
"Seasonal variability and source apportionment of volatile organic compounds (VOCs) in the Paris megacity (France)" by A. Baudic et al.

Authors' Responses to the Handling Co-Editor Dr. Claire Reeves

All the authors would like to thank the Handling Co-Editor, Dr. Claire Reeves, for her general feedback. We apologize for not enough clarifying the issue concerning analytical uncertainties. Minor revisions were taken into account when preparing the new version of this manuscript. In the present document, authors' responses to specific comments addressed by the Co-Editor are mentioned in **blue**, while changes made to the revised manuscript are shown in *italic.*

Minor Revision

Section 2.3.2. You were asked by reviewer #1 to explain how the uncertainties were calculated. For the NMHC measurements there were 2 instruments, one run by LCSE at LHVP and one run by AIRPARIF at Les Halles. In your response to the reviewer you explained that there was insufficient information provided by AIRPARIF to quantify the uncertainties, however you did not explain how the uncertainties in the LHVP data were calculated. The second paragraph is on the Les Halles data and the last sentence states that the analytical uncertainty of "all" data was estimated. It is not clear if this refers to the LHVP data and based on your response to the reviewer, I would expect the uncertainties in the LHVP and Les Halles data to have been derived differently. Also since you lacked sufficient information from AIRPARIF to quantify the uncertainties, you should make this clear in the text. When you say that the uncertainties were estimated to be 15% in agreement with previous studies, it is not clear if they were set at 15% to be in agreement or somehow estimated to be 15% which happened to be in agreement with previous studies.

Throughout this study, two different automated Gas Chromatographs equipped with a Flame Ionization Detector (GC-FID) were used. The first one run during the MEGAPOLI field campaign (January-February 2010) and was operated by LSCE. The same instrument has already been used during another field campaign (AEROCOV program, 2007). The description of experimental set-up and uncertainties evaluation can be found in Gros et al., 2011. Analytical uncertainties were estimated from laboratory tests to be ± 15 %. The second GC-FID run during the FRANCIPOL intensive campaign (March - November 2010) and was operated by the regional air quality monitoring network AIRPARIF. From previous experimental tests (not available to us), AIRPARIF provided NHMC measurements with an accuracy of ± 15 %, which happened to be in agreement with the value estimated for LSCE GC-FID.

Gros, V., Gaimoz, C., Herrmann, F., Custer, T., Williams, J., Bonsang, B., Sauvage, S., Locoge, N., d'Argouges, O., Sarda-Estève, R. and Sciare, J.: Volatile organic compounds sources in Paris in spring 2007. Part I: qualitative analysis, Environ. Chem., 8, 74-90, doi: 10.1071/EN10068, 2011.

In the revised manuscript, it now reads:

Section 2.3.2 – NMHC on-line measurements by Gas Chromatographs (GC)

First paragraph (P6 – L.22-31)

'The analytical uncertainties on LSCE measurements were estimated from laboratory tests (i.e. memory effects, repeatability, accuracy of the gas standard) to be ± 15 % (Gros et al., 2011).'

Second paragraph (P6 – L.32 – P7 – L.5)

'A comprehensive calculation of uncertainties could not be performed due a lack of sufficient information from AIRPARIF. Nevertheless, from previous experimental tests, NMHC measurements were provided with an accuracy of ± 15 %, which happened to be in agreement with the uncertainty value for the LSCE GC-FID.'

Technical Issues

All technical corrections (typing and/or grammar errors) were taken into account in the revised manuscript and are mentioned in blue in its marked-up version. The font size of the axes titles and labels for specific figures was increased in the last version of the manuscript.

1) Page 6, line 18. Relative Humidity should be all in lower case letters.
2) Page 7 lines 29 and 31 "back trajectories".
3) Page 8 line 26 "ranges".
4) Page 8 lines 27-28 "Despite this limitation".
5) Page 10, line 25. Please define w/w in the text.
6) Page 12, line 17 "composition of VOCs".
7) Page 12, line 32 "on VOC trends".
8) Page 12, line 33 "As OVOCs are", "to observe different"
9) Page 15, line 2 "good indicators of".
10) Page 15, line 21 "passenger cars".
11) Page 19, line 11 "shallower".
12) Page 22, lines 18-19 "These kinds of emissions have been".
13) Page 22, line 29 "temperatures recorded …. were".
14) Page 24, line 14 "and the rest of the world".
15) Make sure axes labels are large enough. I have concerns over those for Fig. 7c and Fig. 10.
16) Fig. 13 in the caption which pie chart is which (left and right).

“Seasonal variability and source apportionment of volatile organic compounds (VOC) in the Paris megacity (France)” by A. Baudic et al.

Authors’ Responses to Referee #1

We would like to thank the Referee #1 for her/his general feedback and each of her/his useful comments/questions for improving the quality of this manuscript. All of them have been taken into account when preparing the revised version of the manuscript. In the present document, authors’ answers to the specific comments addressed by Referee #1 are mentioned in **blue**, while changes made to the revised manuscript are shown in *italic.*

General comments

The presented manuscript presents data on NMHCs and OVOCs from two sampling sites in Paris covering about 10 months of quasi-continuous measurements as well as on-site measurements at three potential major source categories. Since long-term datasets covering more than several weeks (usually obtained from short campaigns) are quite rare, in particular for combined NMHCs and OVOCs, this dataset is very valuable. However, the presented paper needs some major improvement before it can be considered for publication in ACP.

The analysis of data was kept at a very descriptive level. Diurnal and annual variations are shown for single compounds and for the resolved factors from the PMF analysis. A detailed source apportionment has been performed and the conducted PMF seems to be solid. All applied methods are quite standard and widely used in the scientific community. But apart from the new data, none of the methods are original. The manuscript could profit a lot from deeper analysis of the very nice dataset. The authors mention that the originality of the work derives from using near-field speciation profiles to refine the identification of the profiles. It is a good approach to compare modeled profiles with real measurements and it definitely helps to apportion source profiles. Despite the fact that this is also not a new approach, since there is a lot of literature, where it has been done before, it is new for Paris, and thus it is very useful to have this additional information for the SA. The authors only mention the measured profiles in one half-sentence for motor vehicle sources and for the background factor and not at all in the apportionment section of the wood burning factor. For the latter factor, the agreement is only partially fulfilled, since a lot of ethane and propane is assigned to the factor by PMF but not detected in the fireplace experiment. The authors should discuss the discrepancy and make more use of the added value by the experiments.

The performed analysis of the wind fields and the usage of the simple wind roses for interpretation of air-mass origin are not sufficient. The authors should consider a more elaborate analysis of the air-mass history, e.g. by trajectory analysis/clusters or to take a look into general weather patterns to really draw conclusions.

The structure of the paper could be improved. Many parts of the results section, where interpretation and already discussion is performed, would better fit into the discussion section. Also some parts of the results could be a bit and more stringent and focused. The English throughout the manuscript is okay; however, some rephrasing could sharpen the expressions and bring the content a bit more to the point. Past tense should be used more consequently when methods and results etc. are described.

The large dataset presented here for Paris constitutes a valuable addition to the VOC apportionment literature as it combines non-methane hydrocarbons (NMHC) and oxygenated VOC (OVOC) measurements performed over an extended time period, almost a year – which is new for Paris. The objectives of this paper are to (i) assess VOC concentration levels measured in Paris for further comparisons in other urban areas, (ii) give some information about seasonal and diurnal variabilities of a selection of VOC and more particularly (iii) present a Source Apportionment (SA) analysis performed using the Positive Matrix Factorization (PMF) standard statistical method. In order to help strengthen the identification of factors/sources issued from PMF simulations, we used speciation profiles determined from near-field measurements (inside a highway tunnel, at a fireplace and from a domestic gas flue). PMF results therefore appear more robust when combining with source profiles studies.

We agree with the reviewer that near-field observations have not been enough presented in the manuscript. We only compared PMF factor profiles with speciated source profiles by exploring major compounds. In the revised version of the manuscript, we now discuss similarities and discrepancies between the two approaches (see comments n° 45 - 50 and sub-sections 3.4.1, 3.4.3, 3.4.6).

To explain the seasonal variability of some emission sources (especially, the PMF factor mixing natural gas and background influences), we had performed monthly wind roses, what was not objectively considered as sufficient. As advised, an elaborate analysis of the air-mass history using the HYSPLIT model was done. To determine main air-mass origins affecting the "Île-de-France" region, a clustering analysis was also undertaken (see comments n° 22, 31, 59). We note that Paris is mainly affected by clean air masses coming from the west and usually associated with local and regional pollution conditions. To a lesser extent, air masses originated from the north and northeast were especially observed during springtime and are typically associated with continental influences from the Eastern Europe, the Benelux area and the north of Germany.

In this new version of the manuscript, results and discussion sections were grouped with the purpose of avoiding redundancies and any overlapping. We chose to present, interpret and discuss VOC data and the modeled PMF factors variabilities and compare them with other SA studies performed in Europe and in the world within one main section. Following reviewer's suggestions, the authors made efforts to write a revised version of the manuscript with conciseness. Some rephrasing also aimed at bringing the content a bit more to the point.

Note that a number of changes have been made to address the issues that emerged from the reviewer. Several figures and tables were added and modified in the revised manuscript (e.g. air masses trajectory analysis – Fig. 5; comparison between "highway tunnel" and "motor vehicle exhaust" profiles – Fig. 9; relative and absolute contributions of reactivity of each PMF factor – Fig. 12) and in the supplementary (e.g. PMF profiles issued from simulations using the FRANCIPOL dataset & comparison of speciated PMF profiles issued from our two different datasets – Section S1; representativeness of meteorological parameters in 2010 – S4; mean absolute contributions of factors per month in 2010). Consequently, figures and tables numeration is now different in this new version.

Specific and technical comments:

1/ P2, L-26: 12 million inhabitants, omit "of".
Correction applied in the revised manuscript: *"Paris and its surroundings (also called the Île-de-France region) constitute the second largest European megacity with about 12 million inhabitants, representing 20% of the French national population distributed over only 2% of its territory (Eurostat, 2014)."*

2/ P3, L-16: Some citations could be helpful.
As suggested, a small selection of citations *(e.g. Crippa et al., 2013; Skyllakou et al., 2014; Ait-Helal et al., 2014; Beekmann et al., 2015 and references therein)* **was added to the revised manuscript. It comprises different scientific studies documenting gaseous and particulate compounds in ambient air within the MEGAPOLI project: variabilities, source apportionment, origins and modeling.**

3/ P3, L-28: Should be elaborated further in the source apportionment section.
The following sentence "The originality of this study stands in [...] to refine the identification of apportioned sources" was kept in the introduction in order to inform readers about the added value of near-field speciation profiles. This approach thus allows reinforcing the source apportionment analysis and the corresponding discussion in Section 3.

4/ P4, L-3ff: Some redundant information from the introduction.
The following sentence "Ambient air measurements of VOCs [...] within the framework of two different research projects" was removed entirely from the text to avoid creating duplication and confusion as main information were given in the introduction.

5/ P4, L-8: Please state which background site.
The authors should discuss and prove that it is valid to merge the two datasets from two locations. Emissions in urban areas can be quite heterogeneous upon receptor location, wind patterns etc. and some further explanation is needed.
As reported in Table 1, VOC measurements were undertaken at two different locations between March and November 2010. The first sampling site is the "Laboratoire d'Hygiène de la Ville de Paris" (LHVP) located in the southern part of Paris (13th district – 48°51'N, 02°20'E). LHVP is juxtaposed with a large public garden (called "Parc de Choisy") at approximately 400 meters from Place d'Italie (grouping a shopping centre and main boulevards). The second monitoring site is the "Les Halles" subway station (1st district – 48°51'N, 02°20'E) located about 2 kilometers away from LHVP.

Table 1. Summary table restating specific features of sampling locations considered in this study.

Site	Coordinates	Experiments	
LHVP	48°82'N – 02°35'E	MEGAPOLI (J-F 2010) FRANCIPOL (M-N 2010)	GC-FID + PTR-MS PTR-MS
Les 'Halles' subway station	48°514N – 02°20'E	FRANCIPOL (M-N 2010)	GC-FID

LHVP = Laboratoire d'Hygiène de la Ville de Paris.
J-F 2010 = January-February 2010 (MEGAPOLI experiment)
M-N 2010 = March-November 2010 (FRANCIPOL experiment)
GC-FID = Gas Chromatograph with a Flame Ionization Detector
PTR-MS = Proton Transfer Reaction – Mass Spectrometer

Due to the low intensity of the surrounding activities, these two locations correspond to actual and previous urban background stations of the Airparif regional monitoring air quality network and have been considered as such within a previous study

(Gros et al., 2011). This station typology is based on two main criteria: (1) the population density is at least 4 000 inhabitants per square kilometre within a 1-km radius of the station and (2) no major traffic road is located within 300 metres (criteria from 2008/50/CE European Directive).

The characterization of site typologies can be conducted by studying the nitrogen monoxide (NO): nitrogen dioxide (NO_2) ratio. NO is known to be a vehicle pollution indicator whereas NO_2 has also an important secondary fraction. To consider a station as an urban background site, the ratio *R* between annual average NO and NO_2 concentrations (NO/NO_2) should be less than 1.5 ppb/ppb as indicated in a report (Mathé, 2010) written at the national level for regulatory purposes. It compiles the criteria for air quality measurement site implementation.

We observe here a very close NO: NO_2 ratio (expressed as ppb: ppb) between both sites in 2010: 0.40 (LHVP) *versus* 0.38 (Les Halles). The same conclusion can be made for the years 2008 and 2009 (0.45 ± 0.01 for LHVP and 0.48 ± 0.01 ppb: ppb for "Les Halles"). With NO: NO_2 ratios very similar and less than 1.5, this confirms that these two locations have the same site typology and can be considered as urban background stations.

The concept of the spatial representativeness is also important in that context. Mathé (2010) considers that the stations should preferably have a radius of the area of representativeness from 100 meters to 2 kilometers or an area of several square kilometers for urban background sites. However, this representativeness area is dependent on many factors: the topographic location, the local variability of emissions, wind patterns and pollutants concerned.

To prove that it is valid to merge these two specific datasets from different locations, we opted for a comparison of PMF results derived from MEGAPOLI-FRANCIPOL (as performed in the present paper) and FRANCIPOL data files respectively. PMF modeling simulations were performed using only the FRANCIPOL dataset (it corresponds to 88% of the total data matrix) and operating with similar processing conditions: *(i)* data quality control, *(ii)* concentration and uncertainty files checking, *(iii)* same species categories assignment and *(iv)* the optimum factors number by exploring several statistical parameters: IM, IS, $Q_{true}/Q_{expected}$, R^2 for VOC. Just like final PMF results, a 6-factor solution was then chosen in terms of interpretability and fitting scores (see Figure 1). A good agreement was found for the "Motor Vehicle Exhaust", "Evaporative sources", "Wood Burning", "Biogenic" sources ($0.93 < R < 0.99$ -- $0.87 < R^2 < 0.98$). More significant differences were identified for "Solvents use" and "Natural Gas & Background" sources ($0.71 < R < 0.78$ – $0.51 < R^2 < 0.60$), especially for aromatic compounds (e.g. benzene, toluene, xylenes + C_8). A ~ 50 % loss of aromatics contributions is observed when considering the "MEGAPOLI-FRANCIPOL" dataset compared to the 'FRANCIPOL" one. It is compensated by a contribution of these compounds in the "Natural Gas and Background" source.

