



Source Apportionment of PM_{2.5} Using Positive Matrix Factorization (PMF) and PMF with Factor Selection

Peter Molnár^{1*}, Sandra Johannesson¹, Ulrich Quass²

¹ Occupational and Environmental Medicine at University of Gothenburg, Sweden

² IUTA e. V., Air Quality & Sustainable Nanotechnology Unit, Duisburg, Germany

ABSTRACT

Personal exposure, indoor, residential outdoor and urban background particulate matter (PM_{2.5}) samples were collected in parallel, for 30 participants and analyzed for their chemical content. Source apportionments for the separate microenvironments were performed using conventional positive matrix factorization (PMF), and for the combined dataset, applying a new PMF method with factor selection.

Regional sources were the largest contributor to the sampled PM_{2.5} in all microenvironments and accounted for 69% in urban background; 55% and 54% in residential outdoor and indoor environment, respectively; and 40% of personal exposure. For personal exposure, personal activities accounted for 21% (2.2 µg/m³), and constituted the main difference in total mass concentration between personal exposure and the other microenvironments.

The PMF method with factor selection was found to be a useful tool in the PMF analysis of multiple microenvironments, since ambient contributions to indoor and personal exposure are less likely to be distorted or misinterpreted. The possibility to more correctly estimate the source contributions will increase by combining the datasets for the different microenvironments into a larger dataset and using the PMF with factor selection method.

Keywords: Personal exposure; Indoor; Residential outdoor; Urban background.

INTRODUCTION

People are exposed to air pollution which originates naturally or due to anthropogenic activities. Air Pollutants can travel thousands of miles from its source (Huneeus *et al.*, 2011), their concentrations vary spatially and temporally (Ito *et al.*, 2004). Moreover, daily activities (traffic, road dust, domestic wood burning, and industrial activities) play an important role in the concentration and distribution of PM (Tang *et al.*, 2004).

Numerous epidemiological studies have found associations between mass concentrations of fine particulate matter (PM_{2.5}) and adverse health effects (Samet *et al.*, 2000; Pope *et al.*, 2002; Brook *et al.*, 2010; WHO, 2013). There are many studies investigating relationships between personal exposure and indoor as well as ambient levels of particles mass concentration. These studies were mostly focused on PM_{2.5} and sometimes also on black smoke, and performed in Europe and North America. Some studies have shown that ambient concentrations do not always reflect personal

exposure (Oglesby *et al.*, 2000; Sarnat *et al.*, 2005; Sorensen *et al.*, 2005; Johannesson *et al.*, 2007). Some studies have also characterized the chemical composition of the PM_{2.5} in personal exposures (Yakovleva *et al.*, 1999; Oglesby *et al.*, 2000; Kinney *et al.*, 2002; Lai *et al.*, 2004; Larson *et al.*, 2004; Janssen *et al.*, 2005; Molnár *et al.*, 2005; Molnár *et al.*, 2006; Zhao *et al.*, 2006; Johannesson *et al.*, 2011). The infiltration of ambient air indoors can be calculated using elements with no indoor sources (e.g., S or Pb), and thus, the influence of ambient air on indoor air can be assessed (Hänninen *et al.*, 2011). Additional information obtained from chemical analysis makes it possible to estimate the contributions from different sources and their influence on personal exposure, using statistical methods like principal component analysis (PCA) (Koistinen *et al.*, 2004) or positive matrix factorization (PMF) (Yakovleva *et al.*, 1999; Larson *et al.*, 2004; Zhao *et al.*, 2006). There are, however, few studies that have data obtained from simultaneous collection of personal, indoor, and outdoor samples. Parallel personal, residential indoor and outdoor sampling is more time consuming and labor intensive compared to a fixed monitoring station study. Therefore, personal exposure datasets are often rather small and less suitable for the application of standard PMF methods. However, it is important to apply PMF on personal exposure data and it is therefore desirable to develop methods for smaller datasets.

* Corresponding author.

Tel.: +46-31-786 2857; Fax: +46-31-82 50 04
E-mail address: peter.molnar@amm.gu.se

In this paper we test a new factor selection script for the multilinear engine (ME-2) program on a small dataset with 30 parallel samples in four different microenvironments (personal exposure, residential indoor, residential outdoor and urban background). The results are discussed and compared to the four separate microenvironment PMF analyses.

The aims were to (1) compare the two different approaches (separate and combined PMF) of source characterization; (2) characterize and quantify sources to personal exposure, indoor, residential outdoor and urban background concentrations of PM_{2.5} and investigate differences between the different microenvironments.

MATERIALS AND METHODS

Study Design, Participants, and Sampling Equipment

The present study took place in Gothenburg, Sweden between April 2nd–June 7th and September 26th–November 6th 2002, and in 2003 (March 27th–June 12th and October 7th–30th) to cover spring and autumn seasons, respectively. Personal exposure, indoor, and residential outdoor measurements of PM_{2.5} for 30 participants were performed in parallel with PM_{2.5} measurement at a stationary outdoor urban background station (UB) (Johannesson *et al.*, 2007). The participants lived within 0.8–15 km from the urban background station (median distance 3.3 km). The sampling time was 24 hours.

Identical sets of equipment were used for personal exposure and indoor and residential outdoor sampling, namely, a GK2.05 (KTL) cyclone for PM_{2.5} connected to a BGI 400S pump (BGI Inc., Waltham, MA, USA) with a flow rate of 4 L/min. For PM_{2.5} measurements at the urban background station, an EPA-WINS impactor (PQ100 EPA-WINS Basel PM_{2.5} (BGI Inc., Waltham, MA, USA), with a flow rate of 16.7 L/min was used. We used 37 mm Teflon filters (Pall Teflo, R2PJ037) in the cyclones and 47 mm Teflon filters (Pall Teflo, R2PJ047) in the WINS impactor. Information about the weather (temperature, wind speed, wind direction, rain, etc.) was provided by the Environmental Agency in Gothenburg. More detailed information regarding the study design, the participants, and the sampling can be found in Molnár *et al.* (2006) and Johannesson *et al.* (2007). The study was approved by the Ethics Committee at the University of Gothenburg.

One participant had a significant workplace exposure, and this personal sample was excluded in the analysis. Another participant was found to have a very high Ni concentration in the personal sample (46 ng/m³ vs. the median personal exposure of 2.6 ng/m³). This outlier is believed to be a contamination, and was replaced by the personal exposure median value and the uncertainty was increased accordingly. The total number of samples were; 29 personal, 30 indoor, 29 residential outdoor, and 28 UB.

