://doi.org/10.106/j.apr.2016.03.001

636 Pokorná, P.; Hovorka, J.; Klán, M.; Hopke, P.K. 2015. Source apportionment of size
637 resolved particulate matter at a European air pollution hot spot. Science of the Total
638 Environment, 502, 172-183. https://doi.org/10.1016/j.scitotenv.2014.09.021

639 Polissar A.V., Hopke, P.K., Poirot, R.L., 2001. Atmospheric aerosol over Vermont:
640 chemical composition and sources. Environmental Science & Technology 35, 4604-4621.
641 https://doi.org/10.1021/es0105865

642 Querol X., Viana M., Alastuey A., Amato F., Moreno T., Castillo S., Pey J., Rosa J.,
643 Sánchez de la Campa A. Artínano B., Salvador P., García Dos Santos S., Fernández-
644 Patier R., Moreno-Grau S., Negral. L., Minguillón M.C., Monfort Gil J. I., Inza A.,
645 Ortega L.A., Santamaría J.M., Zabalza J., 2007. Source origin of trace elements in PM
646 from regional background, urban and industrial sites of Spain. Atmospheric Environment
647 41, 7219 – 7231. https://doi.org/10.1016/j.atmosenv.2007.05.022

648 Ravindra K. Sokhi R., Van Grieken R., 2008. Atmospheric polycyclic aromatic
649 hydrocarbons: source attribution, emission factors and regulation. Atmospheric
650 Environment 42, 2895-2921. https://doi.org/10.1016/j.atmosenv.2007.12.010

651 Riffault, V.; Arnds, J.; Marris, H.; Mbengue, S.; Setyan, A.; Alleman, L. Y.; Deboudt, K.;
652 Flament, P.; Augustin, P.; Delbarre, H.; Wenger, J. 2015. Fine and ultrafine particles in
653 the vicinity of industrial activities: a review. Environmental Science and Technology, 45
654 (21), 2305-2356. https://doi.org/10.1080/10643389.2015.1025636

655 Rogula-Kozłowska, W., 2015. PAH and heavy metals in ambient particulate matter: a
656 review of up-to-date worldwide data, in: Pastuszka, J.S. (Eds.), Synergic influence of
657 gaseous, particulate, and biological pollutants on human health. Crc Press-Taylor &
658 Francis Group, FL USA, pp. 68-108. https://doi.org/10.1201/b19592-5

Rogula-Kozłowska, W., Kozielska, B., Klejnowski, K., 2013. Concentration, origin and health hazard from fine particle-bound PAH at three characteristic sites in southern Poland. Bulletin of Environmental Contamination and Toxicology. 91 (3), 349-355. https://doi.org/10.1007/s00128-013-1060-1

Simoneit B.R.T., Schauer J.J., Nolte C.G., Oros D.R., Elias V.O., Fraser M.P., Rogge W.F., Cass G.R., 1999. Levoglucosan, a tracer for cellulose in biomass burning and atmospheric particles. Atmospheric Environment 33, 173-182. https://doi.org/10.1016/S1352-2310(98)00145-9

Sowlat M. H., Hasheminassab S., Sioutas C., 2016. Source apportionment of ambient particle number concentrations in central Los Angeles using positive matrix factorization (PMF). Atmospheric Chemistry and Physics. 16, 4849-4866. https://doi.org/10.5194/acp-16-4849-2016

Šrám, R. J.; Dostál, M.; Libalová, H.; Rossner, P.; Rossnerová, A.; Svecová, V.; Topinka, J.; Bartonová A. The European hot spot of B[a]P and $PM_{2.5}$ exposure-the Ostrava region, Czech Republic: health research results. ISRN Public Health, 2013a, article ID 416701. https://doi.org/10.1155/2013/416701

Topinka, J., Rossner, P., Milcová, A., Schmuczerová, J., Penčíková, K., Rossnerova, A., Ambrož, A., Štolcpartová, J., Bendl., J., Hovorka, J, Machala, M., 2015. Day-to-day variability of toxic events induced by organic compounds bound to size segregated atmospheric aerosol. Environmental Pollution, 202, 135-145. https://doi.org/10.1016/j.envpol.2015.03.024

Venkatamaraan C., 1999. Comparison of particle lung doses from the fine and coarse fractions of urban PM_{10} aerosols. Inhalation Toxicology, 11:2, 151-169, https://doi.org/10.1080/089583799197221

Vossler, T., Černíkovský, L., Novák, J., Williams, R., 2016. Source apportionment with uncertainty estimates of fine particulate matter in Ostrava, Czech Republic using Positive Matrix Factorization. Atmospheric Pollution Research 7, 503-5012. https://doi.org/10.1016/j.apr.2015.12.004

687 Vu T., Delgado Saborit J. M., Harrison R. M., 2015. Review: particle number size
688 distribution form seven major sources and implications for source apportionment studies.
689 Atmospheric Environment 122, 144-132. https://doi.org/10.1016/j.atmosenv.2015.09.027

690 Vu, T. V., Beddows, D. C. S., Delgado-Saborit, J. M. and Harrison, R. M., 2016. Source
691 apportionment of the lung dose of ambient submicrometre particulate matter. Aerosol and
692 Air Quality Research 16(7), 1548-1557. https://doi.org/10.4209/aaqr.2015.09.0553

693 Wang Q., Shao M., Liu Y., William K., Paul G., Li X., Liu Y., Lu S., 2007. Impact of
694 biomass burning on urban air quality estimated by organic tracers: Guangzhou and
695 Beijing as cases. Atmospheric Environment 41: 8380–8390.
696 https://doi.org/10.1016/j.atmosenv.2007.06.048

697 Weitkamp, E. A.; Lipsky, E. M.; Pancras, P. J.; Ondov, J. M.; Polidori, A.; Turpin, B. J.;
698 and Robinson, A. L., 2005. Fine particle emission profile for a large coke production
699 facility based on highly time-resolved fence line measurements. Atmospheric
700 Environment, 39 (36), 6719–6733. https://doi.org/10.1016/j.atmosenv.2005.06.028

701 Zhou, L., Hopke, P. K., Paatero, P., Ondov, J. M. , Pancras, J. P., Pekney N. J.,
702 Davidson, C. I., 2004. Advanced factor analysis for multiple time resolution aerosol
703 composition data. Atmospheric Environment 38, 4909-4920.
704 https://doi.org/10.1016/j.atmosenv.2004.05.040

Supplementary material for:

Source apportionment of aerosol particles at a European air pollution hot spot using particle number size distributions and chemical composition

Cecilia Leoni, Petra Pokorná, Jan Hovorka, Mauro Masiol, Jan Topinka, Yongjing Zhao, Kamil Křůmal, Steven Cliff, Pavel Mikuška, Philip K. Hopke

Number of Figures: 13

Number of Tables: 2

Number of pages: 18

PMF analysis

In aerosol science, receptor models are mathematical approaches used for identifying and quantifying different atmospheric particles sources. Specifically, PMF model identifies and quantifies the contribution to each source to the samples. The identification is based on key species (tracers), combination of species, or size distribution characteristics.