This observation can be explained by:

i. The challenge of splitting "Solvents use" and "Natural Gas & Background" sources when considering a 5-factor PMF solution (see Section A2 in the initial manuscript). PMF artifacts may occur.

ii. **The summer variability of the "solvents use" source is more meaningful when considering the "FRANCIPOL" dataset. Including the "MEGAPOLI-FRANCIPOL" dataset (or adding a winter influence) might down the weight of aromatics in the industrial source profile.**

Based on all of these different reasons, we believe that the merging of the two datasets is quite valid.

The sub-section 2.1 (e.g. Sampling sites description) was revised for a better understanding of the involved sampling sites. To make the manuscript more readable, another sub-section was created to discuss the representativeness of sampling stations. The graphs showing the comparison of PMF results between MEGAPOLI-FRANCIPOL and FRANCIPOL (Figure 2) have been added in the supplementary section S1.

In the revised manuscript, it now reads:

"Due to the low intensity of the surrounding activities, the LHVP and Les Halles sampling sites were considered as urban background stations by AIRPARIF and by previous scientific studies (Favez et al., 2007; Sciare et al., 2010; Gros et al., 2011). In accordance with the 2008/50/EC European Directive, this station typology is based on two main criteria: (1) the population density is at least 4 000 inhabitants per square kilometre within a 1-km radius of the station and (2) no major traffic road is located within 300 metres.

This characterization of site typologies can be conducted by studying the nitrogen monoxide (NO): nitrogen dioxide (NO_2) ratio. NO is known to be a vehicle pollution indicator whereas NO_2 has an important secondary fraction. To consider a station as an urban background site, the ratio R between annual average NO and NO_2 concentrations (NO/NO_2) should be less than 1.5 ppb/ppb as indicated in a report (Mathé, 2010) written at the national level for regulatory purposes.

We observed a very close NO: NO_2 ratio (expressed as ppb: ppb) between both sites in 2010: 0.40 (LHVP) versus 0.38 (Les Halles). The same conclusion can be made for the years 2008 and 2009 (0.45 ± 0.01 for LHVP and 0.48 ± 0.01 ppb: ppb for "Les Halles"). With NO: NO_2 ratios very similar and less than 1.5, this confirms that these two locations have the same site typology and can be considered as urban background stations.

To prove that it was valid to merge these two specific datasets from different locations, we opted for a comparison of PMF results derived from MEGAPOLI-FRANCIPOL (as performed in the present paper) and FRANCIPOL data files, respectively. PMF modeling simulations were performed using only the FRANCIPOL dataset. A good agreement was found for the majority of the emission sources. The graphs showing the comparison of PMF results between MEGAPOLI-FRANCIPOL and FRANCIPOL are reported in the Supplementary Section S1".

Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe. Official Journal of the European Communities. 11 June 2008.

Mathé, F. (2010). Evolution de la classification et des critères d'implantation des stations de mesure de la qualité de l'air – Participation à la réactualisation du guide de classification des stations, LCSQA-Mines Douai report.

Figure 1 - Source composition profiles of the 6-factor PMF solution derived from the FRANCIPOL dataset (March-November 2010). The concentrations (in µg.m^{-3}) and the percent of each species apportioned to the factor are displayed as a pale blue bar and a color box, respectively. **(a)** F1 – Motor Vehicle Exhaust; **(b)** F2 - Evaporative sources; **(c)** F3 - Wood Burning; **(d)** F4 - Biogenic; **(e)** F5 – Solvents use source; **(f)** F6 - Natural Gas and Background.

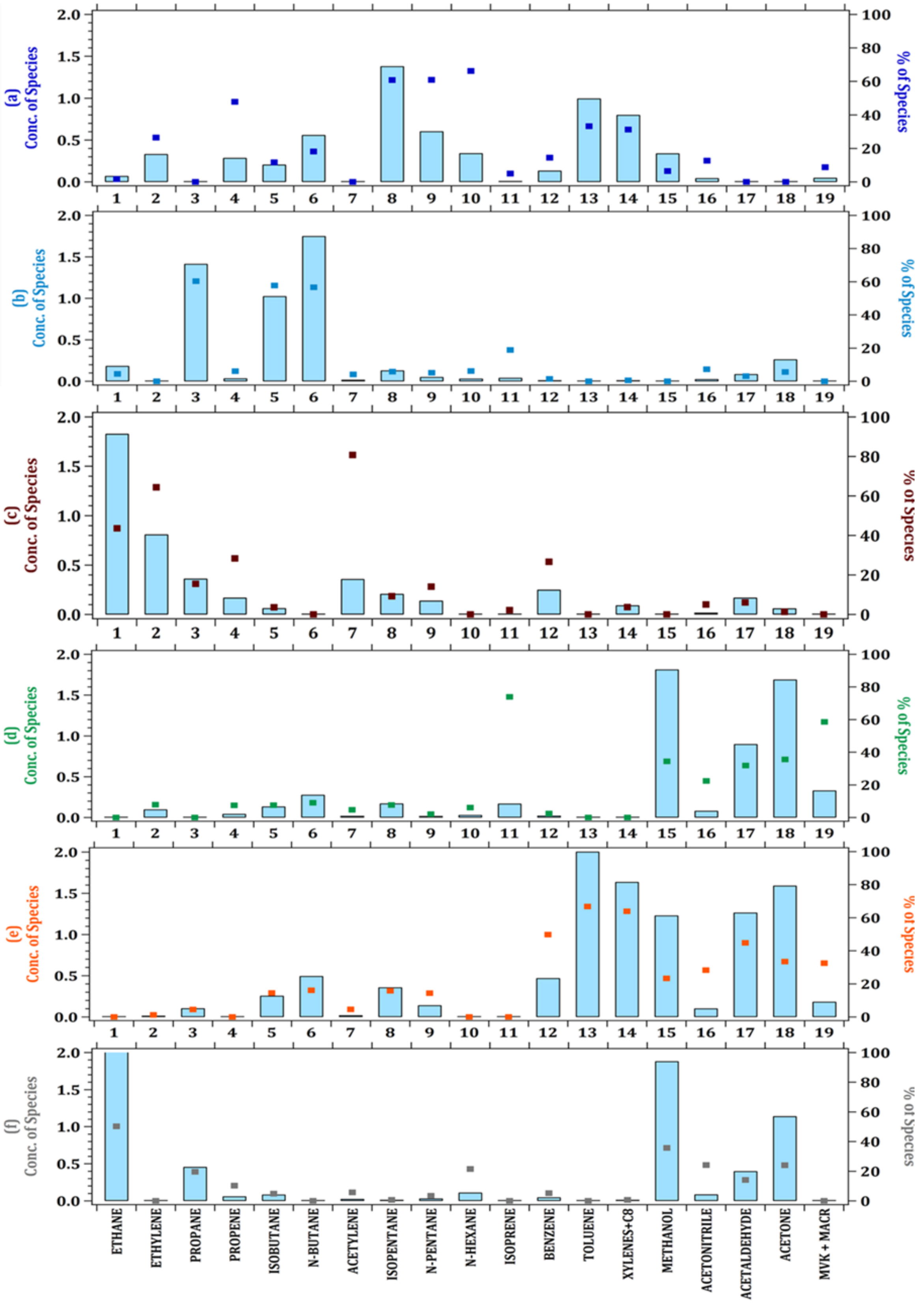
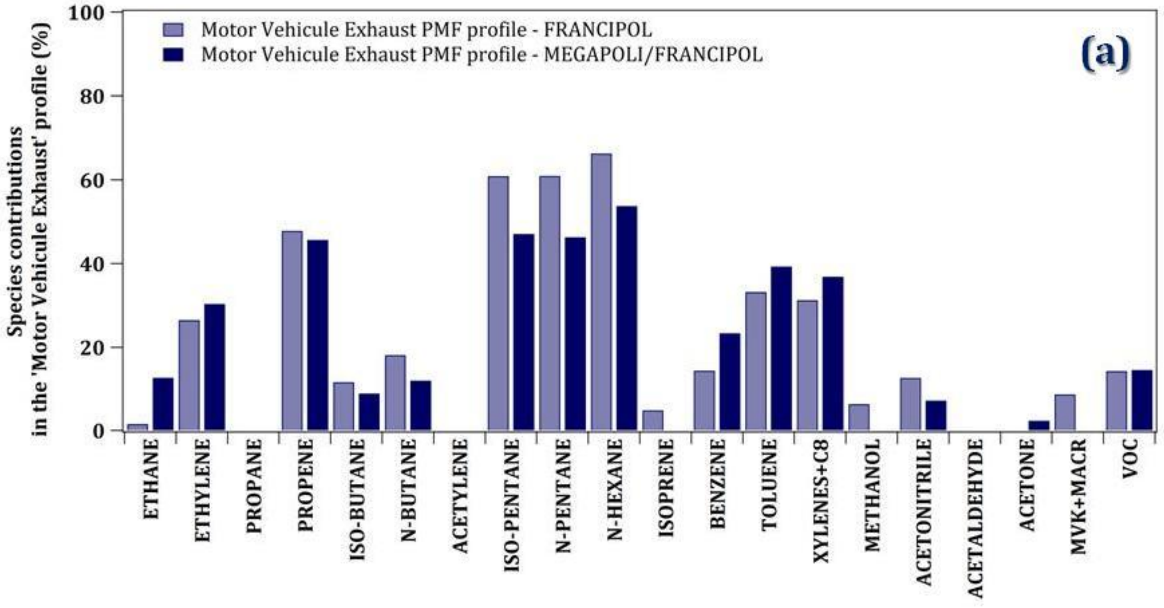
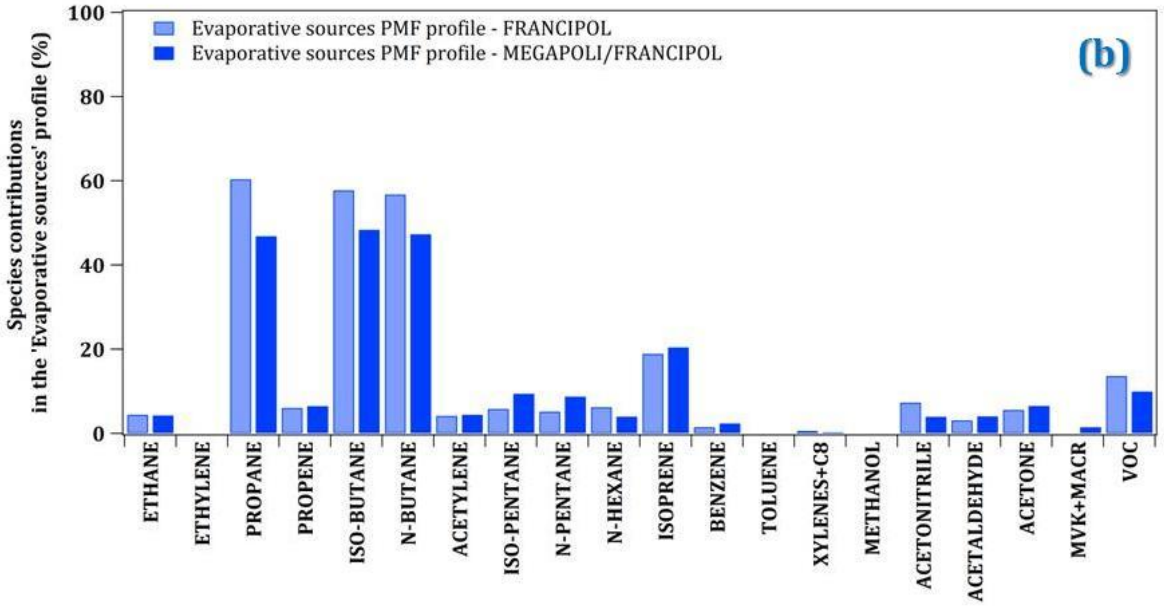
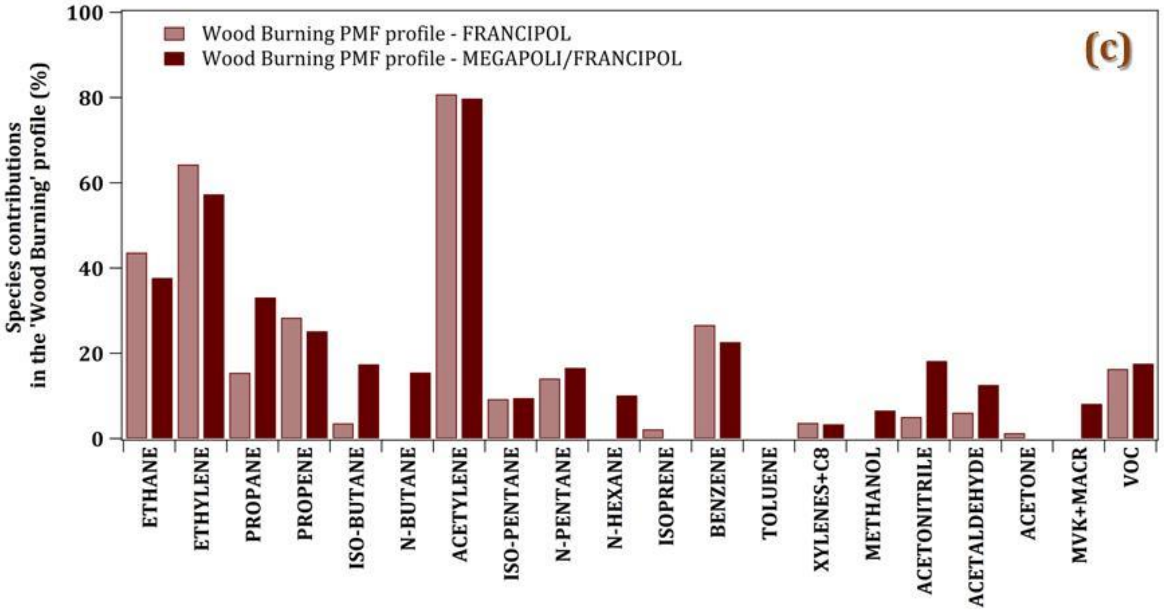
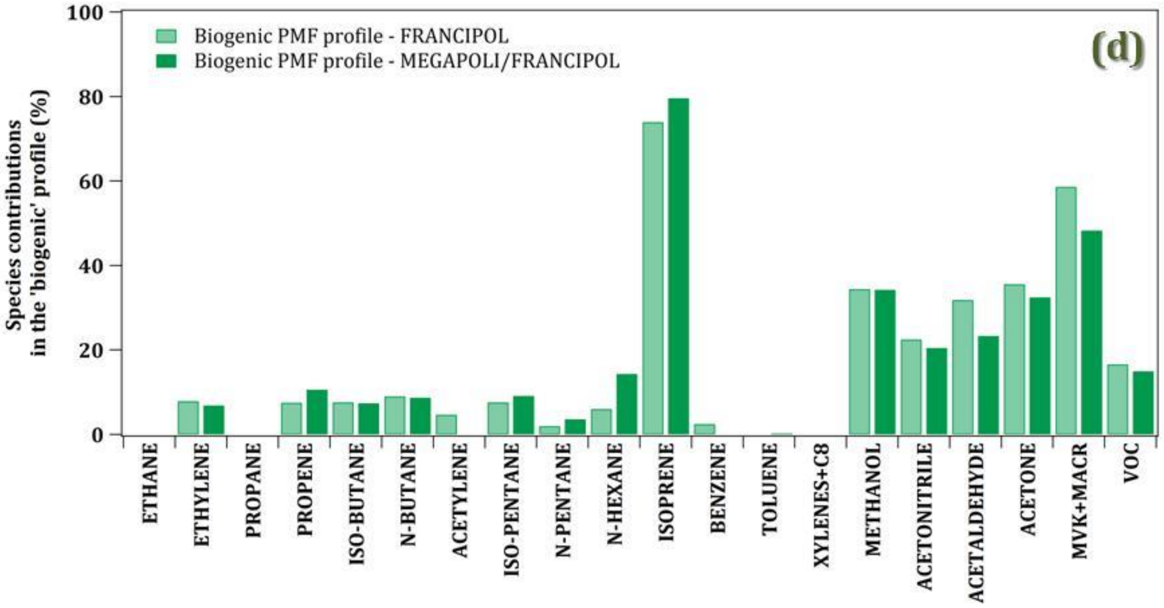