Analytical Techniques

All filters were weighed before and after exposure using a CAHN C-30 microbalance placed in a temperature- and humidity-controlled room (Johannesson *et al.*, 2007). An energy-dispersive X-ray fluorescence (EDXRF) spectrometer

at the Department of Chemistry, Atmospheric Science, University of Gothenburg, was used to analyze the elemental composition of all filter samples (Molnár *et al.*, 2006). The EDXRF spectra were processed and quantified using the Quantitative X-ray Analysis System (QXAS) and the Analysis of X-ray spectra by Iterative Least-square fitting (AXIL) (Bernasconi *et al.*, 2000). The uncertainty is calculated as three times the standard deviation of the fitted peak area, for each element individually. Information about calibration and quality control has been presented elsewhere (Molnár *et al.*, 2005). The mean analytical precision was 5%, as calculated from repeated analysis (N = 5) of two randomly selected filters, one having a low and the other a high mass loading.

After the elemental analysis the filters were examined for black smoke using an EEL 43 smoke stain reflectometer (Johannesson *et al.*, 2007). The reflectance of the filter was transformed into an absorption coefficient, *a*, according to the international standard (ISO9835, 1993) and afterwards converted to black carbon concentration according to the equation presented by (Quincey, 2007).

Source Apportionment by Positive Matrix Factorization

Source apportionment analysis was carried out by positive matrix factorization (PMF) applying the multilinear engine (ME) technique using the ME-2 program (Paatero, 1999). PMF is a multivariate receptor model concept that estimates the source profiles and their contributions based on a weighted least square approach (Paatero and Tapper, 1994; Paatero, 1997). The task of the PMF model, Eq. (1), is to obtain the unknown matrices, *G* and *F*, by the iterative treatment of a least square method:

$$X = GF + E \quad (1)$$

$$Q(E) = \sum_{i=1}^m \sum_{j=1}^n \left(\frac{e_{ij}}{s_{ij}} \right)^2 \quad (2)$$

where *X* is the data matrix (size *m* × *n*) consisting of *n* chemical components analyzed in *m* samples, *G* is the source contribution to each sample (size *m* × *p*) for *p* factors, and *F* is the matrix of source profiles (size *p* × *n*). The matrix *E* is the residual. The main task of the iteration is to minimize the *Q*-value, which is defined in Eq. (2) as the sum of squares of residuals (*e*_{*ij*}) weighted inversely with the error estimates (*s*_{*ij*}) of data points.

In the present study, black carbon (BC) and the following elements were included in the PMF analysis: S, Cl, K, Ca, Ti, V, Mn, Fe, Ni, Cu, Zn, Br, Pb, as well as PM_{rest}. In all the analyses, the elemental mass concentrations have been recalculated to their mean oxidized mass concentrations, when applicable. PM_{rest} is the mass of the unknown part of the sampled PM_{2.5} and includes the major ionic species nitrate and ammonium, as well as silicon and organic species. PM_{rest} was calculated by subtracting the identified concentrations (as listed above, running from BC to Pb) from the total mass concentration.

The different types of microenvironments, ambient (i.e.,

urban background (UB) and residential outdoor), indoor, and the personal exposure were analyzed separately by regular ME-2 to investigate the source contributions without taking into account the other microenvironments.

PMF with Factor Selection

In the analysis combining the four microenvironments UB, residential outdoor, indoor, and personal exposure into one dataset, a new script developed for ME-2 by Pentti Paatero (personal communication) called “PMF with factor selection” was used. In PMF with factor selection, henceforth termed FS, the dataset is complemented with a new selection variable with the values 1 for the ambient samples (UB and residential outdoor), 2 for indoor, and 3 for personal exposure, besides the 15 base variables. In the factor selection script, a factor selection matrix, with three rows (representing the three different types of microenvironments, ambient, indoor, and personal) is used. The columns are the total number of factors in the model. In the matrix, a “1” denotes that the factor should be fitted for the type of sample and a “0” that it should be ignored. An example matrix of a model with four factors for the ambient samples (first row, F1–F4), six factors for the indoor samples (second row, F5 and F6 are indoor specific factors), and seven factors for the personal samples (third row, F7 is the personal activities factor) is presented in Table 1. This means that the whole dataset is used when determining the first four factors, both indoor and personal samples when determining the two indoor factors, and the personal samples only for the personal factor.

Error Estimates and Error Models Used in PMF

The combined dataset contained a number of values below the limit of detection (LoD). For each of these values, the LoD was specified instead of the actual concentration value. For these values, the following modeling procedure was used. The LoD value was used as data X_{ij} , to be fitted. The uncertainty was set to 0.5 LoD. Additionally, the error model used for values $< \text{LoD}$ was set so that the PMF model accepts all fitted values between 0 and $X_{ij} = \text{LoD}$ as a perfect fit, giving rise to zero Q contribution. This corresponds to the spirit of “detection level,” which essentially means that the true value should be somewhere between 0 and LoD, but no preference exists for any value between 0 and LoD.

RESULTS AND DISCUSSION

Mean $\text{PM}_{2.5}$ concentrations for personal exposure, indoor samples, residential outdoor samples, and samples at the urban background station were 11.0, 9.7, 7.8, and $10.1 \mu\text{g}/\text{m}^3$,

respectively (Johannesson *et al.*, 2007). The data on the elemental concentrations can be found in Molnár *et al.* (2006). The chemically undetermined part of $\text{PM}_{2.5}$, PM_{rest} , accounted for, on average, 59% of the total $\text{PM}_{2.5}$.

Conventional PMF of the Different Microenvironments

Urban Background and Residential Outdoors

A four-factor model (*long range transport + ship emissions (LRT + Ships), local combustion, traffic, and sea salt + resuspension*) identified the major sources for $\text{PM}_{2.5}$ at the UB and residential outdoor sampling sites (see Fig. 1 and Table A1 in the supplement material for the factor profiles).

LRT + Ships dominated the contribution both at the UB and residential outdoor locations, (69% and 55%, respectively). The high contribution from *LRT* pollution is in accordance with the results presented by Forsberg *et al.* (2005), who estimated that on average, only 29% of ambient PM_{10} in Sweden was of local origin, and the majority was from regional sources. The *local combustion* factor accounted for 20–23% for UB and residential outdoor $\text{PM}_{2.5}$. Fossil fuel and biomass burning for residential heating were sources in this factor, as were emissions from industries and the refinery in the harbor area. The other two sources made only minor contributions (less than $1 \mu\text{g}/\text{m}^3$ on the average). The small contribution from *traffic* was likely in part due to the fact that the UB station is at an elevated position (at “the medical hill” on the campus area of the faculty of medicine) and not close to any major traffic routes. For the residential outdoor, the majority of the homes participating in the study were also not close to the major roads. *Sea salt + resuspension* together formed the third ambient factor. Since westerly winds (from the ocean) are common and wind speeds are usually higher from this direction, both transport of sea spray and windblown dust are part of this factor.

Similar average levels between UB and residential outdoor were found for the factors *LRT + Ships* and *traffic*, as expected, since both affect the whole city (and major traffic routes tend to be located away from residential areas). The levels of similarities for the other factors; *sea salt + resuspension* and *local combustion* were very low. The reason is probably that these sources are more local and therefore affect the different places (i.e., the locations of the homes) differently due to dispersion.