The fundamental principle of PMF is the mass conservation and the goal of PMF is to solve the chemical mass balance equation. The main physical constraint is that the source composition must be non-negative. In fact, a source does not emit a negative quantity of pollution, or it becomes a sink. Also, the sum of the predicted contributions for each source must be less than or equal to the total measured mass for each variable. PMF performs a matrix factorization, which consists of decomposing a matrix of sample data X into two sub-matrices. The two sub-matrices contain the species profile for each factor (G), and the amount of mass/number contributing by each factor to each individual sample (F), plus the unexplained part E (difference between measured and calculated species concentration) (Brown et al., 2015).

$$X = GF + E$$

When using PMF for source apportionment, there are many parameters involved in the determination of the factor matrices, and the change of one of them may lead to a different solution. To find the optimal solution, several runs with different conditions must be made. In the following section, the strategy applied for the PMF analysis is explained, focusing on the evaluation of the diagnostics provided by the model. The interpretation of the PMF results and the choice of the best solution is the most crucial step that must be done by the operator. All the diagnostics must be evaluated and the best solution is a trade-off between the physical meaning of the solution, the diagnostics of the model and the goodness of the fit.

First step: preparation and analysis of the input data. The PMF input data consist of two matrices: concentration and uncertainty. The matrices must have the same dimensions and must not contain missing data or negative/zero values. The data matrices of the mass chemical composition were prepared in compliance with the procedure described in Polissar et al., 1998. For the NSD uncertainty U, this method was used (Vu et al., 2015):

$$\sigma_{i,j} = \begin{cases} \alpha\left(N_{i,j} + \overline{N}_j\right) & N_{i,j} > 0 \\ 2\overline{N}_j & N_{i,j} = 0 \end{cases}$$

$$U = \sigma_{i,j} + C_3(N_{i,j})$$

σ the error for every size bin *i* and measurement *j*. N*i,j* is the particle number concentration for size bin *i* and measurement *j*. α is an arbitrary constant 0.01. The missing data were replaced with the average particle number concentration $\overline{N}$ and the related uncertainties were assumed to be three times $\overline{N}$. C_3 is a constant value ranging between 0.01 and 0.5, which should be chosen according to the distribution of the scaled residuals and the of uncertainty/concentration ratio of 10-20%. The distribution should be between +2 and -2. If the number of scales residuals >3 is high, it can be necessary to increase the uncertainty or to change the number of factors. In this study C_3 was chosen as 0.08. The calculated uncertainty/concentration ratio ranged from a minimum of 10% and a maximum of 17% in average.

A total of 138 missing data (4.7% of the dataset) were excluded from the modeling. 122 data were missing due to data transfer problems and a power outage that occurred on March 4th; 16 data were erased as outliers because of very high pollution peaks.

When the matrices are ready, it is necessary to upload them in the model and a crucial step is to decide if a variable is strong/weak/bad. First, to understand if a variable is strong/weak/bad, a detailed knowledge of the dataset, of the sampling and the analytical techniques is needed. Than the signal to noise ratio (S/R) must be observed. A variable is weak if it has signal and noise in comparable amounts (S/R=1). Similarly, variables having more noise than signal are termed bad variables (Reff et al., 2007). Downweighing the weak variables by a factor 3 is recommended, to protect in the case that the error level of some variables has been underestimated. Regarding bad variables, where hardly a signal is visible from the noise, the recommendation is to omit these variables, or to strongly downweight by a factor 5 or 10 (Paatero and Hopke, 2003). It is important to know whether a variable can be retained in the analysis. For the particle number size distribution, the weak variables that were excluded from previous studies are typically the first and last SMPS size bins, and the last APS bins, due to the high variability and the presence of zeroes. Lastly, it is necessary to observe the time series of very variable, to determine whether temporal patterns are present and if there are unusual events or outliers. Extreme events can be

85 excluded from the modeling if they perturb the solution, they cause high scale residuals number,
86 they generate spurious factors.
87 *Second step: base model execution*. The model must be run with different number of factors and
88 for each run diagnostics must be evaluated. The choice of the factors number is crucial. Too few
89 factors lead to the combination of sources, while too many factors split one source into two non-
90 existing sources. To understand the correct number of factors is necessary to observe the Q
91 value. The $Q_{theoretical}$ gives information about the quality of the fit since the optimal solution
92 should have a Q_{true} not too different from the $Q_{theoretical}$. The ratio $Q_{true}/Q_{theoretical}$ must be around
93 1. In some cases, the optimal solution does not fully satisfy this requirement, for example when a
94 dataset contains a lot of weak variables. The $Q_{theoretical}$ in this study was calculated according to
95 the EPA PMF 5.0 User Guide. For NSD PMF, $Q_{true}/Q_{theoretical}$ was 0.998.
96 The second diagnostic to analyze are the scaled residuals. If the input data are correct, the plot of
97 scaled residuals values against their frequency should be normally distributed, with the majority
98 between -3 and +3. If some residuals are outside this region, it is possible that there are extreme
99 events or outliers in the data, or the number of factors is not optimal. Distribution with large
100 spread indicates that the uncertainty is too low, while a distribution centered near zero indicates
101 that uncertainty are too high. Third step is to observe the goodness of fit, comparing the
102 predicted species with the original concentrations, using regression and time series. The
103 observed/predicted regression coefficients must be as close as possible to 1 and the time series of
104 modeled variables should fit as best as possible the time series of the original variable (Fig. S1).
105 Table S1: Summary of PMF diagnostics for NSD and chemical composition modeling.

Diagnostic	**NSD**	**Chemical composition**
N. of observations	2905	120
Missing values	4.7%	10%
Qexpected/theoretical	85028	2327
Qtrue	85216	2641
Qrobust	84866	2640
Species with Q/Qexpected>2	18.5 nm	K
DISP swaps	0	0
BS mapping	100%	92 – 100%

106

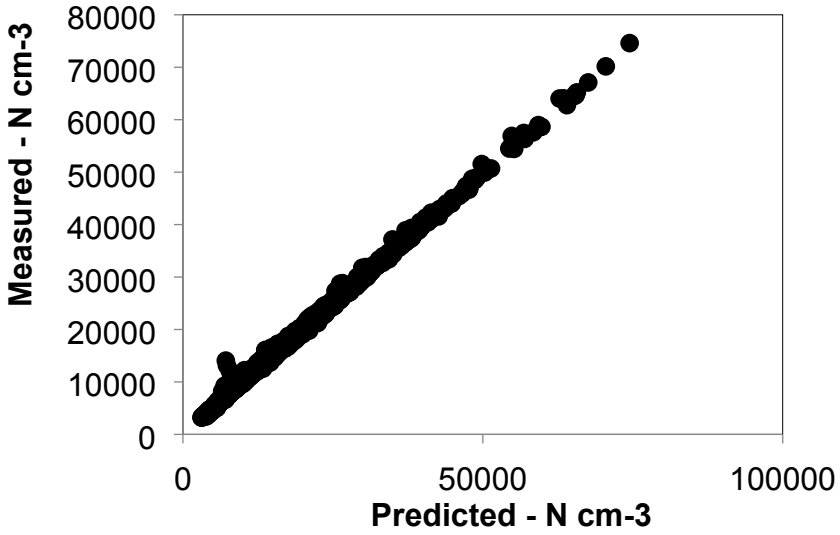


Figure S1. The linear regression between measured/ predicted values.

Lastly, the G-plots must be observed. The G-plots show scatter plots of one factor versus another factor are they are very important to verify the relationship between the factors. The G plots are generated with the assumption that the determined factors are uncorrelated between each other. If the factors are independent one of another, the points should fill the all scatter plot space. The more stable solution will have many samples with zero contribution on both axes. Imagining straight lines passing through the origin of axes including all the points between them, the lines should be as much close to the Cartesian axes. If the points are not aligned with the Cartesian axes, an *edge* can be observed in the plot, which means there might be an overlooked factor or there is rotational ambiguity in the solution.