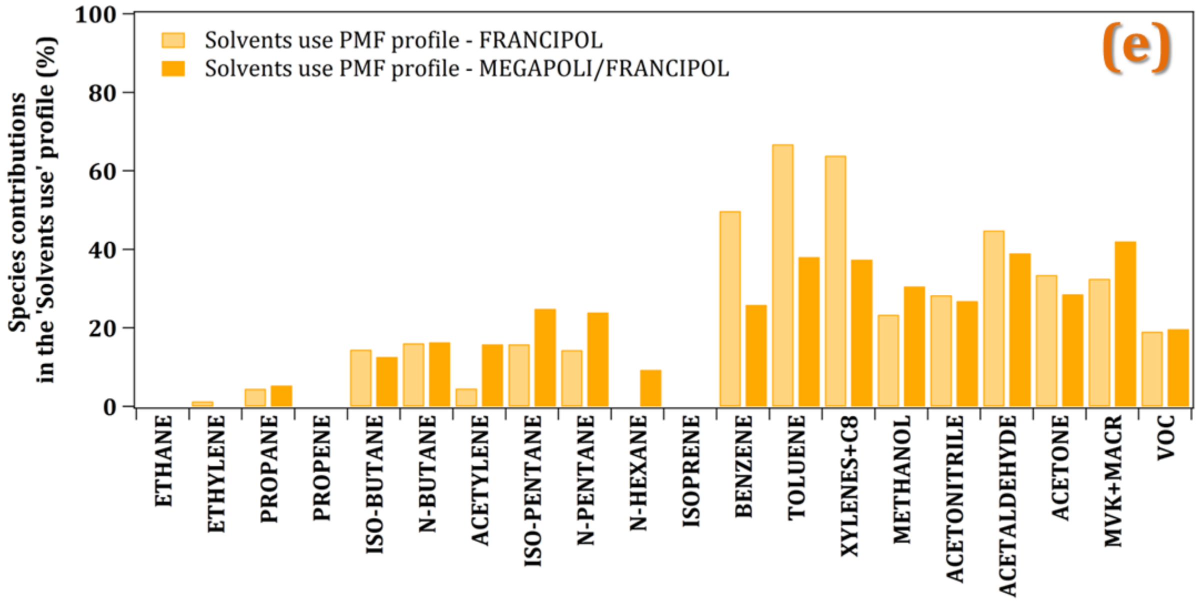
Figure 2 - Comparison of speciated PMF profiles issued from both FRANCIPOL AND MEGAPOLI/FRANCIPOL datasets. The contributions of species in the PMF profile are expressed in %. **(a)** F1 – Motor Vehicle Exhaust; **(b)** F2 - Evaporative sources; **(c)** F3 - Wood Burning; **(d)** F4 - Biogenic; **(e)** F5 – Solvents use source; **(f)** F6 - Natural Gas and Background**.**

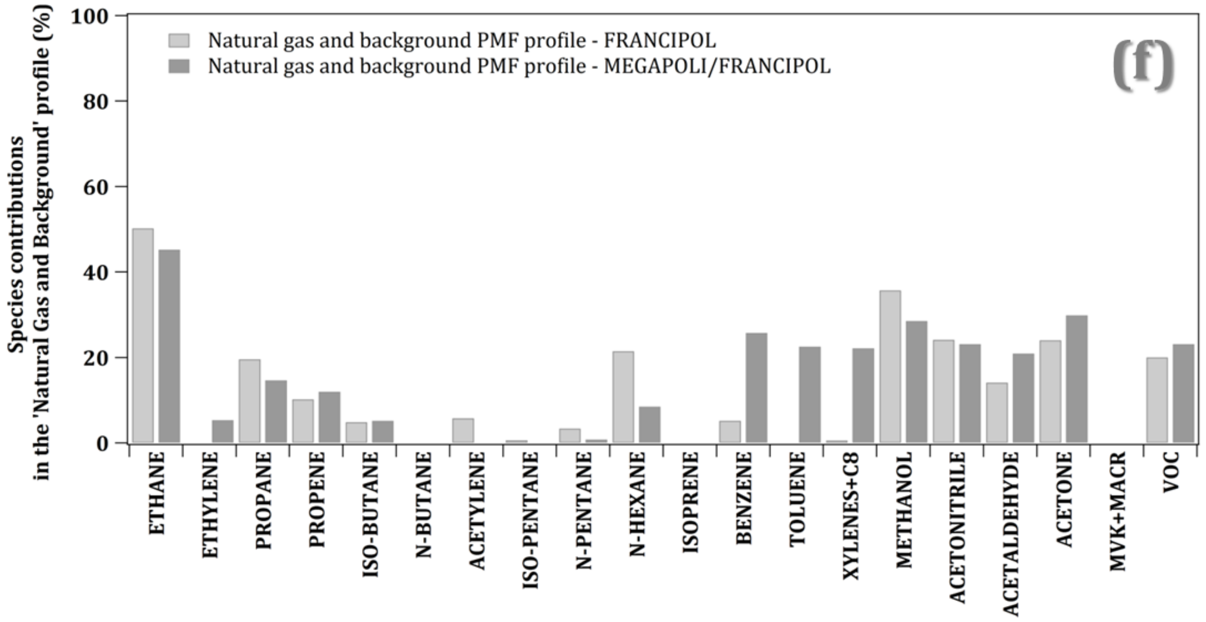





Species contributions in the 'Wood Burning' profile (%)
100
80
60
40
20
0
Wood Burning PMF profile - FRANCIPOL
Wood Burning PMF profile - MEGAPOLI/FRANCIPOL
(c)
ETHANE
ETHYLENE
PROPANE
PROPENE
ISO-BUTANE
N-BUTANE
ACETYLENE
ISO-PENTANE
N-PENTANE
N-HEXANE
ISOPRENE
BENZENE
TOLUENE
XYLENES+C8
METHANOL
ACETONITRILE
ACETALDEHYDE
ACETONE
MVK+MACR
VOC


Species contributions in the 'biogenic' profile (%)
100
80
60
40
20
0
Biogenic PMF profile - FRANCIPOL
Biogenic PMF profile - MEGAPOLI/FRANCIPOL
(d)
ETHANE
ETHYLENE
PROPANE
PROPENE
ISO-BUTANE
N-BUTANE
ACETYLENE
ISO-PENTANE
N-PENTANE
N-HEXANE
ISOPRENE
BENZENE
TOLUENE
XYLENES+C8
METHANOL
ACETONITRILE
ACETALDEHYDE
ACETONE
MVK+MACR
VOC


Species contributions in the 'Solvents use' profile (%)
100
80
60
40
20
0
Solvents use PMF profile - FRANCIPOL
Solvents use PMF profile - MEGAPOLI/FRANCIPOL
(e)
ETHANE
ETHYLENE
PROPANE
PROPENE
ISO-BUTANE
N-BUTANE
ACETYLENE
ISO-PENTANE
N-PENTANE
N-HEXANE
ISOPRENE
BENZENE
TOLUENE
XYLENES+C8
METHANOL
ACETONITRILE
ACETALDEHYDE
ACETONE
MVK+MACR
VOC


Species contributions in the 'Natural Gas and Background' profile (%)
100
80
60
40
20
0
Natural gas and background PMF profile - FRANCIPOL
Natural gas and background PMF profile - MEGAPOLI/FRANCIPOL
(f)
ETHANE
ETHYLENE
PROPANE
PROPENE
ISO-BUTANE
N-BUTANE
ACETYLENE
ISO-PENTANE
N-PENTANE
N-HEXANE
ISOPRENE
BENZENE
TOLUENE
XYLENES+C8
METHANOL
ACETONITRILE
ACETALDEHYDE
ACETONE
MVK+MACR
VOC

6/ P6, L-8f: Calibration once a month seems little; how was the stability of the systems within the months checked? How the zeroing was performed?
Full calibrations were performed once a month by the regional air quality network monitoring AIRPARIF using a standard gas mixture containing only propane (25.5 ppb, Air Liquide). This certified gas is sampled in the same conditions than ambient air (e.g. 30 min. sampling and 30 min. analysis) during 6 hours. Response coefficients were calculated for each compound using the Effective Carbon Number (ECN) method, described in Badol et al. (2004).

In addition, the stability of GC-FID systems was checked over time by controlling chromatographic profiles, the baseline and the stability of intrinsic parameters (pressure, temperature, leak tests). The zeroing was performed every 6 months using a zero air bottle in order to detect any instability or problem with the system.

Correction applied in the revised manuscript: *"A full calibration was performed once a month with a standard gas mixture containing only propane during 6 hours. As the FID response is proportional to the Effective Carbon Number (ECN) in the molecule, calibration coefficients were calculated for each compound and regularly checked so that they drifted no more than ± 5% (tolerance threshold). In addition, a zeroing was carried out every 6 months using a zero air bottle in order to detect any instability or problem with the GC-FID system. LoD were assessed at 0.024 µg.m^{-3} for all the selected compounds, except for n-hexane (0.013 µg.m^{-3}) (0.023 – 0.004 ppb)."*

Badol, C., Borbon, A., Locoge, N., Léonardis, T. and Galloo, J.-C.: An automated monitoring system for VOC ozone precursors in ambient air: development, implementation and data analysis, Anal. Bioanal. Chem., 378, 1815-1824, 2004.

7/ P7, L-5: Since it is only a very small selection of papers out of the many studies from urban areas, I would put in an "e.g.," before the citations.
As suggested, an "e.g." was added before the citations, both for studies from urban and rural areas.

8/ P7, L-6: Not too many studies were performed in rural areas, but the ones could be mentioned, e.g. Lanz et al. 2009 ACP (doi: 10.5194/acp-9-3445-2009) and Leuchner et al.2015 ACP (doi: 10.5194/acp-15-1221-2015).
The two suggested references were added to the selection of papers of the studies from rural areas. Subsequently, bibliographic references were also updated.

Correction applied in the revised manuscript: *" Concerning VOC, PMF studies have been conducted in urban (e.g Brown et al., 2007 – LA, USA; Lanz et al., 2008 – Zürich, Switzerland; Morino et al., 2011 – Tokyo, Japan; Yurdakul et al., 2013 – Ankara, Turkey) and rural areas (e.g. Sauvage et al., 2009 – France; Lanz et al., 2009 – Jungfraujoch, Switzerland; Leuchner et al., 2015 – Hohenpeissenberg, Germany)."*

9/ P7, L-20: "acetonitrile".
The word "acetonitrile" was correctly modified in the revised manuscript.

10/ P7, L-22: "except".
The word "expected" was substituted by "expect".

11/ P7, L-25: Why was the proportion of missing values estimated? It should be know exactly.
The virtual number "-999" was defined as the missing value indicator in the concentration input file. The PMF model was programmed to find all instances of "-999" in the data file. It is also able to provide both the proportions of raw and modeled samples for each species. The percentage of missing values is thus exactly known. It is ranging from 19 % (especially for compounds measured by PTR-MS) to 41 % (only for isoprene). To avoid any misunderstanding, the corresponding sentence was rephrased.

Correction applied in the revised manuscript: *"The proportion of missing values ranges from 19 % (especially for compounds measured by PTR-MS) to 41 % (only for isoprene)".*

12/ P7, L-27: "constraints"
It is quite problematic that isoprene had such a high amount of missing values thus it is nor too surprising that the biogenic factor could not be resolved very clearly from some substances.
On one hand, the word "constrains" was substituted by "constraints".

On the other hand, the high proportion of missing values for isoprene is partly due to analytical issues on the operating GC-FID in July 2010. Each missing data point was substituted with the median concentration of this species over all the measurements and associated with an uncertainty of four times the species-specific median (as suggested in Norris et al., 2014). Missing values are usually downweighted by large uncertainty values. In this case, concentrations are typically below uncertainties. Subsequently, they were determined to have a low Signal-to-Noise (1.7 for isoprene here – the lowest ratio of our species selection). This implies significant difficulties in modeling this compound, as demonstrated by statistical parameters (R^2 = 0.29 – Slope = 0.53). The limitations of the PMF model to simulate isoprene have been kept in mind within the reconstructed results description and discussions.

13/ P7, L-32: Since most of PMF studies describe their results in ppb units, it might be nice to have some information how much ppb/ppt the units equal for comparison (throughout the manuscript).
For further qualitative and quantitative comparisons, VOC concentrations are expressed both in micrograms per cubic meter (µg.m^{-3}) and in parts per billion (ppb). Conversion factors (µg.m^{-3} to ppb) were calculated considering the molecular mass of a gaseous compound (g.mol^{-1}) over the air molar volume (24 L at 20°C).

Correction applied in the revised manuscript: Throughout the manuscript, each concentration/value/result initially in µg.m^{-3} is also mentioned in ppb (except for PMF results). The conversion factor for each species is reported in Table 1.

14/ P8, L-1: How were the uncertainties determined? Was there a standard procedure for each substance following guidelines from GAW or ACTRIS?
The determination of uncertainties on data measured both by PTR-MS and GC-FID did not specically follow a standard procedure from Global Atmosphere Watch (GAW) or Aerosols, Clouds, and Trace gases Research InfraStructure (ACTRIS) network.

(O) VOC data measured by PTR-MS were provided both by LCE (Megapoli period) and LSCE (Francipol period). From the available data, uncertainty calculations were estimated by taking into account errors on standard gas, calibrations, blanks,

reproducibility/repeatability, linearity and relative humidity parameters. The measurement uncertainty was calculated from the theory of propagation of the errors and was found to be ~ 20% (Gros et al., 2011; Dolgorouky et al., 2012). Finally, PMF empirical tests were performed and aimed at changing input analytical uncertainty from 20% to 25% for species measured by PTR-MS. As a consequence of these tests, no significant change was observed in PMF modeling results.

VOC data measured by GC-FID during the Francipol campaign were provided by the regional air quality monitoring network AIRPARIF. Nevertheless, no uncertainty calculations were supplied to us. As these calibration data were no longer available (processing software changes, old releases not adequate, compatibility issues), we could not make an exhaustive uncertainty calculation. In this present study, GC-FID data uncertainties were estimated at 15% in agreement with previous studies (Gros et al., 2011; Gaimoz et al., 2011; Dolgorouky et al., 2012).

Correction applied into the methods section (Subsection 2.3.2 – VOC measurements using a Proton Transfer Reaction-Mass Spectrometer (PTR-MS)): *"The analytical uncertainty on all data was estimated by taking into account errors on standard gas, calibrations, blanks, reproducibility/repeatability, linearity and relative humidity parameters. The measurement uncertainty was estimated at ± 20% in agreement with previous studies (Gros et al., 2011; Dolgorouky et al., 2012)."*

Correction applied into the methods section (Subsection 2.4.2 – NMHC on-line measurements by Gas Chromatography (GC)): *"The analytical uncertainty on all data was estimated at ± 15% in agreement with previous studies (Gros et al., 2011; Gaimoz et al., 2011, Dolgorouky et al., 2012)."*

15/ P8, L-4: What other tests were performed? Might be interesting for other readers?
16/ P8, L-10: It is not entirely clear what exactly was done and what modeling parameters are meant here.
As previously reported, isoprene exhibits a high proportion of missing values (41 %). This actually means that 59% of isoprene data are modeled by the PMF technique, what is not objectively considered as sufficient. To address this problem, input values of isoprene were modified when detecting missing data. We have investigated the possibility of replacing missing values with data more appropriate than the median. Indeed, a "virtual" averaged pattern was calculated from 1-h real samples observed in June and August (to keep the summer variability of isoprene). Another option was to preserve raw data of isoprene (no subsequent changes) and to increase the analytical uncertainty initially estimated at 15% (gradually from 15 % to 30%) – instead of categorizing isoprene as *weak* (as uncertainties are tripled > lower Signal to Noise). This was intended to better display this biogenic compound. PMF simulations were performed by considering these different options. As a consequence of these tests, no significant improvement on the quality of modeling isoprene was observed. Regular statistical parameters such as R^2, slope, slope/intercept SE (Standard Error) were used to draw such conclusions. As empirical tests have not helped, isoprene is still categorizing as *strong*.