Indoor

For the indoor measurements, six different sources/factors could be identified (Fig. 2(a)). *Indoor resuspension* had the highest contribution to the indoor levels followed by *traffic*. The remaining four factors (*marine, indoor Cu, soil*

Table 1. The Factor Selection matrix with seven factors in total. Four ambient factors, two additional indoor factors, and another one for personal exposure.

	F1	F2	F3	F4	F5	F6	F7
Ambient	1	1	1	1	0	0	0
Indoor	1	1	1	1	1	1	0
Personal exposure	1	1	1	1	1	1	1

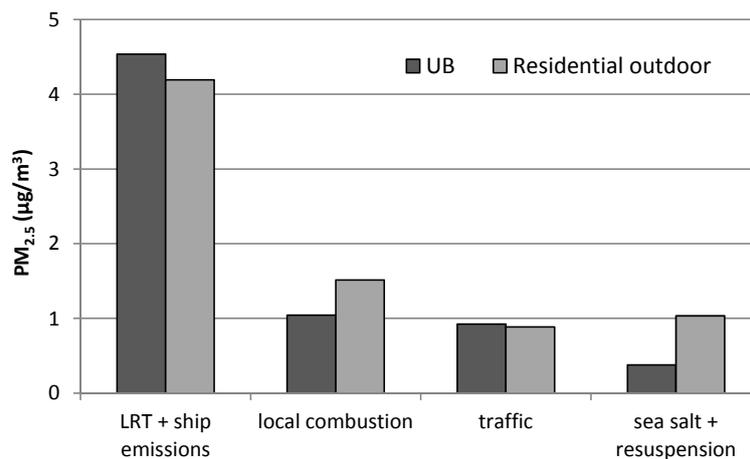


Fig. 1. The resulting factor contributions to PM_{2.5} for the combined urban background and residential outdoor datasets.

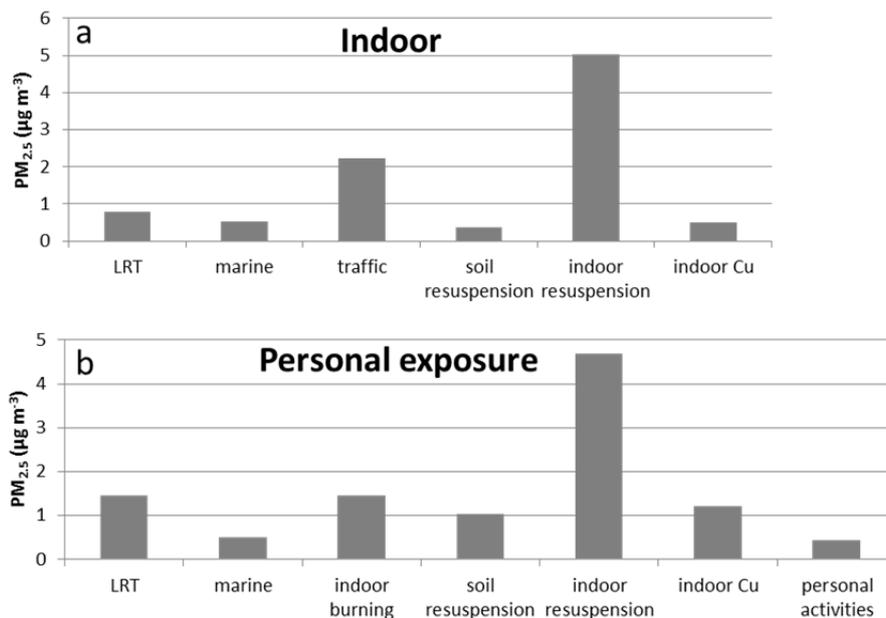


Fig. 2. The resulting factor contributions to PM_{2.5} for indoor (a) and personal exposure (b) sampling.

resuspension and *LRT*) contribute to a lower, but similar extent. The *marine* factor was highly correlated to Cl and the origin of the *Indoor Cu* was most likely electrical appliances and cooking. The *LRT* and *soil resuspension* factors made only small contributions to the total mass. The *indoor resuspension* factor had chemical characteristics similar to those of these latter two factors, indicating outdoor origin, and the settled particles were then re-emitted by indoor activities. The indoor sources, *indoor resuspension* and *indoor Cu sources* were activity driven and are highly variable in strength from day to day, depending on the amount and type of activity in the homes.

Personal Exposure

The personal exposure were best represented by seven factors (Fig. 2(b)), with strong influence from indoor activities (mainly *indoor resuspension*) as was the case for indoor measurements. *Indoor burning*, *indoor Cu*, *LRT*, *soil*, and

the *marine* factor, together with the *personal activities* factor, were minor contributors. Factors that contributed the most were associated with various activities such as movements indoors that cause resuspension of settled dust, and burning candles, cooking/frying, and use of household appliances (the Cu factor).

The resulting source factors from the PMF modeling of the different microenvironments, ambient air (UB and residential outdoor), indoor, and personal exposure, can be found in Tables A1–A3, in the supplement material.

Time Series Comparison

Strong correlations were found between UB and residential outdoor for the factors *LRT + ship* and the *traffic* ($r_p = 0.80$ and 0.75 , respectively). For the two other factors, *sea salt + resuspension*, and *local combustion*, weaker correlations were found between UB and residential outdoor (0.43 and 0.23 , respectively).

The correlations between the factors in the indoor model vs. the factors in the UB and residential outdoor models were weak for all combinations of factors, $r_p < 0.5$. The elemental source profile for the *indoor resuspension* factor indicated that it consisted of a combination of outdoor sources (LRT, oil combustion, traffic, and soil), but was not correlated to these factors in the ambient models. This leads to the conclusion that although the *indoor resuspension* factor was consisted of a considerable quantity of particles that were of outdoor origin, the day-to-day variation indoors was mainly governed by indoor activities and not by the fluctuation of outdoor concentrations.

For personal exposure, there were strong correlations to the corresponding indoor factors *indoor/personal Cu*, *marine*, and *indoor resuspension* ($r_p = 0.93, 0.81, \text{ and } 0.78$, respectively), and moderate correlation to the *LRT* factor ($r_p = 0.57$). The day-to-day variations were therefore most likely directly connected to the corresponding indoor activities.

When comparing the time series contribution for the personal exposure with the UB and residential outdoor, two groups of factors stand out. Firstly, the personal *LRT* factor was moderately to strongly correlated with the UB and the residential outdoor *LRT + Ship* and *traffic* factors ($r_p = 0.66$ and 0.83 compared to the UB time series and $r_p = 0.44$ and 0.55 compared to the residential outdoor time series). Secondly, there was a moderate correlation between the *marine* factors, $r_p = 0.35\text{--}0.44$. For all other factors, no or very weak correlations were found. The time series factor contribution for the different microenvironments can be found in Fig. A1, in the supplement material.