Third step: error evaluation wit bootstrap (BS) and displacement (DISP). A number of phenomena can contribute to the uncertainty of the solution modeled by PMF: temporal variation of PM sources, measurements errors, sampling variability, errors in the model, for example wrong factor number (Reff et al., 2007). The technique of bootstrapping is used to check the validity of the solution. Bootstrapping consist of randomly selecting n samples to create new datasets and executing the PMF on each new dataset. Multiple PMF solutions are generated using the series of data that are resampled version of the original dataset. Several hundreds of bootstrapped samples can be used, and the summary statistics are calculated by the model. In this study for example, 100 BS runs were performed for both NSD and chemical composition

solutions. If the same sources are identified in most of the bootstrapped samples, the solution of the original dataset can be considered stable. Mapping over 80% of the factor indicates that the number of factor can be appropriate (Table S1).

The DISP explores the rotational ambiguity in a PMF solution by assessing the largest range of source profile values without appreciable increase in the Q value. Each fitted element in the factor profile matrix in the base PMF solution is *displaced* from its fitted value far enough so that the Q increases by a predetermined amount called Q_{max}. The variables are *perturbed* within a certain interval, which has a certain extension (the boxplot in the results diagnostics). If the variable is perturbed so much that the factor changes its identity, a swap occurs. If a swap occurs, it means that the solution is not robust enough and another solution might be a better choice.

References

1. Brown S.G., Eberly S., Paatero P., Norris G., A., 2015. Methods for estimating uncertainties in PMF solutions: examples with ambient air and water quality data and guidance reporting PMF results. Science of the Total Environment, 518-519 626-635. https://doi.org/10.1016/j.scitotenv.2015.01.022
2. Paatero P., Hopke P.K., 2003. Discarding or down weighting high-noise variables in factor analytic models. Analytica Chimica Acta 490, 277-89. https://doi.org/10.1016/S0003-2670(02)01643-4
3. Polissar A.V., Hopke P.K., Paatero P., Malm W.C., Sisler J.F., 1998. Atmospheric aerosol over Alaska-2. Elemental composition and sources. Journal of Geophysical Research, 103, 19045-19057. https://doi.org/10.1029/98JD01212
4. Reff A., Eberly S. I., Bhave, P. V. 2007. Receptor modeling of ambient particulate matter data using Positive Matrix Factorisation: Review of existing methods. Air & Waste management Association, 57:146-154. http://dx.doi.org/10.1080/10473289.2007.10465319
5. US Environmental Protection Agency PMF 5.0 User Guide – Norris G., Duvall R.
6. Vu T., Delgado Saborit J. M., Harrison R. M., 2015. Review: particle number size distribution form seven major sources and implications for source apportionment studies, Atmospheric Environment 122, 144-132. https://doi.org/10.1016/j.atmosenv.2015.09.027

158 Table S2. Upper/lower and average diameter of the NSD size ranges, used as input variables in
159 the PMF model. In bold is the overlap size range between SMPS and APS.

Variable n.	Size range		PMF input variable
	Lower midpoint diameter	Upper midpoint diameter	Average diameter
1	17.8	19.2	18.5
2	19.9	21.4	20.6
3	22.1	23.8	22.9
4	24.6	26.5	25.5
5	27.5	29.4	28.4
6	30.5	32.8	31.6
7	34.0	36.6	35.3
8	37.9	40.7	39.3
9	42.2	45.4	43.8
10	47.0	50.5	48.8
11	52.3	56.2	54.3
12	58.3	62.7	60.5
13	65.0	69.9	67.4
14	72.5	77.8	75.1
15	80.6	86.6	83.6
16	89.9	96.5	93.2
17	100.0	107.6	103.8
18	111.5	119.8	115.6
19	124.2	133.5	128.8
20	138.3	148.7	143.5
21	154.1	165.7	159.9
22	171.7	184.5	178.0
23	191.3	205.6	198.4
24	213.1	229.0	221.0
25	237.4	255.1	246.2
26	264.4	284.1	274.2
27	294.5	316.6	305.5
28	328.2	352.6	340.3
29	365.5	392.8	379.1
30	407.2	437.6	422.3
31	453.6	487.4	470.4
32	**505.3**	**542.0**	**523.7**

33	583.0	673.0	627.3
34	723.0	835.0	778.3
35	898.0	1037.0	966.7
36	1114.0	1286.0	1199.0
37	1382.0	1596.0	1488.0
38	1715.0	1981.0	1846.3
39	2129.0	2458.0	2291.7
40	2642.0	3051.0	2844.0
41	3278.0	3786.0	3529.0
42	4068.0	10370.0	7219