Correction applied in the revised manuscript: *"To address this lack of isoprene data, several empirical tests (e.g. simulating an averaged seasonal/diurnal cycle of isoprene or increasing the analytical uncertainty of raw data from 15 % up to 30 %) were conducted within PMF simulations with the aim of better modeling the variability of this compound. As a consequence of these tests, no significant improvement on the quality of modeling isoprene was observed. Finally, isoprene is still categorizing as strong here".*

17/ P8, L-14: For what reason was the total VOC mass included in the model? What is the benefit of it? In the results it basically looks like this parameter is just apportioned equally to all factors more or less.
The total VOC mass was included in the PMF model in order to directly estimate VOC source contributions (as a whole) and not just to quantify the proportion of each species in a factor profile (when considering individually species). ∑VOC was defined as *Total Variable* and automatically categorized as *weak* to lower its influence in the final PMF results. Adding the ∑VOC variable allowed to ensure a good diagnosis of final PMF results (for instance, VOC predicted *vs.* VOC measured – Fig. A.1 in the initial version of the manuscript).

Considering ∑VOC in the model is usually made in PMF studies because it is worthwhile when comparing the contribution of VOC sources with modeling results from previous studies and emission inventories, in which the sum of all individual VOC species emitted from a specific source is taken into account.

18/ P9, L-5: Omit "specially".
Correction applied in the revised manuscript: *"The speciated profiles of these different anthropogenic sources and their representativity are given here."*

19/ P9, L-20: "Omit" instead of "get out from".
Correction applied in the revised manuscript: *"In order to omit any local background, nighttime values (as suggested in Ammoura et al., 2014) were subtracted from the peak VOC concentrations."*

20/ P9, L-22: What is meant by "w/w"?

The term "w/w" means "weight of each compound" out of "weight of the total VOC mass".

21/ P10, L-23: Please re-phrase "weather indicators".

Correction applied in the revised manuscript: *"Meteorological parameters (e.g. temperature, relative humidity, rainfall, sun exposure, boundary-layer height, wind speed and direction) are known to be key factors governing seasonal and diurnal variations of air pollutant levels."*

22/ P11, L-3ff: The analysis of air-mass origin need to be more profound; only wind direction does not explain the origin of the air masses. To determine if air masses were originating from the clean marine BL or from LRT or deriving from industrial areas from the PBL of central Europe a trajectory analysis (e.g. trajectory clusters) or at least an analysis of general weather patterns could bring further information and support to the assumptions in the manuscript. The wind rose plots could go to the supplement or replaced by a more profound analysis.

As recommended in general comments, an elaborate analysis of the air-mass history was performed and presented in the revised manuscript as described below.

Correction applied into the methods section (Subsection 2.3.3 – Additional data available): *"In order to determine the air-masses origin, 5-day back trajectories were calculated every 3 hours from the PC based version of the HYbrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model (Stein et al., 2015) with Global Data Assimilation System (GDAS) meteorological field data. Back trajectories were set to end at Paris coordinates (49°02'N, 02°53'E) at 500 m a.g.l."*

Within this study, 2 164 back trajectories were calculated every 3 hours using the HYSPLIT model. A clustering algorithm was used to group these trajectories depending on their direction, their speed and their altitude (Sirois and Bottenheim, 1995; Sauvage et al., 2009). For each backtrajectory, *xyz* (latitude, longitude, altitude) coordinates of the trajectory endpoint were used as input variables for the clustering technique. To determine the optimum number of clusters (N), plotting the percentage of change in Total Spatial Variance (TSV) against the number of cluster was used. A slope break is observed for N=4 where the change in TSV (%) is at its maximum (~70%). The clustering analysis of backtrajectories was performed throughout the study period (J-N 2010) and is presented in Figure 3.

Paris is mostly influenced over the year by air masses coming from the west (ca. 62 % - CL#1 + CL#2) and usually associated with clean marine air influences from the Atlantic Ocean. They are typically representative of local and regional pollution conditions, as already observed in Gros et al., 2011; Gaimoz et al., 2011; Dolgorouky et al., 2012 and Petit et al., 2015. To a lesser extent, Paris can be affected by northeast air masses (26%) originating from Eastern France, the Benelux area, Northern Germany and are indicative of continental imports of long-lived pollutants. North air masses (passing through the United Kingdom and the north of France) can also affect the receptor site (11%). Air masses coming from the west (LRT and SRT) are generally observed in summer and autumn (ca. 32 – 41%) whereas northeast air masses (CL#3) are found to be significant in winter (34%) - specifically in February - and most frequently in spring (ca. 40%).

This trajectory analysis was then used in Section 3 when discussing the air-mass origin of the background factor.

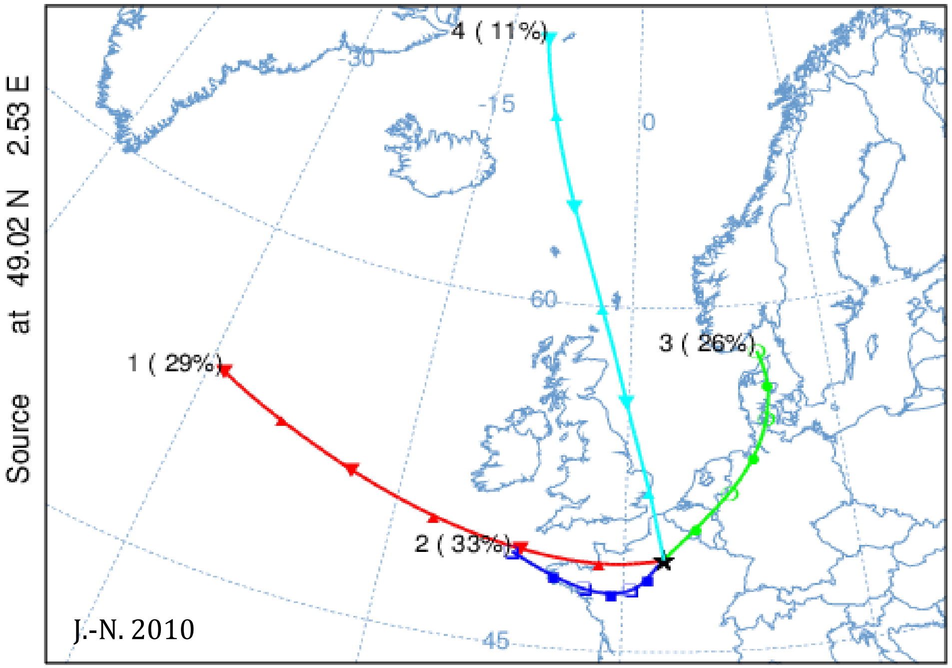
Correction applied in the revised manuscript: *"For the year 2010, Paris was mainly influenced by air masses coming from the west (62%) and usually associated with clean marine air influences from the Atlantic Ocean (see Fig.5).They are typically representative of local and regional pollution conditions, as already observed in Gros et al., 2011; Gaimoz et al., 2011; Dolgorouky et al., 2012 and Petit et al., 2015. To a lesser extent, Paris can be affected by northeast air masses (26%) originating from Eastern France, the Benelux area, Northern Germany and are indicative of continental imports of long-lived pollutants (Gaimoz et al., 2011). Air masses coming from the west are generally observed in summer and autumn (32-41%) whereas northeast air masses are found to be significant in winter (34%) and most frequently in spring (ca. 40%) due to the stagnation of an anticyclone surrounding the British Isles (Monthly weather report for Paris and its surroundings during April 2010, Météo-France) during that period."*

Sauvage, S., Plaisance, H., Locoge, N., Wroblewski, A., Coddeville, P and Galloo, J.C.: Long term measurement and source apportionment of non-methane hydrocarbons in three French rural areas, Atm. Env., 43, 2430-2441, doi: 10.1016/j.atmosenv.2009.01.001, 2009.

Sirois, A. and Bottenheim, J.W.: Use of backward trajectories to interpret a 5-Year Record of PAN and O3 ambient air concentrations at Kejimkujik national Park, Nova-Scotia, Journal of Geographical Research-Atmospheres, 100 (D2), 2867-2881, 1995.

Stein, A.F., Draxler, R.R., Rolph, G.D., Stunder, B.J.B., Cohen, M.D. and Ngan, F.: NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System, Bull. Am. Meteorol. Soc., 96(12), 2059-2077, doi:10.1175/BAMS-D-14-00110.1, 2015.

.



Cluster CL#	Air mass origins	Year (J-N)	Winter	Spring	Summer	Autumn
CL#1	LRT (West)	29 %	24 %	25 %	32 %	32 %
CL#2	SRT (West)	33 %	29 %	25 %	41 %	33 %
CL#3	Continental influences (NE)	26 %	34 %	39 %	19 %	22 %
CL#4	North	11 %	13 %	11 %	8 %	12 %

LRT: Long Range Transport
SRT: Short Range Transport

Figure 3 - Average trajectories obtained after clustering analysis and the relative proportion of clusters (%) over the year and per season.

23/ P11, L-12: Omit "From this table, it was observed that".
Correction applied in the revised manuscript: *"The average composition of VOCs was mainly characterized by oxygenated species (0.7 – 5.9 µg.m*$^{-3}$ *(0.4 – 4.4 ppb); 36.5 % of the TVOC mass), alkanes (0.5 – 4.6 µg.m*$^{-3}$ *(0.1 – 3.7 ppb); 39.1 %) followed by aromatics (1.1 – 3.3 µg.m*$^{-3}$ *(0.3 – 0.9 ppb); 16.9 %) and to a lesser extent by alkenes, alkynes and dienes (0.3 – 1.6 µg.m*$^{-3}$ *(0.1 – 1.3 ppb); 7.5 %)."*

24/ P11, L-23f: What is the reason for the difference? Meteorological conditions? Should be discussed in the discussion section.
Mean concentration levels of aromatics and OVOCs measured by PTR-MS in 2007 (Gaimoz et al., 2011) were reported in Table 2 (of the initial manuscript) and compared with our study (2010). For a relevant comparison, VOC measurements performed during the same sampling time (25 May – 14 June) were considered. A significant decrease in VOC concentrations was observed between spring 2007 and spring 2010, especially for aromatics (xylenes + C_8, benzene). This finding is in agreement with decreasing downward trends (2003-2013) of aromatics concentrations observed during springtime in Paris (Waked et al., 2016) and more globally with the gradual decline of NMHC emissions over the last few years in France.
Note that the study of Waked et al. (2016) did not examine the trends of oxygenated compounds.

Correction applied in the revised manuscript: *"These differences in aromatics concentration levels are consistent with decreasing downward trends of NMHC recorded during springtime in Paris (Waked et al., 2016). We note that the study of Waked et al. (2016) on VOC trends in Paris only concerns NMHC and not oxygenated compounds. As these OVOCs are significantly impacted by biogenic and secondary sources, it is not surprising to observe a different variation between 2007 and 2010."*

Waked, A., Sauvage, S., Borbon, A., Gauduin, J., Pallares, C., Vagnot, M.-P., Léonardis, T. and Locoge, N.: Multi-year levels and trends of non-methane hydrocarbon concentrations observed in ambient air in France, Atmos. Environ., 141, 263-275, 2016.

25/ P11, L-34: Last sentence is quite generic.
The corresponding sentence was removed from the revised manuscript.

26/ P11, L-29: "global studies".
Correction applied in the revised manuscript: *"Average VOC concentrations were also calculated in line with sampling periods of the other European and global studies over different years (see Table 2)."*

27/ P12, L-3: Omit "time".
The title of the subsection 3.3 was modified in the revised manuscript.
Correction applied: *"Seasonal and diurnal variations".*

28/ P12, L-15: Better to use a 24h time instead of am/pm, avoids confusion.
When time is mentioned throughout the manuscript, the 24-h format is used now.

29/ P12, L-17ff: The increase occurs until 21h. And after that the levels stay quite high. Is there another (meteorological) reason for this? Interpretation of the graphs should be done in the discussion section.

CO emissions are mainly related to road-traffic and/or wood burning activities, for which they would be more intense in autumn and in winter, respectively. The diurnal variability of CO is characterized by a "double wave" profile. A first peak is observed between 08-10 h and a second one from 18 to 21 h. After 21 h, CO levels stay quite high (240-260 ppb). This finding can be explained by on-going emissions but also by atmospheric dynamics. Indeed, the PBL is lower at night. This implies an accumulation of CO (and other pollutants) into the atmosphere. In addition, a few photochemical reactions occur due to lower OH concentrations and solar radiations. All these features explain quite high levels of CO at night.

Correction applied in the revised manuscript: *"After 21h, CO levels stay quite high (240-260 ppb) due to several factors: on-going emissions (traffic and wood-burning activities), lower photochemical reactions and atmospheric dynamics (the shallower boundary layer leads to more accumulation of CO (and other co-emitted species))."*

30/ P12, L-26: Format of citations should be "Gros et al. (2011)" etc.
The format of citations was corrected in the revised manuscript.
Correction applied in the revised manuscript: *"These observations are in agreement with the conclusions from Gros et al. (2011) and Gaimoz et al. (2011)."*

31/ P12, L-32ff: The conclusions cannot be drawn from wind roses. As stated before a trajectory analysis or similar is needed for that.
Please refer to the comment n°22.

32/ P12, L-4ff: Entire section: part of the whole section better fit into the discussion.
As reported ahead, we have chosen to group results and discussions section together. Overwise, it would have led to too many repetitions. The whole section 3.3 presents and discusses seasonal and diurnal variabilities of some selected VOC with a focus on trace gases (NO, CO), aromatics and oxygenated species.

33/ P13, L-6: "Peaks" rather than increases; values high for longer time (see CO).
Correction applied in the revised manuscript: *"In winter, methanol shows a "double wave" profile with two peaks at 10-11 h and 19-20 h, suggesting the influence of anthropogenic activities (e.g. road-traffic, wood burning sources)."*

34/ P13, L-10: It is a bit hard to read because the description jumps between the normal manuscript and the supplement. It might be worth to think about including important graphs that are discussed here into the main part and leave the supplement to really supplemental information.
The diurnal cycle of methanol observed in winter and summer has been moved to Figure 7(c) in the revised manuscript.

35/ P13, L-8: "to" instead of "with".

Correction applied in the revised manuscript: *"A slight delay (1-2 hours) is observed for methanol in comparison to other primary species (for instance, aromatics), that highlights the secondary origin of this species."*

36/ P13, L-10: "…, however, …"

Correction applied in the revised manuscript: *"This night-time maximum of methanol has already been observed in urban environments, however, with no clear explanation (Solomon et al., 2015)."*

37/ P13, L-18: In the plots no clear difference can be seen between the two species.