Combined Analysis, PMF with Factor Selection

Analysis of the full dataset in one combined model using the factor selection script for the ME-2 resulted in a model with four outdoor factors (*LRT(FS)*, *local combustion(FS)*, *traffic(FS)*, and *sea salt + resuspension(FS)*), two additional indoor factors (*indoor resuspension(FS)* and *indoor heated sources(FS)*), and a *personal activities(FS)* factor. Combined model factors are indicated in the text with a (FS) in the end of the name to distinguish from the factors obtained by the separate PMF models. The source profiles are presented in Fig. 3.

The proportion of each factor in the different microenvironments is presented in Fig. 4. The contribution from *LRT(FS)* was the main source and dominated the ambient environments, UB (69%) and residential outdoor (55%), and was also the largest single source indoors (54%) and for personal exposure (40%). The *local combustion(FS)* factor was stronger for residential outdoor than UB, but contributed only to a lesser extent indoors and for personal exposure. The contribution from the *traffic(FS)* factor was small in all environments. The contribution from the *sea salt + resuspension(FS)* factor to UB was small, probably due to the elevated position of the sampling site, and like the *traffic(FS)* factor, contributed more to residential outdoor and personal exposure than to UB. Regarding the two indoor factors, *indoor heated sources(FS)* was stronger than *indoor resuspension(FS)* and the mass contributions were

nearly equal for these factors for both indoor and personal exposure. The factor contributions for the indoor and personal exposure were similar (excluding the *personal activities(FS)* factor). In the personal exposure, the *personal activities(FS)* factor accounted for $2.2 \mu\text{g}/\text{m}^3$ (21% of the total personal exposure). The estimated mean $\text{PM}_{2.5}$ mass for personal exposure was $10.6 \mu\text{g}/\text{m}^3$, and $7.6, 8.0, \text{ and } 8.6 \mu\text{g}/\text{m}^3$ for UB, residential outdoor, and indoor, respectively. All factors had fairly similar contribution for personal exposure compared to indoor and residential outdoor. Consequently, the *personal activities(FS)* add an extra mass contribution, resulting in a higher personal exposure compared to the other microenvironments. Similar results have been shown in previous studies (Wallace, 2000; Williams et al., 2003; Ito et al., 2004; Wallace et al., 2006).

The correlations between the microenvironments in the FS model for the different factors were generally high (Table 2). For the traffic factor however, significant correlations were only found between residential outdoor and indoor and for residential outdoor and personal exposure, and no correlations were found between UB and any of other microenvironments. This is likely due to the fact that traffic is a local source and varies over the city. The two indoor sources, *indoor resuspension(FS)* and *indoor heated sources(FS)*, were strongly correlated to the corresponding personal exposure ($r_p = 0.94$ and 0.65 , respectively).

The Factor Selection Model versus Separate PMF Models

The four outdoor factors resolved in the FS model were strongly correlated with the corresponding factors in the separate four-factor models for both UB and residential outdoor ($r = 0.88\text{--}0.98$). This suggests that the chosen FS model does not introduce additional ambiguity, at least not regarding the ambient sources.

The separate indoor and personal exposure models, however, deviated more from the FS model for several factors. Since the separate indoor model does not take into account the fact that parts of the sampled concentrations origin from outdoor sources, some sources that had partly the same composition, for example, the factors connected to combustion, LRT, and traffic, showed different correlation patterns. The *local combustion(FS)* factor was correlated to the indoor model factors *LRT* and *traffic* ($r_p = 0.69$ and 0.62 , respectively) while the *LRT(FS)* factor was not correlated to any factor in the indoor model. The factors *indoor heated sources(FS)* and *indoor resuspension(FS)* were correlated to the *indoor Cu* and *indoor resuspension* factors in the separate models ($r_p = 0.90$ and 0.86 , respectively). That is, the separate indoor and personal exposure models may not be able to correctly identify ambient sources. For example, the day-to-day variation of the ambient sources and the resuspension of these at later times, due to indoor activities, will not be separated into different factor types in the separate indoor and personal exposure models.

Comparison with Other Studies

There are only a limited number of studies that have measured personal, indoor and residential outdoor PM simultaneously and then performed source apportionment