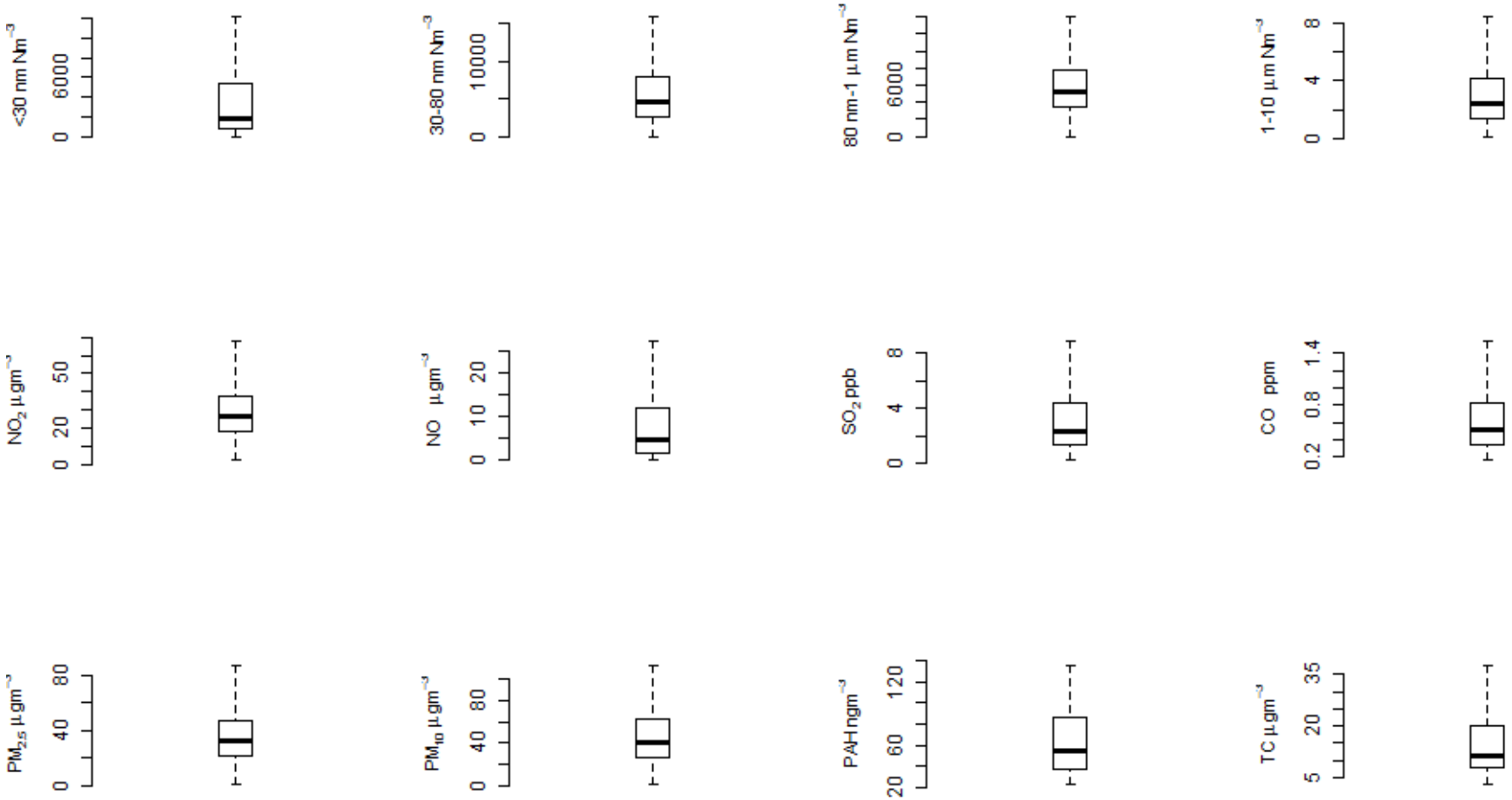
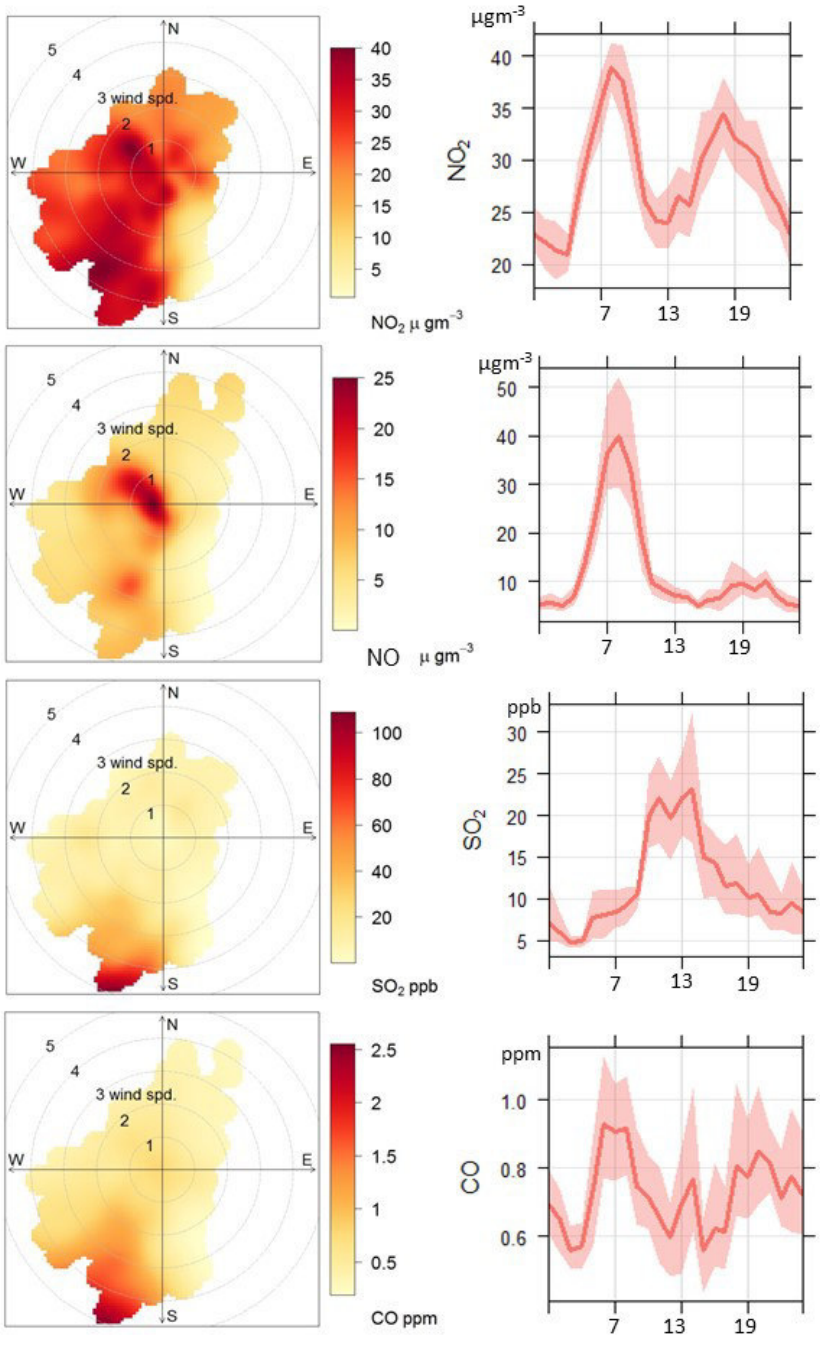


Figure S2. Boxplots of PNCs, gaseous pollutants, $PM_{2.5}$, PM_{10}, PAHs and TC. Middle line represents the median, the box is the inter quartile range, the whiskers indicate 1.5 inter quartile range.



164

165

166

167

168

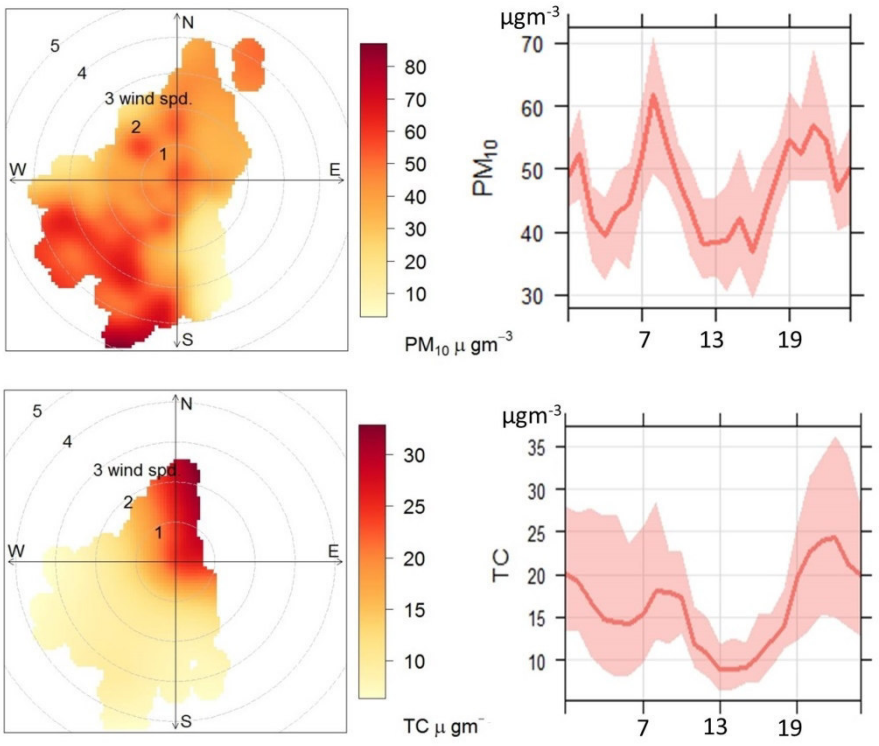


Figure S3. Polar plots and daily patterns of PM_{10}, gaseous pollutants (NO, NO_2, SO_2, CO) and TC, shaded area in the daily pattern plot represents the 95th % confidence interval.