Acetaldehyde and acetone exhibit fairly comparable seasonal and diurnal cycles, suggesting that they share the same common source(s). Mean species concentrations remain stable throughout the year, with however, higher levels in summer. They increase from sunrise to a maximum at noon and slightly decrease during the afternoon. We found that this decreasing did not occur in the same way between the two species. From 12 – 18 h, average acetaldehyde concentrations are linearly decreasing (~ 1 µg.m^{-3} or 0.5 ppb) while mean acetone levels show a slower decline rate (~0.5 µg.m^{-3} or 0.25 ppb) with a tiny raise at 17h. This difference may be explained by emission strengths, but also by the reactivity of species (acetone is less reactive than acetaldehyde).

Correction applied in the revised manuscript: *"For these two OVOC species, the reduction of concentrations does not occur in the same way. From 12 – 18 h, average acetaldehyde concentrations are linearly decreasing (~ 1 µg.m^{-3} or 0.5 ppb) while mean acetone levels show a slower decline rate (~ 0.5 µg.m^{-3} or 0.25 ppb) with a tiny raise at 17h. This finding depends on their emission sources and strengths (e.g. biogenic, solvents use), but also on their respective photochemical reaction rates (1.5 x 10^{-11} cm^{3} molecule^{-1} s^{-1} for acetaldehyde and 1.8 x 10^{-13} cm^{3} molecule^{-1} s^{-1} for acetone)."*

38/ P13, L-24: These substances are not "emissions" but secondary products of isoprene oxidation.

We agree methylvinylketone (MVK), methacrolein (MACR) and isoprene hydroxy hydroperoxides (ISOPOOH) are known to be the first-order isoprene oxidation products (Spaulding et al., 2003). Sorry for the confusion.

Spaulding, R. S. (2003). Characterization of secondary atmospheric photooxidation products: Evidence for biogenic and anthropogenic sources. Journal of Geophysical Research, 108.

Correction applied in the revised manuscript: *"The formation of these secondary compounds mostly occurs in summer, but also in winter (Fig.7f)."*

39/ P13, L-27f: Into methods section.

Introductory sentences to this section were removed in the revised manuscript as corresponding information was already introduced in the methods section.

40/ P13, L-30: "such as" instead of "like".

Correction applied in the revised manuscript: *"The temporal source variations were assessed with independent parameters used as tracers of specific sources such as inorganic gases (NO, NO_2, CO), black carbon (BC) and meteorological data (temperature)."*

41/ P14, L-7: “fair” instead of “good”

Correction applied in the revised manuscript: *“This factor 1 displays fair correlations with nitrogen monoxide/dioxide (NO/NO_2), carbon monoxide (CO) and black carbon (from its fossil fuel fraction), which are known to be relevant vehicle exhaust markers (0.53 < r < 0.64).”*

42/ P14, L-9ff: Link to respective figure should be inserted; a table or small figure of the absolute contributions could be useful.

The link to the respective figure has been inserted in the revised manuscript.

A table of the mean absolute contributions per factor and per month is reported in Section S7 in the Supplement of the new manuscript.

43/ Chapter 3.4: In general, reactivity of the different factors could be investigated in addition to stating only mass contributions; the biogenic factor and the motor vehicle factor would potentially gain importance compared to the factors with more stable gases such as background/natural gas and gas evaporation.

As advised, the reactivity of different resolved PMF sources has been investigated taking into account the factor concentration of each species with their OH rate constant (k_{OH}).

Using the definition of the OH reactivity, the total reactivity of a source $R_{total(k)}$ can be estimated as follows:

$$R_{total(k)} = \sum_{j=1}^{m} f_{kj} * k_{compound+OH}$$

Where f_{kj} is the *j*th species mass concentration from the *k*th source and $k_{compound+OH}$ represents the constant rate of a compound (*j*) with the OH radical ($k_{compound+OH}$ expressed in $cm^3.molecule^{-1}.s^{-1}$) (Atkinson et Arey, 2003).

The absolute total reactivity of the *k*th source (s^{-1}) is reported in Table 2. The reactivity of Motor Vehicle Exhaust (F1), Solvents use (F5) and Biogenic (F4) sources is among the highest ranging from 29 to 40 s^{-1}. These factors appear as the most reactive emission sources (as they are constituting of reactive species such as isoprene, aromatics, propene and ethylene) compared to factors with more stable gases (ethane, propane, butanes). As an example, the absolute reactivity of evaporative sources or natural gas and background factor is ranging from 13-18 s^{-1}.

	F1(MVE)	F2 (EVAP)	F3 (WB)	F4 (BIO)	F5 (SOLV.)	F6 (NGB)
$R_{total(k)}$ (s^{-1})	40	13	19	29	33	18

Table 2 – Absolute total reactivity of PMF factors expressed in s^{-1}.

Then, the relative contribution of reactivity R_{ik} of each factor (*k*) to the *i*th sample was calculated from the contribution g_{ik} from the *k*th source to *i*th sample and the $R_{total(k)}$ from the source *k*:

$$R_{ik} = \frac{g_{ik} \times R_{total(k)}}{\sum_{k=1}^{p} g_{ik} \times R_{total(k)}} \times 100$$

The relative contribution of reactivity of each factor is illustrated in Figure. 4 (and Fig. 12 in the revised manuscript). Solvents use and motor vehicle exhaust factors are considered here as the two main sources exhibiting a high reactivity contribution (26 % and 23 %, respectively). This can be explained by high constant rates of aromatics and alkenes which are the major components of profiles associated to these two emission sources. The contribution of the biogenic source is surprisingly weak (17 %). Although isoprene is an extremely reactive species, this factor exhibits a dominating weight of OVOCs for which constant rates can be low. Instead, the relative contribution of the mixed source "natural gas and background" is surprisingly high (16%) due to the presence of aromatics (toluene and xylenes) in the factor profile. The lower contribution of reactivity is represented by the evaporative sources factor (5%) which contains more stable gases (propane, butanes).

Atkinson, A. and Arey, J.: Atmospheric Degradation of Volatile Organic Compounds, Chem. Rev., 103(12), 4605-4638, doi:10.1021/cr0206420, 2003.

This approach is not being given in a specific section, but is included in the few words in the Sub-section 3.5 (VOC source contribution).

In the revised manuscript, it now reads: *"The reactivity of each modeled factor has also been investigated by considering the factor concentration of each species with their OH rate constant (kOH) (Atkinson and Arey, 2003) and is reported in relative (and absolute) contributions in Fig. 12. Among all the emission sources identified by PMF, solvents use and motor vehicle exhaust factors appear as the main reactive sources [26 % (33 s^{-1}) and 23 % (40 s^{-1}), respectively]. This can be explained by high constant rates of aromatics and alkenes mainly associated to these two emission sources. The contribution of the biogenic source is surprisingly weak (17 %). Although isoprene is an extremely reactive species, this factor exhibits a high weight of OVOCs for which constant rates can be low. Instead, the relative contribution of the mixed source "natural gas and background" is surprisingly high (16%) due to the presence of aromatics (toluene and xylenes) in the factor profile. The lower contribution of reactivity is represented by the evaporative sources factor [5 % (13 s^{-1})] which contains more stable gases (propane, butanes)."*

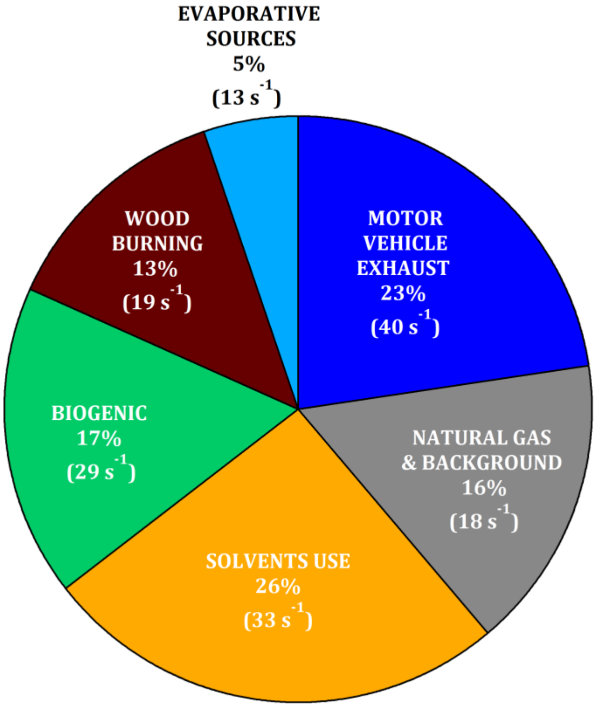


Figure 4 – Relative and absolute contributions of reactivity of each PMF factors (% and s^{-1}, respectively).

44/ P14, L-13: Why are the values so high in the middle of the night? Meteorology?

The diurnal variability of the "Motor Vehicle Exhaust" source is mainly driven by the daily cycle of species constituting the factor profile (e.g. ethylene, propene, *iso-/n*-pentane and toluene). It is characterized by a "double wave" profile: a first peak is observed between 08-10h and a second one from 18 to 21 h, usually corresponding to morning and evening rush hours during weekdays. After 21h, this factor exhibits quite high contributions (7-8 µg.m^{-3}). This observation can be explained by several factors: on-going emissions (until midnight), lower photochemical reactions and atmospheric dynamics (the lower PBL leads to more accumulation of pollutants at night). Only after 1h and until 6h, the reduction of the source contributions is significant as additional vehicular emissions are now limited.

Correction applied in the revised manuscript: *"After 21h, the absolute contributions of this factor stay quite high (7-8 µg.m^{-3}) due to several factors: on-going emissions (until midnight), lower photochemical reactions and atmospheric dynamics (the shallower boundary layer leads to more accumulation of pollutants at night)."*

45/ P14, L-19: No reference to the tunnel experiment is given here. Since it is the value added to have these kinds of data it should be described here (and discussed later).

Near-field measurements performed in a highway tunnel allowed us to characterize the fingerprint of the traffic source in Paris. This intensive experiment helps to highlight an average speciated profile representative of local primary vehicular emissions (exhausts). Photochemical reactions were considered of minor importance.

The originality of this study stands in using this near-field profile to refine the identification of a traffic source derived from modeling simulations. PMF results therefore appear more robust when comparing with independent speciated profiles.

To evaluate the relevance of the modeled traffic source, a comparison between speciated profiles from the tunnel experiment and the PMF analysis was performed and is reported in Figure 5 (in the Figure 9 in the revised manuscript). Traffic profiles are in general coherent and consistent amongst themselves (R= 0.83; R^2 = 0.69). Indeed, a good agreement is observed between these two profiles for the major species such as *iso-/n*-pentane, toluene, ethylene, propene. Instead, significant differences in mass contributions of ethane (almost a factor of 10), acetylene (considered as a key combustion compound emitted from traffic not identified in the PMF profile), isoprene (represented by evaporative sources) and oxygenated species were found.

These differences can potentially be explained by several reasons. Firstly, the proportion of VOC emitted from traffic may be different depending upon the types of vehicles/engines/fuels (Montero et al., 2010). VOC emissions can also be dependent on the use of vehicles (age, maintenance...), driving situations and thermal conditions (hot soak...). Secondly, the vehicle fleet composition can be another reason that explains these differences. In Paris intra-muros, the share of passengers cars was estimated between 60 and 80 % of the total composition of the fleet in circulation. In addition, the proportions of two-wheelers (10-20%) and Light Duty Vehicles (5-10%) have significantly increased in the recent years (Airparif, 2013). Traffic limitations prohibit heavy goods vehicles entering in the centre of Paris. In the tunnel, heavy vehicles are allowed (5% of the fleet composition). Light Duty Vehicles account for nearly 90% of the total fleet composition. Motorized scooters represent less than 2% of the total vehicles (Ammoura et al., 2014). Finally, PMF artefacts cannot be excluded. The apportionment of ethane

concentrations/contributions into the different PMF factors can also play an important role in the differences observed (factor of 10 for this species between near-field and PMF profiles). From the interpretability and fitting scores, we suppose that the contribution of ethane in the “Motor Vehicle Exhaust” factor is overestimated (same as the “Wood Burning” factor) at the expense of the mixed “Natural Gas and Background” source (for which ethane contributions seem to be underestimated – see the comment n°58).

Correction applied in the revised manuscript: *“To evaluate the relevance of this factor, a comparison between speciated profiles from tunnel measurements (Fig. 1) and PMF simulations was done and is reported in Figure. 9. Traffic profiles are in general coherent and consistent amongst themselves, thus allowing to label this factor as a “Motor Vehicle Exhaust” source. Indeed, a good agreement is observed between these two profiles for the major species such as iso-/n-pentane, toluene, ethylene, propene. Instead, significant differences in mass contributions of ethane (almost a factor of 10), acetylene (considered as a key combustion compound emitted from traffic not identified in the PMF profile), isoprene (represented by evaporative sources) and oxygenated species were found.*

These differences can potentially be explained by several reasons. Firstly, the proportion of VOC emitted from traffic may be different depending upon the types of vehicles/engines/fuels (Montero et al., 2010). VOC emissions can also be dependent on the use of vehicles (age, maintenance…), driving situations and thermal conditions (hot soak…). Secondly, the vehicle fleet composition is different in the centre of Paris and in a highway tunnel. Although the proportion of passengers cars and Light Duty Vehicles (LDV) accounts for 60-90% of the total composition of the fleet in circulation in both cases, the share of two-wheelers and heavy goods vehicles can be different. Indeed, heavy vehicles are subject to traffic limitations prohibiting their entry in Paris whereas as they are allowed in the highway tunnel (5%). The proportion of two-wheelers is significant in Paris (10-20%) (Airparif, 2013) while they represent less than 2% of the total vehicles in the tunnel. Finally, PMF artefacts cannot be excluded. We suppose that the contribution of ethane in the “Motor Vehicle Exhaust” factor is overestimated (same as the “Wood Burning” factor) at the expense of the mixed “Natural Gas and Background” source (for which ethane contributions seem to be underestimated).”

Airparif: Evolution de la qualité de l’air à Paris entre 2002 et 2012, Juillet 2013.

Montero, L., Duane, M., Manfredi, U., Astorga, C., Martini, G., Carriero, M., Krasenbrink, A. and Larsen, B.-R. : Hydrocarbon emission fingerprints from contemporary vehicle/engine technologies with conventional and new fuels, Atmos. Environ., 44, 2167-2175, 2010.