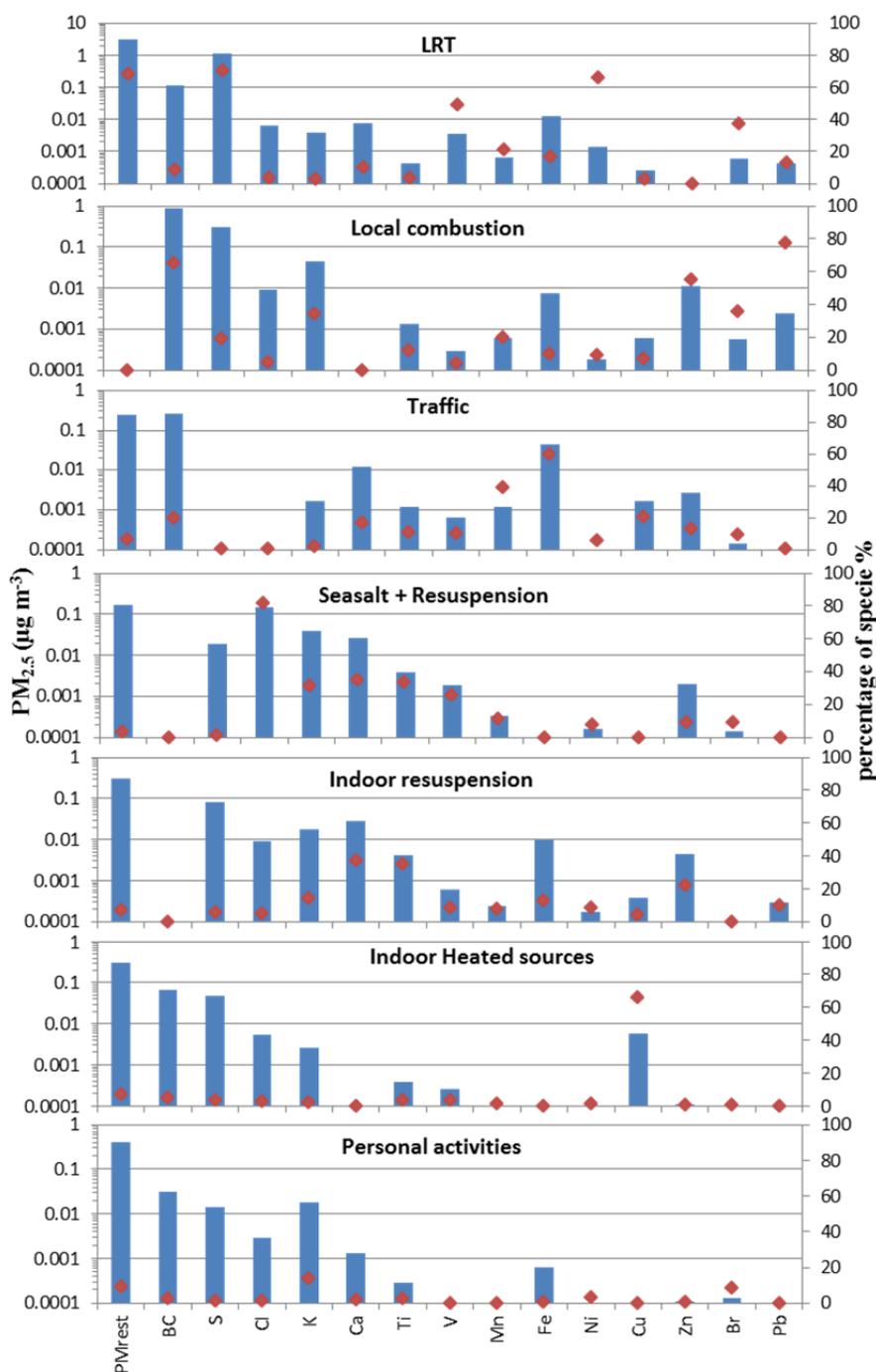


Fig. 3. The seven-factor model using factor selection in ME-2 for the whole dataset, with the settings four factors outdoors, two additional factors indoors, and one additional factor for the personal exposure. Concentrations ($\mu\text{g}/\text{m}^3$) are shown on the left Y-axis and percentage of each species apportioned in the factors (diamonds) on the right Y-axis. Note that the scales on the left axis are different for the factors *LRT* and *sea salt + resuspension*.

using the PMF technique on the datasets (Yakovleva *et al.*, 1999; Larson *et al.*, 2004; Zhao *et al.*, 2006).

In the Riverside, CA, study involving 178 participants, Yakovleva *et al.* (1999) measured PM_{10} outdoors (O), indoors (I) and personal exposure (P), as well as $\text{PM}_{2.5}$ outdoors ($\text{O}_{2.5}$) and indoors ($\text{I}_{2.5}$). The following seven factors (and in which environments they contribute) were found; ambient soil (O), resuspended indoor soil (P), indoor soil (I, P),

personal activities (P, I), sea salt (O), nonferrous operations and motor vehicle exhaust (O, $\text{O}_{2.5}$, I, $\text{I}_{2.5}$, and P), and secondary SO_4 (O, $\text{O}_{2.5}$, I, $\text{I}_{2.5}$, and P).

In a study involving 20 participants (83 personal samples), in Seattle, WA, Larson *et al.*, (2004) found the following outdoor $\text{PM}_{2.5}$ sources: vegetative burning, secondary sulfate, mobile emissions, fuel oil, crustal, and Cl rich. The indoor $\text{PM}_{2.5}$ sources were vegetative burning, secondary sulfate,

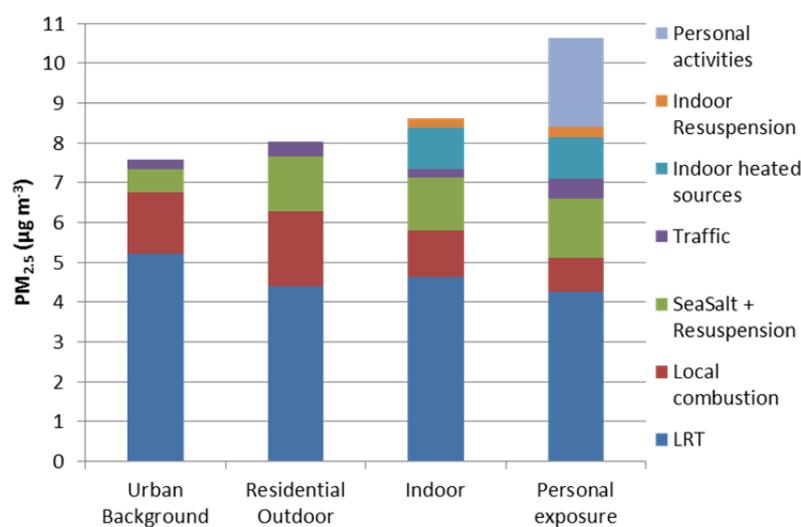


Fig. 4. The contribution from the different factors in the FS model to each microenvironment.

Table 2. Correlations between the different microenvironments and corresponding factors in the factor selection model. Correlations marked in bold are statistical significant ($p < 0.05$).

	LRT	Local combustion	Traffic	Sea salt + resuspension	Indoor resuspension	Indoor heated sources
UB vs. Out	0.81	0.85	0.20	0.64	—	—
UB vs. In	0.69	0.66	0.17	0.55	—	—
UB vs. Pers	0.71	0.85	0.07	0.64	—	—
Out vs. In	0.70	0.63	0.62	0.75	—	—
Out vs. Pers	0.69	0.66	0.46	0.46	—	—
In vs. Pers	0.73	0.73	0.26	0.61	0.94	0.65

crustal, mobile source, and Cl rich. For the personal exposure, the following PM_{2.5} sources were found: vegetative burning, crustal 2, crustal, Cl rich, and S, Mn, Fe (a combination of secondary sulfate, the mobile source, and cooking).

In the study by Zhao *et al.* (2006) on PM_{2.5} conducted in Raleigh and Chapel Hill, NC, involving 38 participants, four ambient/residential outdoor factors (secondary SO₄, motor vehicle, soil, and secondary nitrate) were identified. For indoor/personal, four additional factors (cooking, personal care and activity, environmental tobacco smoke (ETS) and its mixture, and a Cu factor mixed with indoor soil) were identified.

In the present study and the studies mentioned above, there were several factors of similar type/origin in the different environments. Typical ambient/outdoor sources were regionally emitted sources, traffic-related sources, crustal material, and marine influence (when relevant). Several of these sources also contributed in various degrees to the indoor and personal exposure. Common indoor and personal sources were indoor resuspension of soil, indoor activities such as cooking, and other personal activities. The absolute (or relative) contribution of these, and other sources, may differ between the studies due to local and regional conditions, for example, size of the city, vehicle fleet composition, building types and ventilation, climate, season, and industries nearby the sampling stations. It also depends on which measured substances were used in the

models. It is therefore hard to make quantitative comparisons between studies from different locations, but qualitative comparisons can be helpful especially during the factor identification process.

Potentials and Limitations of the Factor Selection Model

An advantage of the FS script is that by combining several microenvironments, all observations are put into one large model. This can be important in studies involving time-consuming sampling strategies (e.g., parallel outdoor, indoor, and personal exposure sampling), where the number of sampling days is a limiting factor. A similar approach has recently been reported in combined two-site PMF analysis (Beuck *et al.*, 2011; Molnar and Sallsten, 2013).