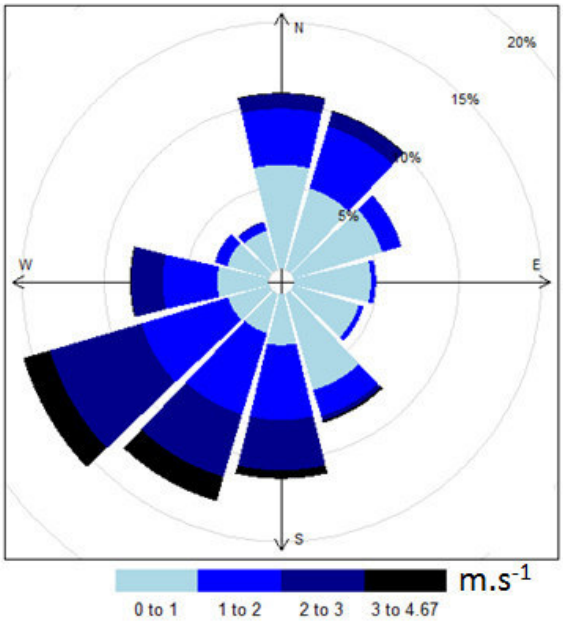


Figure S4. Wind rose for the whole sampling campaign (February 5 to March 7, 2014).

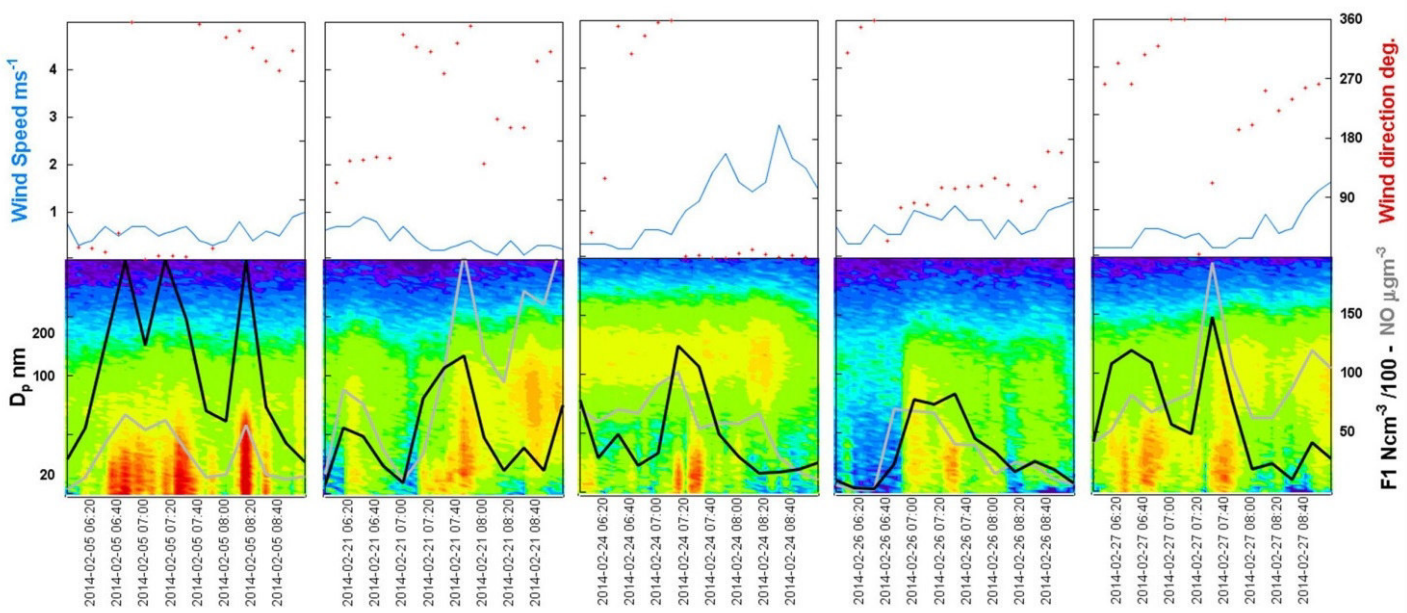


Figure S5: Time series for wind speed and direction (top) and particle number size distribution (bottom). Black line indicates concentration of industrial/traffic fresh emitted nanoparticles (factor 1) and the grey line of NO.`

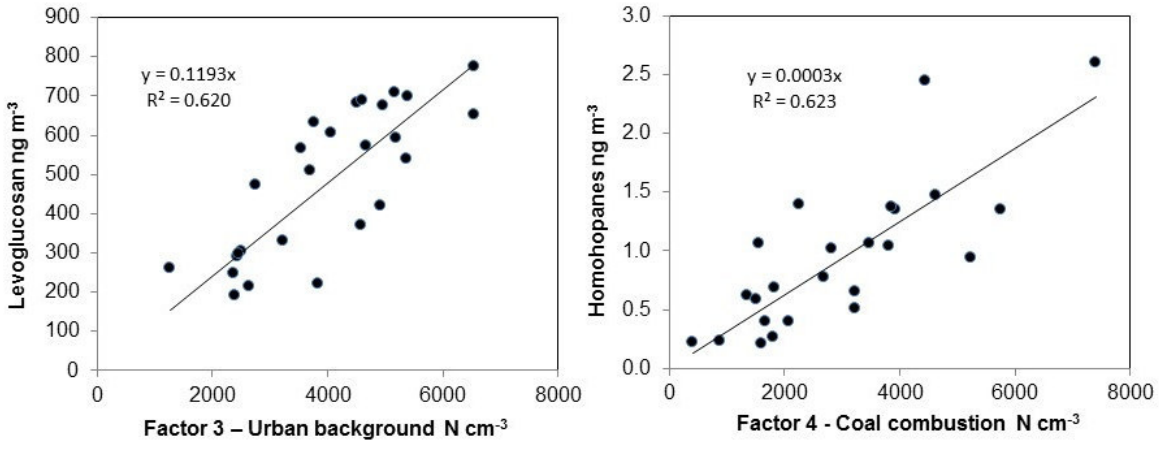


Figure S6. Left: linear regressions between 24-hours concentration of levoglucosan and Factor 3 -urban background PNC. Right: linear regression of 24-hours averages of R homohopane and Factor 4 - CC PNC.