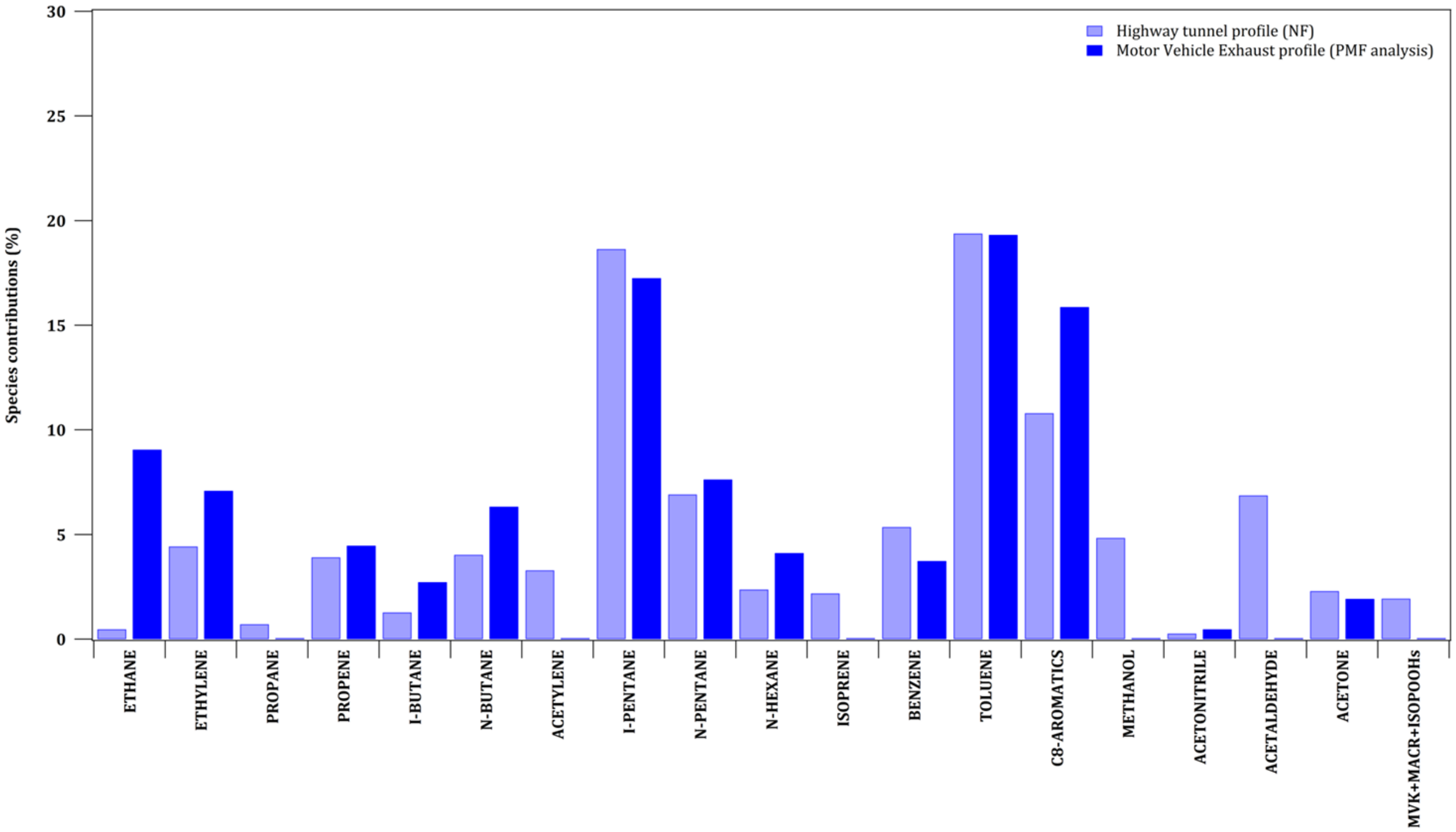


Figure 5 –Comparison of speciated profiles issued from the highway tunnel experiment and PMF simulations (F1 – Motor Vehicle Exhaust). The species contributions are expressed in %. NF = Near-Field

46/ P14, L-26: If part of the isoprene is from vehicles, it should be show up in F1, since it would be a combustion product instead of an evaporative source.
Borbon et al. (2001) have shown that isoprene could be emitted by traffic emissions. In the present study, it also shows up in the tunnel profile (Fig. 2 in the initial manuscript) in which vehicle exhausts and gasoline evaporation are represented. Isoprene emissions usually depend on temperature and could then be associated with evaporative sources.

Note that the label of this source (previously named "Gasoline Evaporation") has been changed by a more generic term "Evaporative sources" as it includes gasoline evaporation from vehicles, storage, extraction and distribution of gasoline or Liquid Petroleum Gas (LPG).

47/ P14, L-31ff: Link to figure 9 should be given.
The link to the respective figure has been inserted in the revised manuscript.

48/ P15, L-1: Correct citation: "Frachon (2009: pers. Communication)".
As suggested above, the format of this citation was corrected in the revised manuscript.

49/ P15, L-30: A lot of acetonitrile was apportioned also to F4, what is that?
The biogenic source is essentially associated with a large contribution of isoprene and oxygenated species (such as methanol or acetone). Among OVOC, acetonitrile also appears in this source. It could be attributed to a mixing with other temperature-related sources or an artefact from the PMF model (Leuchner et al., 2015).

Correction applied in the revised manuscript: *"Amounts of light alkanes (butanes, iso-pentane, n-hexane) and acetonitrile were found in this profile and could be attributed to a mixing with other temperature-related sources or artefacts from the PMF model (Leuchner et al., 2015)".*

50/ P16, L-6: "on" instead of "with". Chapter 3.4.3: reference to experiment needs to be included; a lot of ethane and propane is allocated to this factor that should very likely not be in there.
Correction applied in the revised manuscript: *"Wood burning emissions linked to home/building heating are obviously highly dependent on meteorological conditions and particularly on cold temperatures".*

Near-field measurements performed at a fireplace allowed us to have an idea of the wood-burning fingerprint in Paris. As reported in the main text (see Section 2.4.2), this speciated profile is based on a limited number of data. This confers it a more qualitative (predominant species identification) than quantitative added-value. It can only be compared to see if main compounds are the same (as conducted here), but this limitation should be kept in mind.

Usually, a low contribution of ethane and propane in the "Residential heating/wood-burning" profile is observed from near-field measurements (Barrefors and Petersson, 1995; Wang et al., 2014) and in source apportionment (SA) analyses based on urban ambient air measurements (Lanz et al., 2007). However, it may be that higher contributions of these two species can be found (Badol et al., 2008, Waked et al., in preparation).

As reported for the "Motor Vehicle Exhaust" factor (see the comment n°45), we suspect there are some artefacts within modeled PMF results. As an example, the contributions of ethane and propane in the "Wood Burning" factor seem to be significantly overestimated (compared to the fireplace profile) at the expense of the mixed "Natural Gas and Background" factor (for which ethane contributions seem to be underestimated – see the comment n°58).

Badol, C, Locoge, N., Léonardis, T. and Galloo, J.-C.: Using a source-receptor approach to characterize NMHC behaviour in a French urban area influenced by industrial emissions. Part I: Study area description, data set acquisition and qualitative data analysis of the data set, Sci. Tot. Environ., 389(2-3), 441-452, doi: 10.1016/j.scitotenv.2007.09.003, 2008.

Waked, A., Sauvage, S., Gros, V., Baudic, A., Sanchez, O., Rio, C. and Locoge, N. : Source apportionment of NMHCs in French urban sites (Paris and Strasbourg) for a period of 8 years (2005-2013), Atmos. Environ., in preparation.

Wang, H., Lou, S., Huang, C., Qiao, L., Tang, X., Chen, C., Zheng, L., Wang, Q., Zhou, M., Lu, S. and Yu. S.: Source Profiles of Volatile Organic Compounds from Biomass Burning in Yangtze River Delta, China, Aerosol Air Qual. Res., 14, 818-828, doi: 10.4209/aaqr.2013.05.0174, 2014.

Correction applied in the revised manuscript: *"It also includes ethylene (57.4%), benzene (22.7%) and oxygenated compounds, such as acetonitrile, acetaldehyde and methanol (with 18.3%, 12.6% and 6.6%, respectively). Acetonitrile is a hydrocarbon commonly used as a marker of biomass burning (Holzinger et al., 1999). All these chemical species typically reflect an anthropogenic source related to wood combustion processes (Lanz et al., 2008; Leuchner et al., 2015) in agreement with the fireplace emission profile (see Sub-section 2.4.2, Fig. 3). No full comparison between both speciation profiles was possible as the fireplace profile was based on a limited number of data. With this in mind, only a qualitative approach allowed to identify predominant species emitted from this source and confirm the term "wood burning" assigned to this factor."*

51/ P16, L-16: “includes isoprene’s”

Correction applied in the revised manuscript: *“In addition, this factor profile includes isoprene’s oxidation products [...] with more than 48 %, methanol and acetone – a selection of compounds having a large contribution from biogenic emissions (Kesselmeier and Staudt, 1998; Guenther, 2002).”*

52/ P16, L-24: r-value of the PAR correlation missing.

During the MEGAPOLI-FRANCIPOL field campaigns, Photosynthetically Active Radiation (PAR) was not measured. That’s the reason why r-value of the PAR correlation is missing. Usually, solar radiations and temperature are key factors governing biogenic emissions (Steiner and Goldstein, 2007). To avoid confusion, the corresponding sentence has been rephrased.

Correction applied in the revised manuscript: *“Biogenic emissions are directly related to solar radiations (Steiner and Goldstein, 2007) and ambient temperature (r > 0.7).”*

53/ P16, L-27: very high values at night, why? Stability of the atmosphere?

The biogenic factor is mainly driven by isoprene and oxygenated species (MVK+MCR+ISOPOOH; methanol, acetone). Biogenic emissions are usually dependent on solar radiations and temperatures (Steiner and Goldstein, 2007) – gradually increasing during the day. High nighttime contributions of the biogenic source can be explained by the presence of oxygenated species (long-lived compounds already present in the atmosphere and/or secondarily formed from the oxidation of isoprene) in the profile combined with lower photochemical reactions and atmospheric dynamics (a low PBL height) at night.

Correction applied in the revised manuscript: *“Biogenic emissions are directly related to solar radiations (Steiner and Goldstein, 2007) and ambient temperature (r > 0.7). For that reason, the highest factor contributions occur in summer (10 µg.m^{-3} on average) with a contribution up to 14.3 µg.m^{-3} in July. Daily mean contributions gradually increase from 09h. A slight delay (03h) is observed in comparison with diurnal temperature/solar radiations variations (for which values increase from sunrise at 06 h). We assume that chemistry affects this source factor as it takes part in the formation of secondary species (MAK+MVK+ISOPOOH, for instance) from the oxidation of primarily emitted compounds (isoprene, OVOC). Diurnal contributions reach their maximum at the end of the day (~19h). Highest nighttime contributions of this source can be explained by the presence of oxygenated species (long-lived compounds already present in the atmosphere and/or secondarily formed from the oxidation of isoprene) in the profile combined with lower photochemical reactions and atmospheric dynamics (a low PBL height) at night.”*

54/ P17, L-14: what does “panel 5” mean?

Panel 5 represents “Monthly box and whiskers plots of the solvents use source. To avoid confusion, numbers from 1 to 5 were graphically attributed to each source. Consequently, Panel 1 = Motor Vehicle Exhaust; Panel 2 = Evaporative sources; Panel 3 = Wood Burning, Panel 4 = Biogenic; Panel 5 = Solvents use; Panel 6 = Natural Gas & Background.

55/ P17, L-14: the PBL is higher in the afternoon, that contradicts the statement of the diurnal pattern. Emissions might be higher and also a temperature dependency.

The seasonal and diurnal variabilities of the “Solvents use” source are presented in Figure 10, Panel 5 (in the initial version of the manuscript). We assume that the daily variation of

this factor is dependent on the season considered. Diel variations of the factor is compared here between winter (top right) and spring/summer/autumn (bottom right).

In winter, factor contributions increase after 6 h and reach their maximum between 11 h and 19 h (15-20 µg.m^{-3}) before a long and gradual decline in the evening. Higher contributions in winter can be explained by lower photochemical reactions and atmospheric dynamics. A shallower PBL (and consequently, less intense vertical dynamics) leads to more accumulation of pollutants.

In spring/summer/autumn, factor contributions also increase at sunrise, but reach their maximum between 08 and 10 h (typical of anthropogenic activities). They progressively decrease during the afternoon. This gradual decline (not earlier observed in winter) is influenced by greater photochemical reactions and more intense vertical dynamics during these three seasons, leading to dispersion and dilution processes (and consequently, lower source contributions during the afternoon).

Correction applied in the revised manuscript: *"The highest contribution of this factor is observed during winter (14.2 µg.m^{-3}) with a contribution of up to 20.9 µg.m^{-3} in January. In winter, factor contributions increase at 6h and reach their maximum between 11h and 19h (15-20 µg.m^{-3}) before a long and gradual decline in the evening (see Fig.10, panel 5 – top right). Higher contributions in winter can be explained by lower photochemical reactions (combined with weaker OH concentrations/UV radiations) and atmospheric dynamics. Indeed, a shallower PBL (and consequently, less intense vertical dynamics) leads to more accumulation of pollutants and thus to higher source contributions. The daily wintertime variability of this source is in agreement with the diel cycle of independent tracers (ethanol, butan-2-one).*

Reconstructed contributions associated to this factor are also significant in summer (12.6 µg.m^{-3} in July), which could be mainly explained by the evaporation of solvent inks, paints and other applications during that month due to higher temperatures. In spring/summer/autumn, factor contributions also increase at sunrise, but reach their maximum between 08 and 10 h (typical of anthropogenic activities). They progressively decrease during the afternoon (see Fig.10, panel 5 – bottom right). This gradual decline (not earlier observed in winter) is influenced by greater photochemical reactions and more intense vertical dynamics during these three seasons, leading to dispersion and dilution processes (and consequently, lower source contributions during the afternoon)."

56/ P17, L-18: not shown.
The diurnal cycles of the "Solvents use" factor observed in winter and spring/summer/autumn are presented in Figure 10, panel 5 in the revised manuscript.

57/ P17, L-22: "except"
The word "expected" was substituted by "expect".

58/ P17, L-31: BLH not defined. Better to use only one abbreviation, e.g. PBL Chapter 3.4.6: the annual course of the background does not fit with similar observations; there should be a maximum in February and minimum in late summer due to the OH reactivity (delayed by 2 months to the annual solar cycle because of the long-lifetime of ethane).
Only the abbreviation PBL (Planetary Boundary Layer) has been used throughout the revised manuscript.

The factor mixing "Natural Gas & Background" (NGB) sources is mainly driven by ethane (45% of species sum). The annual course of this PMF factor does not fit with that usually observed in the literature whereas the diurnal cycle does (see Fig.10 – panel 6). As reported in previous comments n°45 and 50, PMF artefacts cannot be ruled out. Indeed, a problem with the distribution of ethane within PMF factors was raised. We assume that higher ethane contributions were assigned to the "Motor Vehicle Exhaust" and "Wood Burning" factors. Consequently, this allocation issue has effects on the NGB source (as ethane is a key component in the factor profile). For that reason, we suspect that the NGB factor contributions were underestimated (especially in winter) for the benefits of the "Wood Burning" factor (another source significantly contributing during this season).

Correction applied in the revised manuscript: *"The average contribution of this mixed source (combining both natural gas and background emissions) is in the range of 9.2 µg.m^{-3} during the whole studied period. Lowest source contributions were observed in winter, which does not fit with that reported in the literature. As mentioned in the "Motor vehicle exhaust" and "Wood burning" sub-sections (3.4.1 and 3.4.3, respectively), PMF artefacts cannot be ruled out. Indeed, a problem with the distribution of ethane (considered as the key species of the mixed source) within PMF factors was raised. We assumed that higher ethane contributions were partly assigned to the "motor vehicle exhaust" and "wood burning" factors. Consequently, we assumed that the "natural gas and background" factor contributions were underestimated (especially in winter) for the benefits of the "wood burning" factor (another source significantly contributing during this season).*

59/ P18, L-7ff: cannot be concluded only from wind roses. Statements need to be supported that it really results from different air masses, since annual cycle does not fit.
Please refer to previous comments n°22 and 31.

60/ Chapter 3.5: pure discussion, also better fits into discussion section.
Please refer to the previous comment n°32.

61/ P18, L-21: background not really fresh emission, but aged air mass
To avoid confusion, the corresponding sentence has been rephrased.
Correction applied in the revised manuscript: *"This Source Apportionment (SA) analysis concluded that the predominant sources at the receptor site were road-traffic-related activities (including motor vehicle exhaust, 15 % of the Total VOC (TVOC) mass on the annual average, and evaporative sources, 10 %), with the remaining sources from natural gas and background (23 %), solvents use (20 %), wood burning (17 %) and biogenic activities."*

62/ P18, L-21ff: additional information on the contribution of reactivity of each factor would be very useful.
Please refer to the comment n°43.