For many of the chemical species, strong correlations were found between UB and residential outdoor and between indoor and personal exposure (Molnár *et al.*, 2006; Johannesson *et al.*, 2011). However, the correlations between ambient (UB and residential outdoor) and indoor, and between ambient and personal exposure were much weaker. Nevertheless, the FS script managed to determine source factors that were realistic, and the correlations between the different factors were generally strong when comparing the different microenvironments.

However, there is a principal limitation with regard to the application of PMF on indoor and personal data. Indoor activities vary between different homes. Thus the actual

composition profile of indoors-generated aerosol is strongly variable from residence to residence, and even from day to day within one residence. This variation violates the basic assumption of PMF modeling that source profiles are stable over time, and such variation cannot be correctly modeled by two or three factors. For this reason, the modeling of indoor and personal data must be regarded with caution. However, it should be noted that earlier PMF-based analyses of indoor and personal data have not been free of this problem either.

In an analysis of the personal exposure variability in this dataset, the within-person variance component dominated the total variability for nearly all investigated chemical species (Johannesson *et al.*, 2011).

CONCLUSIONS

The new PMF with factor selection model was found to be a useful tool in the PMF analysis of multiple microenvironments, since ambient contributions to indoor levels and personal exposure are less likely to be distorted or misinterpreted. By combining the datasets for the different microenvironments and using the PMF with factor selection method, more correct estimates of the source contributions to personal exposure can be accomplished. In particular, resuspension of settled ambient particles due to indoor activities could only be identified with the factor selection model and not with the separate microenvironment models. For personal exposure to PM_{2.5}, personal activities accounted for 21% (2.2 µg/m³), and constituted the difference in total mass concentration between personal exposure and the other microenvironments.

ACKNOWLEDGEMENT

This project was funded by the Swedish Council for Working Life and Social Research (FAS) under grant number 2008-0450. Pentti Paatero is greatly acknowledged both for providing the new ME-2 script “PMF with factor selection” and for valuable suggestions and advice in both the data analysis and the preparation of the manuscript. Gerd Sallsten for valuable comments on the manuscript.

SUPPLEMENTARY MATERIALS

Supplementary data associated with this article can be found in the online version at <http://www.aaqr.org>.

REFERENCES

- Bernasconi, G., Tajani, A. and Kregsamer, P. (2000). Manual for QXAS/AXIL. version 3.5. IAEA, Vienna.
- Beuck, H., Quass, U., Klemm, O. and Kuhlbusch, T.A.J. (2011). Assessment of Sea Salt and Mineral Dust Contributions to PM₁₀ in NW Germany Using Tracer Models and Positive Matrix Factorization. *Atmos. Environ.* 45: 5813–5821.
- Brook, R.D., Rajagopalan, S., Pope, C.A., Brook, J.R., Bhatnagar, A., Diez-Roux, A.V., *et al.* (2010). Particulate Matter Air Pollution and Cardiovascular Disease An Update to the Scientific Statement From the American Heart Association. *Circulation* 121: 2331–2378.
- Forsberg, B., Hansson, H.C., Johansson, C., Areskoug, H., Persson, K. and Jarvholm, B. (2005). Comparative Health Impact Assessment of Local and Regional Particulate Air Pollutants in Scandinavia. *Ambio* 34: 11–19.
- Huneus, N., Schulz, M., Balkanski, Y., Griesfeller, J., Prospero, J., Kinne, S., *et al.* (2011). Global Dust Model Intercomparison in AeroCom Phase I. *Atmos. Chem. Phys.* 11: 7781–7816.
- Hänninen, O., Hoek, G., Mallone, S., Chellini, E., Katsouyanni, K., Gariazzo, C., Cattani, G., Marconi, A., Molnár, P., Bellander, T. and Jantunen, M. 2011. (2011). Seasonal Patterns of Outdoor PM Infiltration into Indoor Environments: Review and Meta-analysis of Available Studies from Different Climatological Zones in Europe. *Air Qual. Atmos. Health* 4: 221–233.
- ISO9835 (1993). Ambient Air-Determination of a Black Smoke Index.
- Ito, K., Xue, N. and Thurston, G. (2004). Spatial Variation of PM_{2.5} Chemical Species and Source-apportioned Mass Concentrations in New York City. *Atmos. Environ.* 38: 5269–5282.
- Janssen, N.A.H., Lanki, T., Hoek, G., Vallius, M., de Hartog, J.J., van Grieken, R., Pekkanen, J. and Brunekreef, B. (2005). Associations between Ambient, Personal, and Indoor Exposure to Fine Particulate Matter Constituents in Dutch and Finnish Panels of Cardiovascular Patients. *Occup. Environ. Med.* 62: 868–877.
- Johannesson, S., Gustafson, P., Molnar, P., Barregard, L. and Sallsten, G. (2007). Exposure to Fine Particles (PM_{2.5} and PM₁) and Black Smoke in the General Population: Personal, Indoor, and Outdoor Levels. *J. Exposure Sci. Environ. Epidemiol.* 17: 613–624.
- Johannesson, S., Rappaport, S.M. and Sallsten, G. (2011). Variability of Environmental Exposure to Fine Particles, Black Smoke, and Trace Elements among a Swedish Population. *J. Exposure Sci. Environ. Epidemiol.* 21: 506–514.
- Kinney, P.L., Chillrud, S.N., Ramstrom, S., Ross, J. and Spengler, J.D. (2002). Exposure to Multiple Air Toxics in New York City. *Environ. Health Perspect.* 110: 539–546.
- Koistinen, K.J., Edwards, R.D., Mathys, P., Ruuskanen, J., Kunzli, N. and Jantunen, M.J. (2004). Sources of Fine Particulate Matter in Personal Exposures and Residential Indoor, Residential Outdoor and Workplace Microenvironments in the Helsinki Phase of the EXPOLIS Study. *Scand. J. Work Environ. Health* 30: 36–46.
- Lai, H.K., Kendall, M., Ferrier, H., Lindup, I., Alm, S., Hänninen, O., Jantunen, M., Mathys, P., Colvile, R. and Ashmore, M.R. (2004). Personal Exposures and Microenvironment Concentrations of PM_{2.5}, VOC, NO₂ and CO in Oxford, UK. *Atmos. Environ.* 38: 6399–6410.
- Larson, T., Gould, T., Simpson, C., Liu, L.J.S., Claiborn, C. and Lewtas, J. (2004). Source Apportionment of Indoor, Outdoor, and Personal PM_{2.5} in Seattle, Washington,