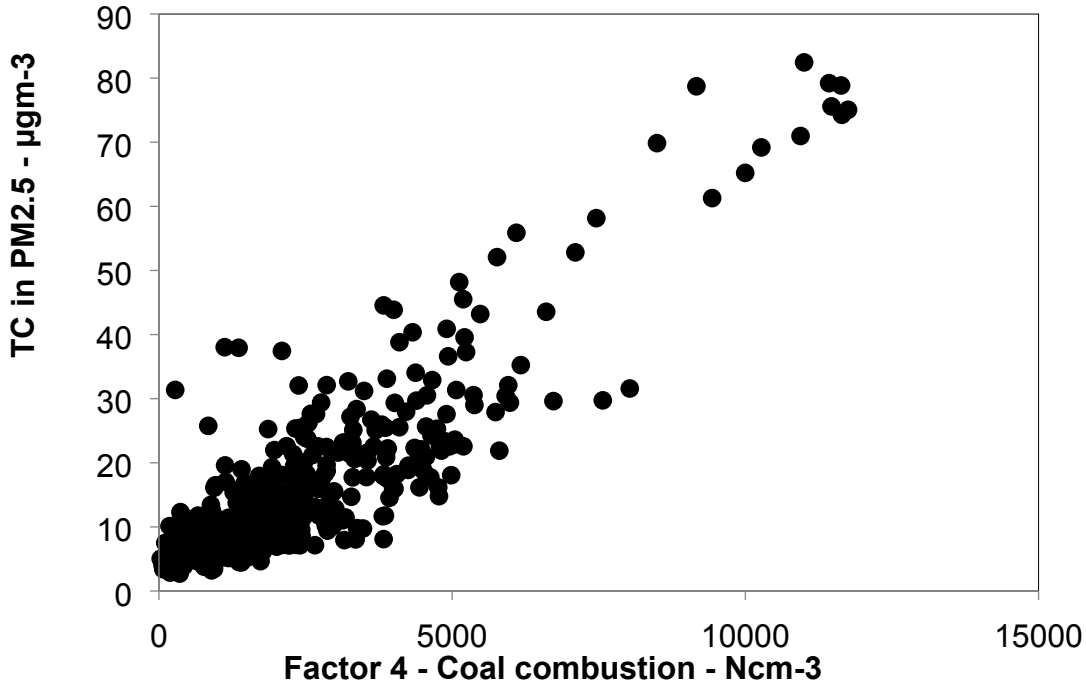


Figure S7. Linear regression between total carbon and NSD factor 4 – coal combustion.

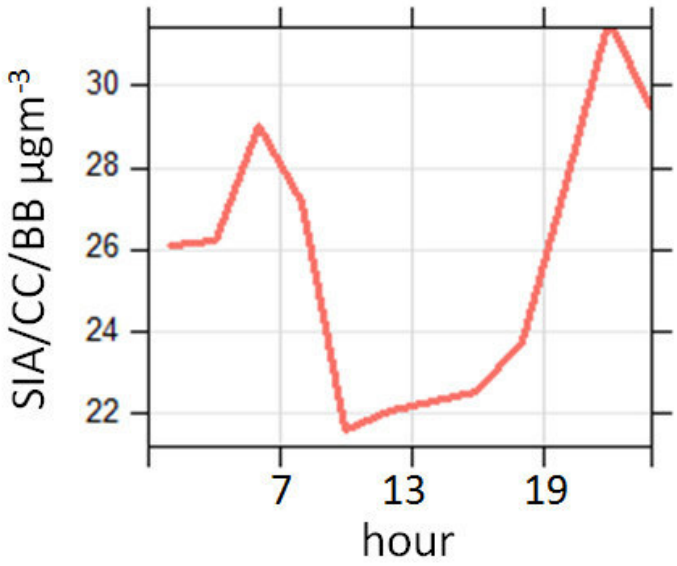
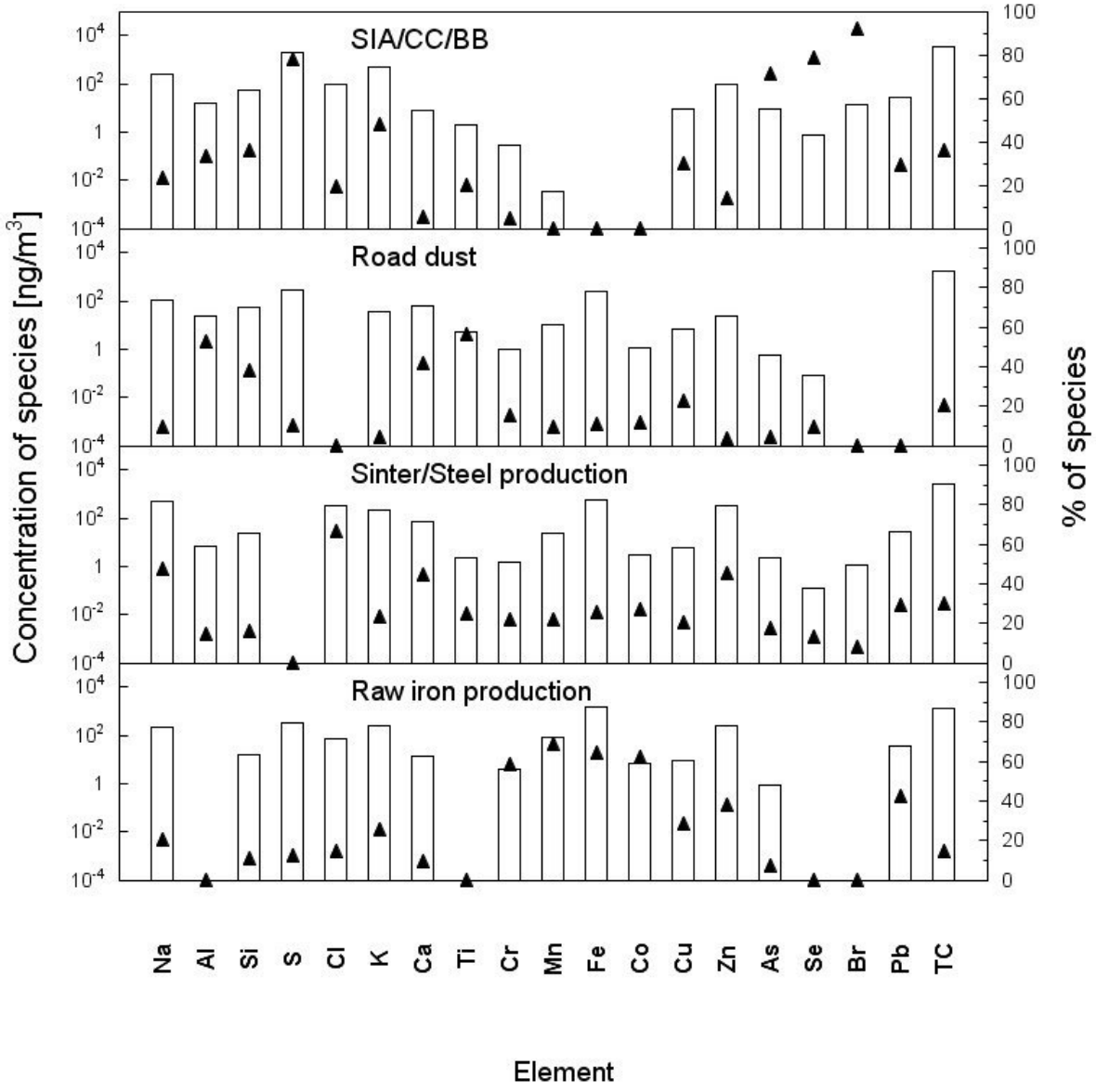
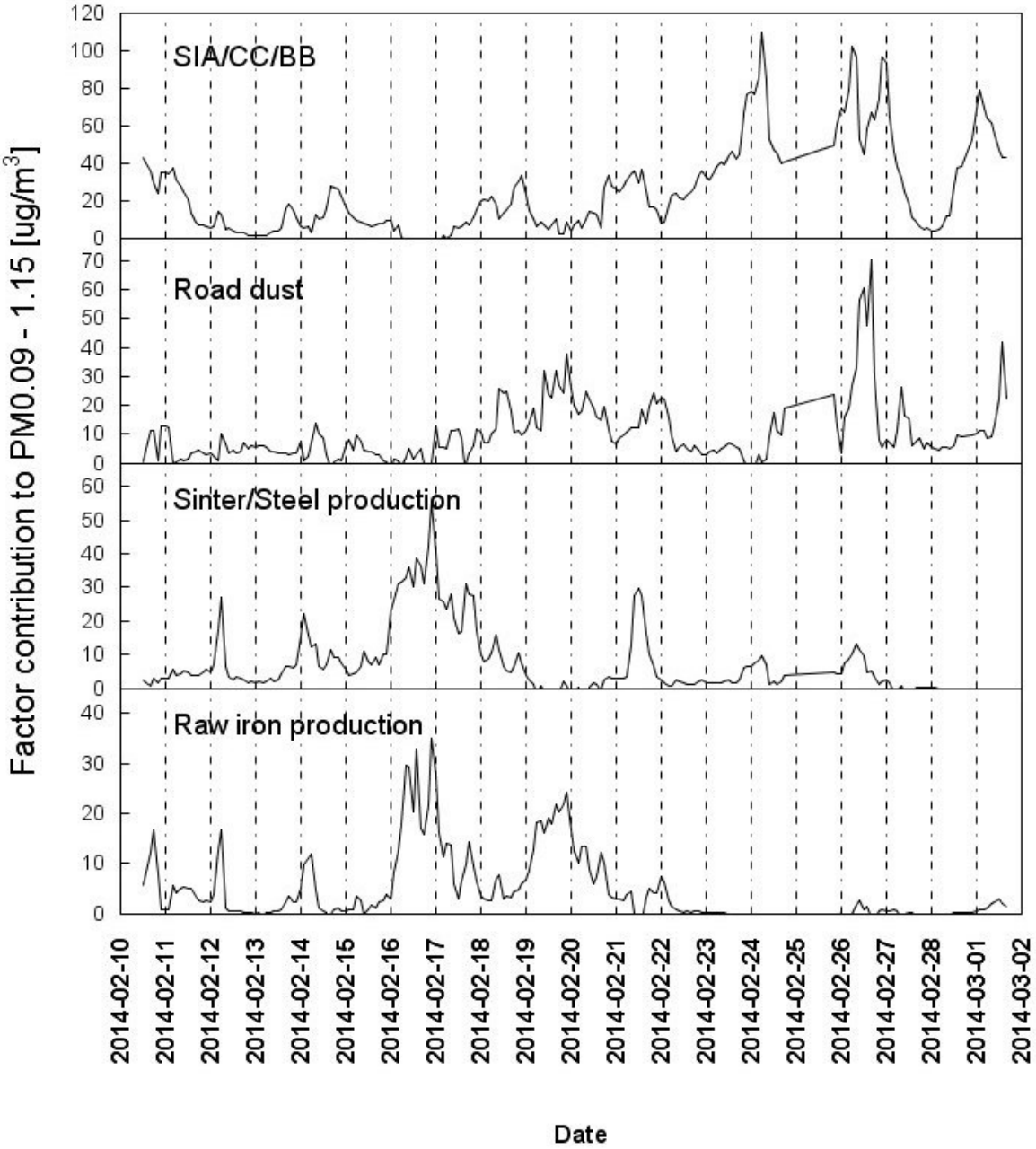