63/ P19, L-4: most of the conclusions are quite speculative and need to be supported by more profound analysis.
Please refer to previous comments n°22, 31 and 59.

64/ P20, L-9: this discussion would fit to the discussion of factors.
Please refer to the previous comment n°32.

65/ P20, L-13: “combination”

Correction applied in the revised manuscript: *“To accurately compare VOC sources proportions between 2007 and 2010 (for a similar combination of hydrocarbons and masses), the contribution of each main factor was recalculated for the specific time period May-June 2010 (Fig.13, right pie chart).”*

66/ P20, L-33: there are some other longer time series, but the dataset indeed is very useful.

67/ P21, L-23: absolute values could also be interesting in addition for comparison to other studies

Please refer to the previous comment n°42.

68/ P22, L-12: why were exactly these studies chosen? There are many other studies described in the literature from e.g., Houston, Santiago de Chile, etc.

This Source Apportionment (SA) study was initially compared with some SA studies performed in Europe and in the World in order to evaluate the consistency and the representativeness of PMF results at a large scale. The selection of SA studies is nowhere near complete. SA studies presented in the manuscript were chosen on the basis of one or more criteria: (i) VOC (NMHC and/or OVOC) measurements should be undertaken in urban background sites (to omit any local influences of emission sources), (ii) SA should be performed from long-term VOC time series, (iii) similar source categories (factor identification) should be reported.

Following reviewer’ suggestions, we decided to keep the comparison between our results with SA studies conducted in Europe, but to shorten the comparison with other cities in the world. Nevertheless, we now mention more references of global SA studies without commenting in detail their results.

Correction applied in the revised manuscript: *“The consistency in VOC contributions in European urban areas raises the question of their representativity at a larger scale. There are currently many other urban SA studies described in the literature (e.g. Jorquera and Rappenglück, 2004 – Santiago/Chile; Buzcu et al., 2006 – Houston/TX; Brown et al., 2007 – LA; Cai et al., 2010 - Shanghai/China; Morino et al., 2011 – Tokyo/Japan; Yurdakul et al., 2013 – Ankara/Turkey; Zheng et al., 2013 – Mexico). Results of these studies are not detailed here but one common feature for European and global scales is the importance of the road-traffic source (between 30 % and 50%). One difference concerns the industrial sector which plays (in the investigated European cities) a lower role than in studied urban areas from other continents.*

Governmental regulations and standards to control pollutants emissions and economic developments may differ between European countries and the world. The location of sampling points (or distances from main sources) and meteorological conditions can strongly affect VOC concentrations and their respective emission sources in the considered urban environments.”

Jorquera, H. and Rappenglück, B.: Receptor modeling of ambient VOC at Santiago, Chile, Atmos. Environ., 38, 4243-4263, doi:10.1016/j.atmosenv.2004.04.03, 2004.

69/ Graphics: Some of the axis titles and labels (e.g., Figs. 8, 9 but also all others) are quite small and might be hard to read when formatted to the actual page size.

The font size of the axis titles and labels for all figures was increased in the revised manuscript for an easier reading.

70/ Fig. 5 could go to supplement or omitted if replaced with a more profound air mass investigation.

The performed analysis of the wind fields was replaced with a more elaborate air mass investigation presented in Figure 5. Wind roses plots were moved to the supplement (Section S5).

71/ Fig. 8: Logarithmic scale makes it hard to see the "real" profile for absolute contributions.

Usually, absolute contributions of each species apportioned in the factor (resolved by PMF) are plotted on a logarithmic scale (see Norris et al., 2014). For a better appreciation of PMF outputs, we opted for a linear scale. We also omitted the Total Variable called "VOC".

Norris, G., Duvall, R., Brown, S. and Bai, S. (2014). EPA Positive Matrix Factorization (PMF) 5.0: Fundamentals & User Guide. Prepared for the US. Environmental Protection Agency (EPA), Washington, DC, by the National Exposure Research Laboratory, Research Triangle Park; Sonoma Technology, Inc., Petaluma.

72/ P47, L-12: The factors could be shown in supplement, would be interesting to see.

PMF modeling simulations derived from the FRANCIPOL dataset (March – November 2010) were performed in order to support to the assumptions related to the accuracy and robustness of final PMF results. These PMF outputs were incorporated into the Section S1 in the Supplement.

73/ P47, L-6ff: quite important information. Could be moved to methods section of main text.

The dataset composition was already mentioned in the methods section of the main text. In the A1 section, only a recap is now provided.

74/ Appendix A: here the calculation of uncertainties could be shown.

As reported in the previous comment n°14, a comprehensive calculation of uncertainties could not be performed on all data and therefore, we chose to not dedicate an additional appendix to this issue.

Acp-2016-185: "Seasonal variability and source apportionment of volatile organic compounds (VOC) in the Paris megacity (France)" -- Baudic et al.

Authors' Responses to Referee #2

We would like to thank the Referee #2 for her/his general feedback and each of her/his useful comments/questions for improving the quality of this manuscript. All of them have been taken into account when preparing the revised version of the manuscript. In the present document, authors' answers to the specific comments addressed by Referee #2 are mentioned in **blue**, while changes made to the revised manuscript are shown in *italic.*

General/scientific comments

This paper reports the results of a source apportionment by positive matrix factorization (PMF) of the concentrations of a suite of ambient VOCs measured in urban background air in Paris over a period of several months from January to November 2010. VOCs were measured both by on-line GC and by PTR-MS. In order to help in the assignment of some of the factors emanating from the application of PMF, the authors compared the speciated VOC profiles of the factors with speciated VOC profiles the authors separately measured at three locations where they assume that a single emission source will dominate the ambient VOC, specifically: (1) measurements during busy (and traffic jam) periods in a highway tunnel, to represent vehicle related VOC emissions; (2) measurements close to a domestic gas flue, to represent natural gas source; and (3) measurements at a fireplace facility, to represent residential wood-burning emissions (the authors acknowledge that their measured VOC profile from this source may be less quantitative than for other source profiles). To further assist in the assignment of PMF factors to particular VOC emission sources the authors also make use of additional co-located atmospheric compositional data available to them, such as NO, CO and black carbon (BC).

The authors present results in which the ambient urban background VOC has been apportioned into six sources, each of which has been assigned an identification, albeit that one factor is assigned to be a mixed natural gas/background source. The analysis presented includes speciated VOC profiles for each factor and monthly and average-hourly variation in the absolute contribution (ug/m3) of each of the six identified factors/sources. The largest contributions in total are from traffic-related activities (through two identified factors: motor-vehicle exhaust, and gasoline evaporation), although all six identified sources have not that dissimilar relative contributions, on average. A noteworthy observation is significant contributions to ambient VOC from wood burning, 18% on average but up to ~50% at times in winter. Biogenic emissions were also reported to be significant, 15% on average but more in summer.

The authors have a large dataset of time-resolved speciated VOC over an extended time period, almost a year. They have used standard, but appropriate, statistical methods to endeavour to decompose the ambient measurements into individual source contributions. These are statistical, rather than dispersion-/chemistry-based source apportionments. These methods have been widely employed to apportion ambient PM, but less so for VOC.

The presentation of results and their discussion are largely descriptive, in the sense that the authors present the details for their factor/source contributions and their monthly and hourly variations, which are rationalised with general text about anticipated behaviour of particular sets of pollutant mixtures in the urban background atmosphere. The authors also present a qualitative and quantitative comparison of their VOC source apportionment with previous literature.

Although in one sense the presentation of this work could be described as 'formulaic' – following previous data presentation and analysis styles – nonetheless the large dataset presented here for a large European city presents a valuable addition to the VOC apportionment literature. A particular feature is the presentation of VOC speciated profiles for three potential VOC sources, although the authors acknowledge some shortcomings in these. The paper is already lengthy and contains much new data, supported by detailed descriptions of the data collection and processing protocols and results. As an overall summary, the content of the paper is suitable for publication in ACP.

The authors thank Reviewer#2 very much for his/her attention to our manuscript. All comments addressed by both reviewers have been taken into account in the preparation of the revised version of the manuscript. In this respect, several figures and tables were notably added and modified (e.g. air masses trajectory analysis – Fig. 5; comparison between "highway tunnel" and "motor vehicle exhaust" profiles – Fig. 9; relative and absolute contributions of reactivity of each PMF factor – Fig. 12) and in the supplementary (e.g. PMF profiles issued from simulations using the FRANCIPOL dataset & comparison of speciated PMF profiles issued from our two different datasets – Section S1; representativeness of meteorological parameters in 2010 – S4; mean absolute contributions of factors per month in 2010). Please note that figures and tables numeration is now different in this new version.

Technical comments

1/ The tables and figures are generally clearly presented. The written text is largely unambiguous in conveying its meaning, but it is overly long in places. There are instances where introductory sentences to a section could be substantially abbreviated, or even deleted as repeating what the reader will have picked up from the methods section. The authors should be encouraged to edit text further for conciseness of expression.
Following reviewer's suggestions, the authors made efforts to write a revised version of the manuscript with conciseness. Some rephrasing also aimed at bringing the content a bit more to the point.

The following are more specific comments.

2/ P6, L-18: should read "Raw data were corrected using…."?
The initial sentence "Raw data were compensated using…" was substituted by the following one: *"Raw data were corrected using the algorithm described in Weingartner et al. (2003) and Sciare et al. (2011)"*.

3/ P6, L-25: delete "is" before "analyzer"
Correction applied in the revised manuscript: *"Nitrogen monoxide and dioxide (NO, NO_2) were measured by chemiluminescence using an AC31M analyzer (Environment SA, Poissy, France) and ozone (O_3) was monitored with an automatic UltraViolet absorption analyzer (41M, Environment SA, Poissy, France)."*

4/ P6, L-32: delete "at"
Correction applied in the revised manuscript: *"Standard meteorological parameters […] were provided by the French national meteorological service "Météo-France" from continuous measurements recorded at the Paris-Montsouris monitoring station (14th district, 48°49'N, 02°20'E), located about 2 km away from the LHVP site."*

5/ P7, L-22: should read "except for"
Correction applied in the revised manuscript: *"This combination of hydrocarbon species and masses is similar to that from Gaimoz et al. (2011), except for iso-butene."*

6/ P8, L-9: I don't understand what is being described in the sentence beginning "Finally, these different processings…." Please rephrase this sentence to make clear to what "it" is referring in this sentence and to clarify what the procedure undertaken was.
In this study, only 59% of isoprene data are modeled by the PMF technique, what is not objectively considered as sufficient. To address this problem, input values of isoprene were modified when detecting missing data. We have investigated the possibility of replacing missing values with data more appropriate than the median. Indeed, a "virtual" averaged pattern was calculated from 1-h real samples observed in June and August (to keep the summer variability of isoprene). Another option was to preserve raw data of isoprene (no subsequent changes) and to increase the analytical uncertainty initially estimated at 15% (gradually from 15 % to 30%) – instead of categorizing isoprene as *weak* (as uncertainties are tripled implying a lower Signal to Noise). This was intended to better display this biogenic compound. PMF simulations were performed by considering these different options. As a consequence of these tests, no significant improvement on the quality of modeling isoprene was observed. Regular statistical parameters such as R^2, slope, slope/intercept SE (Standard Error) were used to draw such conclusions. As empirical tests have not helped, isoprene is still categorizing as *strong*.

In the previous version of the manuscript, "processings" referred to "empirical tests" and "it" to isoprene. To avoid any ambiguity in the meaning of this sentence, corresponding lines were rephrased.

Correction applied in the revised manuscript: *"To address this lack of isoprene data, several empirical tests (e.g. simulating an averaged seasonal/diurnal cycle of isoprene or increasing the analytical uncertainty of raw data from 15 % up to 30 %) were conducted within PMF simulations with the aim of better modeling the variability of this compound. As a consequence of these tests, no significant improvement on the quality of modeling isoprene was observed. Finally, isoprene is still categorizing as strong here".*

7/ P10, L-26: the phrase "fairly comparable" is not scientifically precise.
The corresponding sentence was rephrased.
More specific statements about weather conditions are given here.
Correction applied in the revised manuscript: *"Air temperatures observed during the campaign were comparable to standard values determined by the French national meteorological service "Météo-France" (available at: http://meteofrance.com), with however an uncommon cold wintertime (Bressi et al., 2013- Fig. S1a). Temperatures recorded in January and February 2010 were respectively between -2°C and -3.5°C below normal values (see. Section S2 in the Supplement). Extreme unusual cold-air outbreaks and a few snow flurries affected the Paris region, thus explaining higher temperature anomalies during that period. Levels of hours of sunshine and rainfall were globally consistent with standard values, with however some discrepancies in winter/autumn and spring, respectively (Fig. S3)."*

8/ P10, L-29: "a few flurries" of what? Please write specific statements about the nature of the weather.
Please refer to the previous comment n°7.

9/ P11, L-15: the phrase "contribute to the tune of" is too colloquial; please use more direct wording.
Correction applied in the revised manuscript: *"Both alkanes and OVOCs significantly contribute up to 75 % of the TVOC concentrations."*

10/ P11, L30-L35: there are several instances in these lines of text of negative values for VOC concentrations. These are surely some aberration (albeit repeated aberration) of typing error. Please correct.
A comparison between mean concentrations of aromatics and OVOCs measured in this study and ambient air levels reported in the literature for different urban atmospheres (see Table 2) was made. It aims at highlighting existing differences in VOC levels monitored for a given time period between our work (Paris) and another European or global studies (the selection is not exhaustive here). We note that VOC concentrations measured in European cities (Paris, Barcelona, and London) are in the same order of magnitude depending on the compounds. Instead, more significant differences in VOC levels were found between Paris and Houston, Beijing, Mohali or Mexico City.

The negative values were due to differences between VOC concentrations from our study compared to others. But to avoid confusion, differences are now only given with absolute values.

Correction applied in the revised manuscript: *"Average VOC concentrations were also calculated in line with sampling periods of the other European and global studies over different years (see Table 2). In this study, measured VOC levels were in the range of those found with some European cities (Barcelona, London – from 0.1 to 2.1 ppb concentration differences). However, average VOC levels observed in Paris were significantly lower than those measured in Houston (USA – from 0.1 to 6.9 ppb concentration differences) and more particularly in Beijing (China - from 2.5 to 8.9 ppb), in Mexico City (Mexico – from 0.1 to 27.4 ppb) and in Mohali (India – from 0.9 to 32.7 ppb)".*

It is important to note that Table 2 was modified and all reported values are given in ppb (see the comment n° 13 – Referee 1).

11/ P13, L-3: I do not understand the scientific sense of the sentence starting "With an atmospheric residence time…" How does the statement at the end of this sentence (about methanol emissions contributing to background levels) derive from, or otherwise relate to, the text at the start of the sentence about methanol residence time? Please reword to clarify.
We agree with this comment. There is no direct link between the beginning and the end of this sentence. This additional information did not bring an added value to the scientific approach. To avoid confusion, we opted to omit the start of this sentence.

Correction applied in the revised manuscript: *"Methanol is usually released into the atmosphere by vegetation and man-made activities contributing to a relatively high background levels during most of the year."*

12/ P14, L-24: I do not follow the scientific logic here. The text appears to state that iso-pentane is known to be a key tracer for gasoline evaporation, but also to say that iso-pentane was not present in the speciated profile the authors have assigned in their work to gasoline evaporation.
In this study, the "evaporative sources" factor is mainly characterized by the presence of propane and butanes (*iso-/n-*). The species composition of the F2 profile is consistent

with those obtained from other SA studies (Brown et al., 2007; Gaimoz et al., 2011; Waked et al., 2016). In addition to C_3-C_4 species, pentanes (*iso-/n-*) and toluene are also considered as key tracers for gasoline evaporation (Salameh et al., 2016 – in preparation). A small contribution of these species was identified in F2. However, their highest contributions were assigned to the "Motor Vehicle Exhaust" factor" (F1) in agreement with that observed during the tunnel experiment.
To avoid any misunderstanding, the corresponding sentence was removed.