- Using Positive Matrix Factorization. *J. Air Waste Manage. Assoc.* 54: 1175–1187.
- Molnár, P., Gustafson, P., Johannesson, S., Boman, J., Barregard, L. and Sallsten, G. (2005). Domestic Wood Burning and PM_{2.5} Trace Elements: Personal Exposures, Indoor and Outdoor Levels. *Atmos. Environ.* 39: 2643–2653.
- Molnár, P., Johannesson, S., Boman, J., Barregard, L. and Sallsten, G. (2006). Personal Exposures and Indoor, Residential Outdoor, and Urban Background Levels of Fine Particle Trace Elements in the General Population. *J. Environ. Monit.* 8: 543–551.
- Molnár, P. and Sallsten, G. (2013). Contribution to PM_{2.5} from Domestic Wood Burning in a Small Community in Sweden. *Environ. Sci. Processes Impacts* 15: 833–838.
- Oglesby, L., Kunzli, N., Roosli, M., Braun-Fahrlander, C., Mathys, P., Stern, W., Jantunen, M. and Kousa, A. (2000). Validity of Ambient Levels of Fine Particles as Surrogate for Personal Exposure to Outdoor Air Pollution - Results of the European EXPOLIS-EAS Study (Swiss Center Basel). *J. Air Waste Manage. Assoc.* 50: 1251–1261.
- Paatero, P. and Tapper, U. (1994). Positive Matrix Factorization - A Nonnegative Factor Model with Optimal Utilization of Error-Estimates of Data Values. *Environmetrics* 5: 111–126.
- Paatero, P. (1997). Least Squares Formulation of Robust Non-negative Factor Analysis. *Chemom. Intell. Lab. Syst.* 37: 23–35.
- Paatero, P. (1999). The Multilinear Engine - A Table-driven, Least Squares Program for Solving Multilinear Problems, Including the n-way Parallel Factor Analysis Model. *J. Comput. Graph. Stat.* 8: 854–888.
- Pope, C.A., Burnett, R.T., Thun, M.J., Calle, E.E., Krewski, D., Ito, K. and Thurston, G.D. (2002). Lung Cancer, Cardiopulmonary Mortality, and Long-term Exposure to Fine Particulate Air Pollution. *J. Am. Med. Assoc.* 287: 1132–1141.
- Quincey, P. (2007). A Relationship between Black Smoke Index and Black Carbon Concentration. *Atmos. Environ.* 41: 7964–7968.
- Samet, J.M., Dominici, F., Currier, I., Coursac, I. and Zeger, S.L. (2000). Fine Particulate Air Pollution and Mortality in 20 US Cities, 1987–1994. *New Engl. J. Med.* 343: 1742–1749.
- Sarnat, J.A., Brown, K.W., Schwartz, J., Coull, B.A. and Koutrakis, P. (2005). Ambient Gas Concentrations and Personal Particulate Matter Exposures: Implications for Studying the Health Effects of Particles. *Epidemiology* 16: 385–395.
- Sorensen, M., Loft, S., Andersen, H.V., Nielsen, O.R., Skovgaard, L.T., Knudsen, L.E., Ivan, V.N.B. and Hertel, O. (2005). Personal Exposure to PM_{2.5}, black smoke and NO₂ in Copenhagen: Relationship to Bedroom and Outdoor Concentrations Covering Seasonal Variation. *J. Exposure Sci. Environ. Epidemiol.* 15: 413–422.
- Tang, W., Raymond, T., Wittig, B., Davidson, C., Pandis, S., Robinson, A. and Crist, K. (2004). Spatial Variations of PM_{2.5} during the Pittsburgh Air Quality Study. *Aerosol Sci. Technol.* 38: 80–90.
- Wallace, L. (2000). Correlations of Personal Exposure to Particles with Outdoor Air Measurements: A Review of Recent Studies. *Aerosol Sci. Technol.* 32: 15–25.
- Wallace, L., Williams, R., Rea, A. and Croghan, C. (2006). Continuous Weeklong Measurements of Personal Exposures and Indoor Concentrations of Fine Particles for 37 Health-impaired North Carolina Residents for up to Four Seasons. *Atmos. Environ.* 40: 399–414.
- WHO (2013). Review of Evidence on Health Aspects of Air Pollution – REVIHAAP Project, Copenhagen, Denmark.
- Williams, R., Suggs, J., Rea, A., Sheldon, L., Rodes, C. and Thornburg, J. (2003). The Research Triangle Park Particulate Matter Panel Study: Modeling Ambient Source Contribution to Personal and residential PM Mass Concentrations. *Atmos. Environ.* 37: 5365–5378.
- Yakovleva, E., Hopke, P.K. and Wallace, L. (1999). Receptor Modeling Assessment of Particle Total Exposure Assessment Methodology Data. *Environ. Sci. Technol.* 33: 3645–3652.
- Zhao, W., Hopke, P.K., Norris, G., Williams, R. and Paatero, P. (2006). Source Apportionment and Analysis on Ambient and Personal Exposure Samples with a Combined Receptor Model and an Adaptive Blank Estimation Strategy. *Atmos. Environ.* 40: 3788–3801.

Received for review, November 15, 2013

Accepted, February 13, 2014

Supplement material for the manuscript:Source apportionment of PM_{2.5} using Positive Matrix Factorization (PMF) and PMF with factor selectionTable A1. Contributions ($\mu\text{g}/\text{m}^3$) for the four factor model for ambient air (urban background and residential outdoor).

	Seasalt + Resuspension	Traffic	Local Combustion	LRT + Ship emission	Sum of species
PM _{2.5}	0.708	0.904	1.273	4.348	7.23
PMrest	0.361	0.194	0	2.712	3.27
BC	0	0.548	0.541	0.353	1.44
S	0	0	0.538	1.103	1.64
Cl	0.094	0	0.009	0.017	0.119
K	0.032	0.018	0.028	0.013	0.091
Ca	0.023	0.007	0	0.010	0.041
Ti	0.003	0.001	0.001	0.001	0.006
V	0.002	0	0.001	0.004	0.007
Mn	0.001	0.001	0.001	0.001	0.003
Fe	0.008	0.034	0.009	0.009	0.060
Ni	0	0	0	0.001	0.002
Cu	0	0.001	0	0	0.002
Zn	0.004	0.005	0.011	0	0.020
Br	0	0	0.001	0.001	0.002
Pb	0	0.001	0.002	0.001	0.004
Sum of Factor	0.528	0.810	1.142	4.226	6.71

Table A2. Contributions ($\mu\text{g}/\text{m}^3$) for the six factor model for the indoor samples.