Figure S8. Daily pattern of the factor SIA/CC/BB obtained with chemical composition PMF.



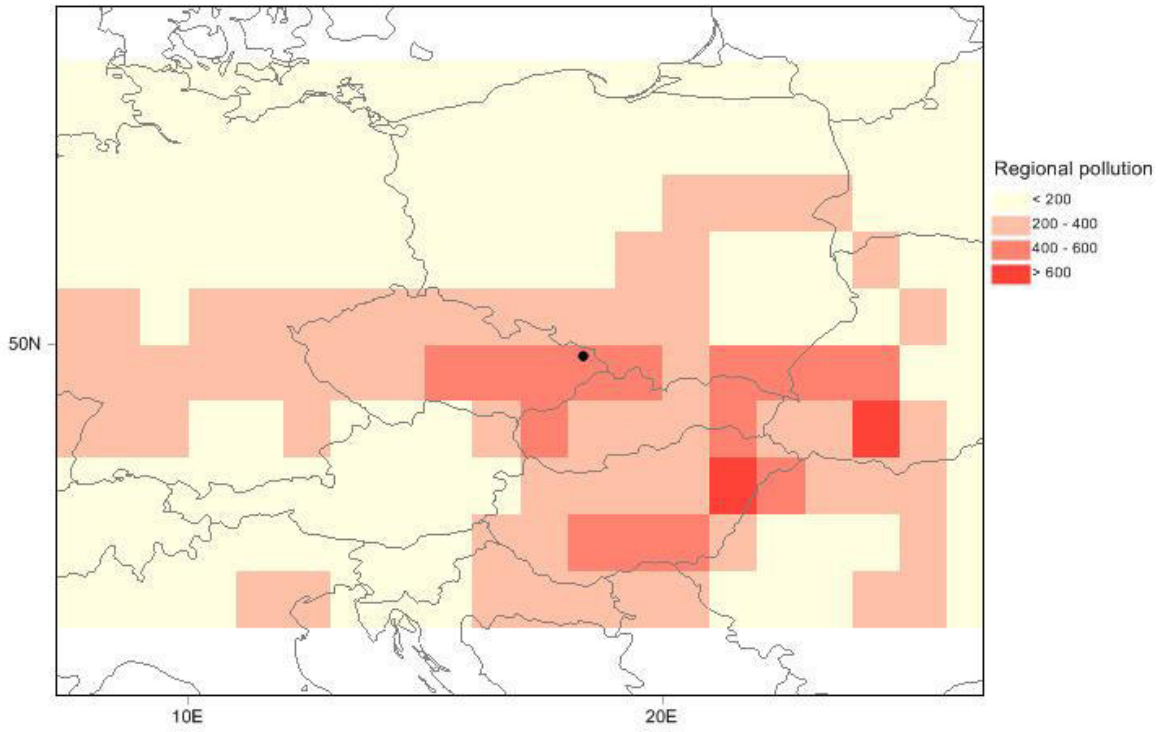
207
208 Figure S9. Factor profiles for the resolved factors of $PM_{0.09\text{-}1.15}$ by PMF.



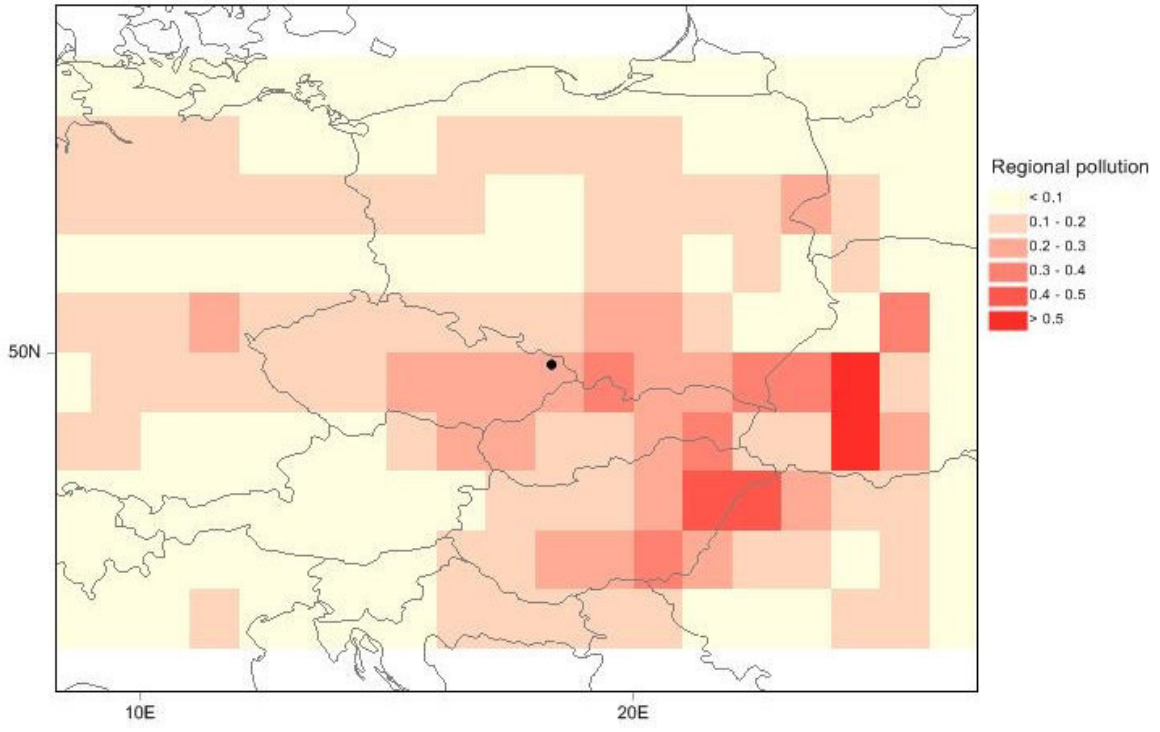
209

210 Figure S10. Temporal variations in the estimated contributions from the four factors resolved by

211 PMF.

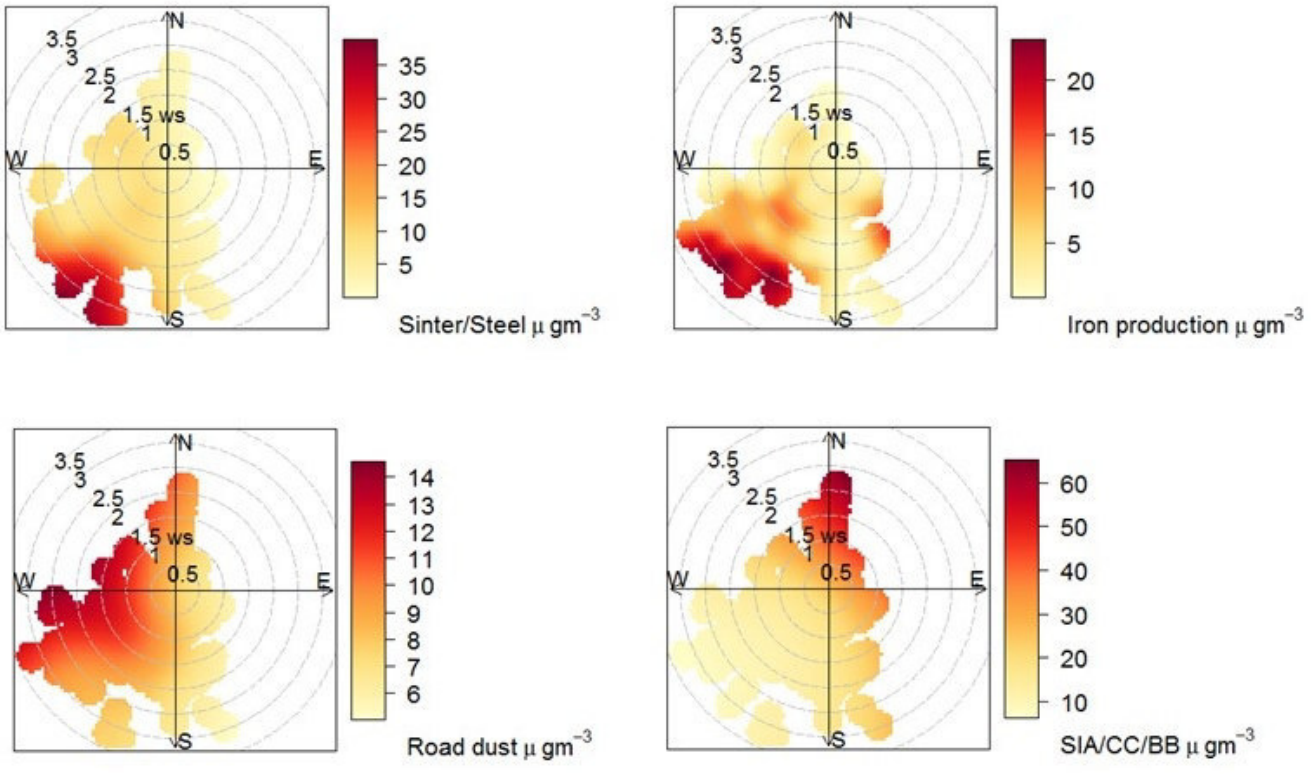


212

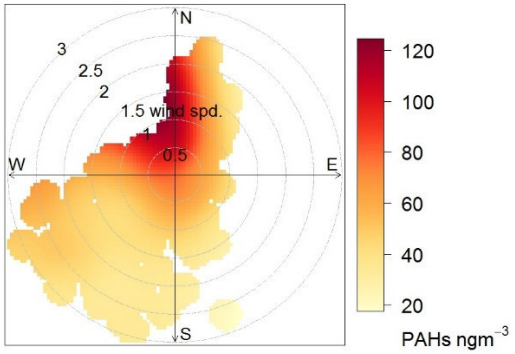


213

214 Figure S11. CWT of Factor 5 regional pollution (above) and PSCF (below). The data in the
215 CWT are expressed in Ncm^{-3}.



216



217

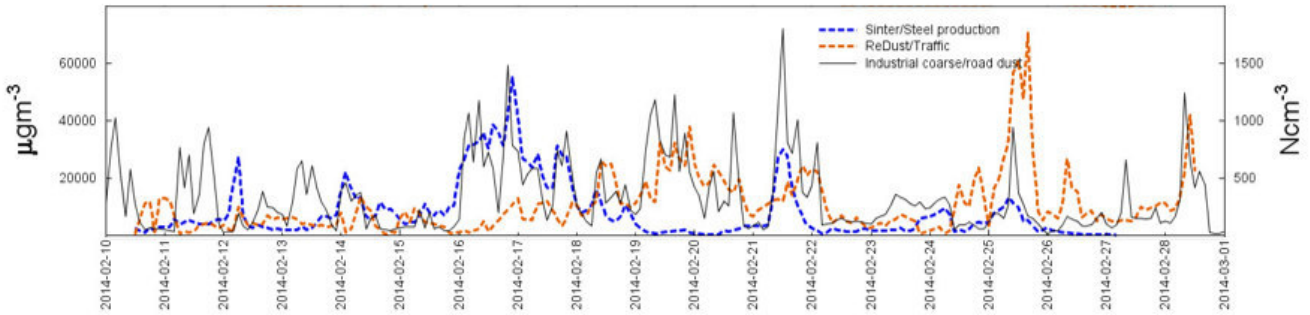
218 Fig. S12. Polar plots for 2-hours chemical composition factors and 24-hours PAHs in PM_1. The
219 24 hours PAHs averages were extended to the 15-minutes vector averages of WS and WD.

220

221

222

223



224

225 Fig. S13. Time series of Industrial coarse particles/road dust factor resolved with NSD (solid
226 line) and factors road dust/ Sinter/steel production resolved with chemical composition (dashed
227 lines).