	LRT	Marine	Traffic	Soil Resuspension	Indoor Resuspension	Indoor Cu	Sum of Species
PM _{2.5}	0.0325	0.6106	2.6650	0.5716	5.1352	0.5022	9.5171
PMrest	0	0.0044	1.0232	0	4.7777	0.3509	6.1562
BC	0.2150	0	1.0049	0	0	0.0523	1.2722
S	0.4929	0.3300	0.0982	0.2678	0.1436	0.0816	1.4141
Cl	0	0.1640	0.0468	0	0.0077	0.0050	0.2235
K	0.0598	0.0057	0	0	0.0854	0.0090	0.1599
Ca	0	0.0189	0.0175	0.0506	0.0016	0	0.0886
Ti	0.0009	0.0022	0.0017	0.0071	0.0010	0.0010	0.0139
V	0.0026	0.0021	0.0007	0.0009	0.0008	0.0008	0.0079
Mn	0.0005	0.0005	0.0010	0.0009	0	0.0001	0.003
Fe	0.0007	0.0044	0.0185	0.0281	0.0043	0.0023	0.0583
Ni	0.0008	0.0003	0	0.0005	0.0003	0.0001	0.002
Cu	0	0.0002	0.0029	0.0007	0	0.0102	0.014
Zn	0.0057	0.0013	0.0043	0.0065	0.0013	0.0006	0.0197
Br	0.0002	0.0002	0.0004	0.0002	0.0003	0	0.0013
Pb	0.0015	0	0.0007	0.0004	0	0	0.0026
Sum of Factor	0.78	0.53	2.22	0.36	5.02	0.51	9.44

Table A3. Contributions ($\mu\text{g}/\text{m}^3$) for the seven factor model for the personal exposure.

	LRT	Marine	Resuspension	Indoor Resuspension	Indoor Cu	Indoor burning	Personal	Sum of Species
PM _{2.5}	1.3166	0.4463	1.2122	4.8515	1.1249	1.6740	0.5012	11.1267
PMrest	0	0	0.8990	4.3045	0.9139	0.8486	0	6.966
BC	0.7208	0	0	0	0.1051	0.4564	0.2042	1.4865
S	0.6370	0.2866	0	0.3093	0.1392	0	0.1836	1.5557
Cl	0.0201	0.1733	0	0	0.0225	0.0383	0.0120	0.2662
K	0	0.0144	0.0535	0.0327	0	0.0970	0	0.1976
Ca	0.0285	0.0268	0.0480	0.0175	0.0098	0.0009	0	0.1315
Ti	0.0044	0.0029	0.0069	0.0015	0.0007	0.0016	0.0019	0.0199
V	0.0040	0.0016	0.0008	0.0006	0.0010	0.0003	0	0.0083
Mn	0.0007	0.0004	0.0004	0.0007	0.0005	0	0.0006	0.0033
Fe	0.0172	0	0.0160	0.0218	0.0119	0.0034	0.0361	0.1064
Ni	0	0.0001	0.0005	0.0004	0.0001	0.0001	0.0012	0.0024
Cu	0	0.0007	0.0014	0	0.0092	0.0014	0.0017	0.0144
Zn	0.0121	0	0.0088	0.0021	0.0007	0	0.0008	0.0245
Br	0	0	0	0.0013	0	0.0003	0.0002	0.0018
Pb	0.0015	0.0001	0.0006	0	0	0.0001	0.0001	0.0024
Sum of Factor	1.45	0.51	1.04	4.69	1.21	1.45	0.44	10.79

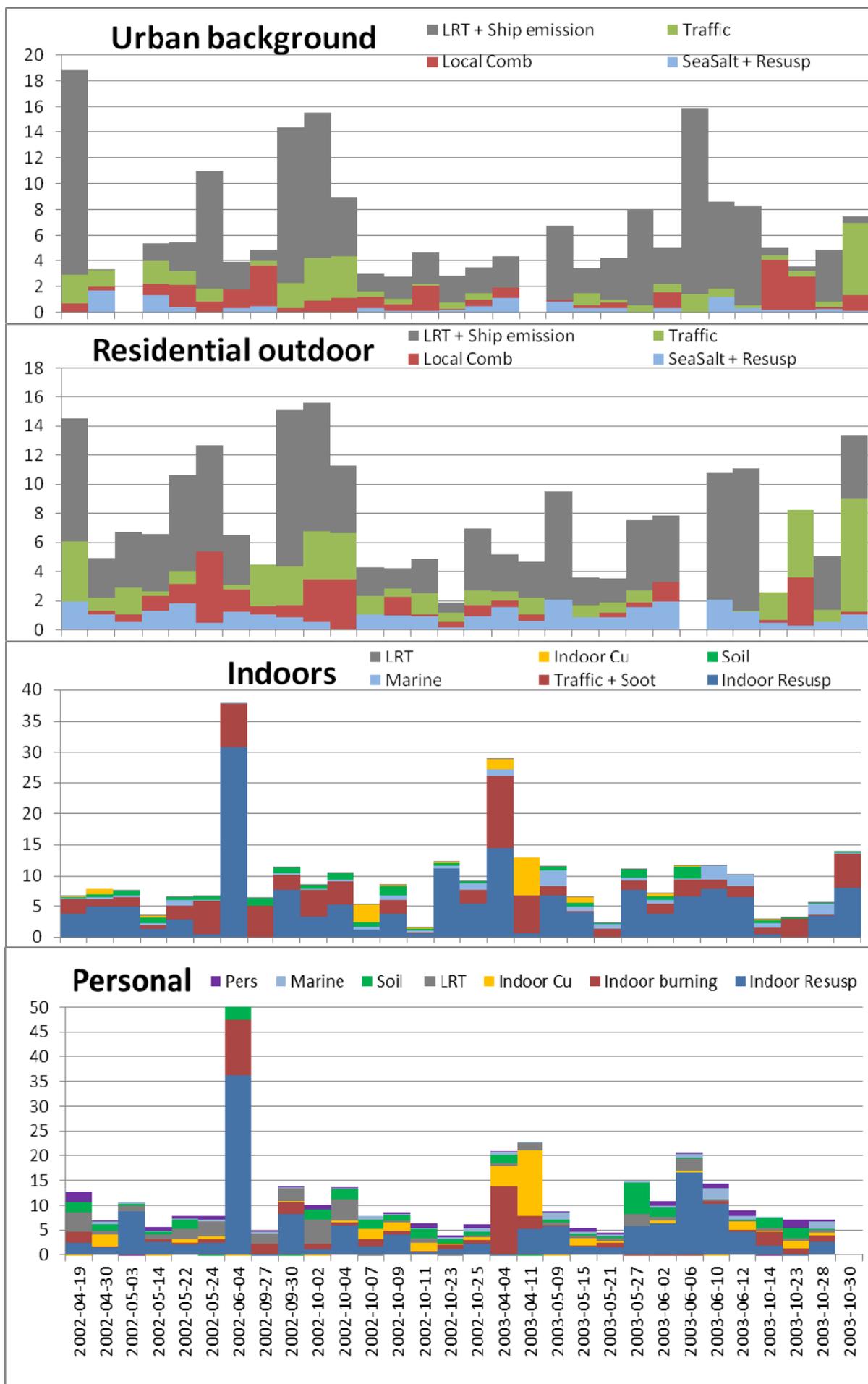


Fig. A1. Time series of the UB/O four factor models for Urban Background (top graph) and the Residential outdoor, the six factor model Indoors, and the seven factor model for Personal exposure. Units are in $\mu\text{g}/\text{m}^3$.