13/ P14, L-32: please explain more clearly what about the monthly change "remains ambiguous".
The seasonal cycle of "Evaporative sources" profile is expected to be significant in spring and summer (when higher temperatures are observed). The minimum mean contribution of this source is observed in July when road-traffic emissions are usually significant (See Figure 10 – Panel 1).
This explanation is given with more details in the main text (P14, L-31 – P15, L-5).

14/ P20, L19: The phrasing that the mean temperature was "in the range +/- 20 degC" does not make sense. Either quote the range, or quote the mean and some recognized statistic of the variation about the mean. Likewise for later in this sentence in connection with "+/- 16 degC".
Correction applied in the revised manuscript: *"The temperatures recorded in May-June 2007 and 2010 were 20°C and 16°C, respectively".*

15/ Caption of Table 1: It would be helpful for the caption to remind the reader with a statement of the time resolution of the raw data from which these statistical summaries are derived, and of the time duration/dates of the total dataset.
Caption of Table 1:
Statistical summaries (µg.m^{-3}) of selected VOC concentrations measured at urban background sites. Statistics were calculated based on hourly mean data, initially obtained every 30 min. (ethane > isoprene) and every 5 to 10 min. (for aromatics and OVOCs). These measurements were undertaken from 15 January to 22 November 2010 (~ 10 months. A conversion factor is provided here to convert VOC concentrations (µg.m^{-3}) into (ppb) mixing ratios.

"Seasonal variability and source apportionment of volatile organic compounds (VOC) in the Paris megacity (France)" by A. Baudic et al.
acp-2016-185

Please find below a list of all relevant changes applied in the revised version of the manuscript:

ABOUT THE TEXT :

Technical corrections :

i) Corrections to typing and grammar errors
ii) Rephrasing and rewriting of the content when it was necessary (see reviewers' comments).
iii) Information about the concentration units conversion (µg.m^{-3} > ppb using a conversion factor – Table 1)
iv) The 24-h format time used
v) New key references included – Updated references
vi) Sections and sub-sections numeration modified

Fundamental corrections :

Sampling sites :

- More details about the location of sampling stations (Sub-section 2.1)
- Representativeness of measurement sites (characterization of the sites typology using NO/NO_2 ratio - Sub-section 2.2)

Analytical section :

- More details about uncertainties on VOC data provided by GC-FID and PTR-MS
- Information about the stability of analytical systems

Positive Matrix Factorization (PMF) simulations :

- PMF simulations only performed from the FRANCIPOL dataset (source composition profiles presented in the Section 1.1 in the Supplement)
- Comparison of speciated PMF profiles issued both from FRANCIPOL and MEGAPOLI/FRANCIPOL data to prove that it was valid to consider these two datasets in this Source Apportionment (SA) study (Section 1.2 in the Supplement)

More details about **weather conditions** in Paris in 2010 (Section 3.1)

Analysis of the air-mass history and backtrajectories clustering using the HYSPLIT model (see **Fig.5**)

VOC variabilities and emission sources :

- Information/Explanation about higher concentrations of CO and some OVOCs
- The generic term ‘Evaporative sources’ was used and replaced the ‘Gasoline Evaporation’ source label
- New table of absolute contributions per PMF factor and per month (Section S7 in the Supplement)
- Information/Explanation about variabilities (daily and/or seasonal) of the ‘Motor Vehicle Exhaust’, ‘Solvents use’, ‘Biogenic’ and ‘Natural Gas & Background’ PMF factors
- Comparison of speciated profiles issued from PMF simulations and near-field measurements (i.e. Motor Vehicle Exhaust [**Fig. 9**] and Wood Burning)
- Absolute and relative contributions of reactivity of PMF factors (Section 3.5 – **Fig. 12**)
- Focus on the comparison between PMF results with SA studies performed in Europe. A few words about global SA studies.

ABOUT FIGURES :

- Figures numeration changed
- Wintertime and summertime diurnal variation of methanol included
- Font size of the axis titles and labels modified

Seasonal variability and source apportionment of volatile organic compounds (VOCs) in the Paris megacity (France)

A. Baudic[1], V. Gros[1], S. Sauvage[2], N. Locoge[2], O. Sanchez[3], R. Sarda-Estève[1], C. Kalogridis[1,*], J.-E. Petit[1,4,**], N. Bonnaire[1], D. Baisnée[1], O. Favez[4], A. Albinet[4], J. Sciare[1,***], and B. Bonsang[1]

[1]LSCE, Laboratoire des Sciences du Climat et de l'Environnement, Unité Mixte CEA-CNRS-UVSQ, CEA/Orme des Merisiers, 91191 Gif-sur-Yvette, France
[2]Mines Douai, Département Sciences de l'Atmosphère et Génie de l'Environnement (SAGE), 59508 Douai, France
[3]AIRPARIF, Association de Surveillance de la Qualité de l'Air en Île-de-France, 75004 Paris, France
[4]INERIS, Institut National de l'EnviRonnement Industriel et des risqueS, DRC/CARA/CIME, Parc Technologique Alata, BP2, 60550 Verneuil-en-Halatte, France
[*]*Now at: Institute of Nuclear Technology and Radiation Protection, Environmental Radioactivity Laboratory, National Center of Scientific Research 'Demokritos', 15310 Ag. Paraskevi, Attiki, Greece*
[**]*Now at: Air Lorraine, 20 rue Pierre Simon de Laplace, 57070 Metz, France*
[***]*Now at: Energy Environment Water Research Center (EEWRC), The Cyprus Institute, Nicosia, Cyprus*

Correspondence to: V. Gros (valerie.gros@lsce.ipsl.fr)

Abstract. Within the framework of air quality studies at the megacity scale, highly time-resolved volatile organic compounds (C_2 - C_8) measurements were performed in downtown Paris (urban background sites) from January to November 2010. This unique dataset included non-methane hydrocarbons (NMHCs) and aromatic/oxygenated species (OVOCs) measured by a GC-FID (Gas Chromatograph with a Flame Ionization Detector) and a PTR-MS (Proton Transfer Reaction - Mass Spectrometer), respectively. This study presents the seasonal variability of atmospheric VOCs being monitored in the French megacity and their various associated emission sources. Clear seasonal and diurnal patterns differed from one VOC to another as the result of their different origins and the influence of environmental parameters (solar radiation, temperature). Source Apportionment (SA) was comprehensively conducted using a multivariate mathematical receptor modeling. The United States Environmental Protection Agency's Positive Matrix Factorization tool (US EPA, PMF) was used to apportion and quantify ambient VOC concentrations into six different sources. The modeled source profiles were identified from near-field observations (measurements from three distinct emission sources: inside a highway tunnel, at a fireplace and from a domestic gas flue, with hence a specific focus on road-traffic, wood burning activities and natural gas emissions) and hydrocarbon profiles reported in the literature. The reconstructed VOC sources were cross-validated using independent tracers such as inorganic gases (NO, NO_2, CO), black carbon (BC) and meteorological data (temperature). The largest contributors to the predicted VOC concentrations were traffic-related activities (including motor vehicle exhaust, 15 % of the total mass on the annual average, and evaporative sources, 10 %), with the remaining emissions from natural gas and background (23 %), solvents use (20 %), wood burning (18 %) and a biogenic source (15 %). An important finding of this work is the significant contribution from wood burning, especially in winter, where it could represent up to ~ 50 % of the total mass of VOCs. Biogenic emissions also surprisingly contributed up to ~ 30 % in summer (due to the dominating weight of OVOCs in this source). Finally, the mixed natural gas and background source exhibited a high contribution in spring (35 %, when continental air influences were observed) and in autumn (23 %, for

home heating consumption).

Keywords : Urban atmospheric pollution, Volatile organic compounds (VOCs), Ozone precursors, Vehicular emissions, Positive Matrix Factorization (PMF), Source apportionment.

1 Introduction

More than half of the world's population is now living in urban areas and about 70 % will be city-dwellers by 2050 (United Nations, 2014). Many of these urban centers are ever-expanding, leading to the gradual growth of megacities. Strong demographic and economic pressures are exerting increasing stress on the natural environment, with impacts at local, regional and global scales. Megacities are hotspots of atmospheric gaseous and particulate pollutants, which are subjects of concern for sanitary, scientific, economic, societal and political reasons. The adverse health effects of outdoor air pollutants are recognized today. Indeed, ambient air pollution has been classified as *carcinogenic to humans* by the International Agency for Research on Cancer since October 2013 (IARC, 2013). In recent decades, air pollution has become one of the most widespread problems in many megacities and should be more investigated.

The understanding of the pollutants in urban areas remains complex given the diversity of their emission sources (unequally distributed in space and time) as well as their formation and transformation processes. Volatile organic compounds (VOCs) are of a great scientific interest because they play an important role in atmospheric chemistry. In the troposphere, primary VOCs take part in chemical and/or photochemical reactions, thus contributing to the formation of ground-level ozone (O_3) (Logan et al., 1981; Liu et al., 1987; Chameides et al., 1992; Carter et al., 1994) and secondary organic aerosols (SOA) (Tsigaridis and Kanakidou, 2003 and references therein; Ng et al., 2007). While some megacities face very poor air quality (such as Beijing, Gurjar (2014)) with pollutant concentrations way above recommended thresholds, European megacities experience stagnant pollution levels at the annual scale. However, pollution episodes related to high O_3 and PM concentrations still regularly occur, leading to detrimental health consequences.

Epidemiological studies revealed that outdoor air pollution, mostly from $PM_{2.5}$ and O_3, could lead to 17 800 premature deaths in France (with 3 100 for the Paris megacity) in 2010 and projections for the future are even worse (3 800 in 2025 and 4 600 in 2050 for Paris) (Lelieved et al., 2015). Paris and its surroundings (also called the *Île-de-France* region) constitute the second largest European megacity with about 12 million inhabitants, representing 20 % of the French national population distributed over only 2 % of its territory (Eurostat, 2014). Although this region is surrounded by a rural belt, it is considered as a large central urban area where a strong pollution signal can be detected. Deguillaume et al. (2008) have shown that the urban area of Paris was frequently associated with a "VOC-sensitive" chemical regime (also called "NO_x-saturated" regime according to Sillman, 1999), for which VOC anthropogenic emission reductions are more effective to decrease ozone levels than NO_x anthropogenic emission reductions. Obtaining accurate knowledge on VOC emissions and sources is consequently essential for O_3 and SOA abatement measures.

Qualitative and quantitative assessments of VOC variability and sources have already been conducted within the Paris area during May-June 2007 (Gros et al., 2011; Gaimoz et al., 2011). This study concluded that road-traffic activities (traffic exhaust and fuel evaporation) influenced the total VOC fingerprint, with a contribution of ~ 39 %. This finding was considered as being in disagreement with the local emission inventory provided by the air quality monitoring network AIRPARIF (http://www.airparif.asso.fr), for which the main contribution was related to solvents usage (from industries and from residential sectors). However, this work was performed over a short period of time (only few weeks). Although it provided valuable information about ambient VOC emissions and sources during a specific period (spring), it did not show their seasonal variations over longer time scales. More resolved observations are therefore required to check the representativity of these first conclusions.

In this context, the EU-F7 MEGAPOLI (Megacities: Emissions, urban, regional and Global Atmospheric POLlution and climate effects, and Integrated tools for assessment and mitigation) (Butler, 2008) and the French PRIMEQUAL-FRANCIPOL research programs involving several (inter)-national partners in the atmospheric chemistry community have been implemented. These MEGAPOLI - FRANCIPOL projects partly consisted in documenting a large number of gaseous and particulate compounds and determining their concentration levels, variabilities, emission sources and geographical origins (local or imported) within the Paris urban area. These experiments go therefore beyond the scope of this paper and a full description of scientific studies conducted under the programme can be found in the special issue "MEGAPOLI - Paris 2009/2010 campaign" available in the Atmospheric Chemistry and Physics (ACP) journal (e.g. Crippa et al., 2013; Skyllakou et al., 2014; Ait-Helal et al., 2014; Beekmann et al., 2015 and references therein).

Here, this work presents near real-time measurements of VOCs performed at urban background sites in downtown Paris from 15 January to 22 November 2010. Its objectives are to (1) assess ambient levels of a VOC selection, (2) describe their temporal (seasonal and diurnal) variabilities, (3) identify their main emission sources from statistical modeling and (4) quantify and discuss their source contributions on yearly and seasonal bases.

In order to identify and apportion ambient VOC levels to their emission sources, the advanced multivariate receptor modeling technique Positive Matrix Factorization (PMF) was applied. As no prior knowledge of the number or the chemical nature of source profiles is explicitely required (Paatero and Tapper, 1994), the identification of PMF source profile outputs must be made a posteriori. It usually relies on speciation profiles available in the literature. Within this study, near-field additional measurements (at source points: inside a highway tunnel, at a fireplace and from a domestic gas flue) were performed to help strengthen this identification of VOC profiles derived from PMF simulations. This experimental approach is dedicated to provide a specific fingerprint of VOC sources related to road-traffic, residential wood burning activities and domestic natural gas consumption, respectively. The originality of this work stands in using these near-field speciation profiles to refine the identification of apportioned sources.

First, Section 2 will describe (i) sampling sites, (ii) analytical techniques conducted and (iii) two combined approaches for identifying and characterizing the main VOC emission sources. Then, Section 3 will investigate VOC levels and their seasonal and diurnal patterns from long-term ambient air measurements. An accurate identification of PMF factors to real physical sources will be proposed in the Sub-section 3.4. Finally, yearly and seasonal contributions of each modeled source will be discussed in the Sub-section 3.5 and compared to previous studies performed in Paris and widely in Europe in Sub-sections 3.6 and 3.7.

2 Material & Methods

2.1 Sampling sites description

As part of the European EU-F7 MEGAPOLI (Megacities: Emissions, urban, regional and Global Atmospheric POLlution and climate effects, and Integrated tools for assessment and mitigation (http://www.megapoli.info, 2007-2011)) program, a winter campaign involved measurements of a large amount of atmospheric compounds – with techniques including GC-FID and PTR-MS for VOCs – from 15 January to 16 February 2010 at the "Laboratoire d'Hygiène de la Ville de Paris" (LHVP) (Baklanov et al., 2010; Beekmann et al., 2015). Located in the southern part of Paris centre (13th district – 48°82′ N, 02°35′ E – 15 m above ground level, a.g.l.), LHVP dominates a large public garden (called "Parc de Choisy") at approximately 400 m from Place d'Italie (grouping a shopping centre and main boulevards).

A second measurement campaign involving less instrumentation (only PTR-MS for VOCs) was conducted at the LHVP site from 24 March to 22 November 2010 (as part of the French PRIMEQUAL-FRANCIPOL program: Impact of long-range transport on particles and their gaseous precursors in Paris and its region (http://www.primequal.fr, 2010-2013)). At the same time, hydrocarbon measurements by GC-FID were carried out by the regional air quality monitoring network AIRPARIF at the "Les Halles" subway station (48°51′ N, 02°20′ E – 2.7 m a.g.l.) located 2 km away from LHVP. The location of these two sampling sites is presented in Figure 1.

2.2 Representativeness of sampling sites

Due to the low intensity of the surrounding activities, the LHVP and Les Halles sampling sites were considered as urban background stations by AIRPARIF and by previous scientific studies (Favez et al., 2007; Sciare et al., 2010; Gros et al., 2011). In accordance with the 2008/50/EC European Directive, this station typology is based on two main criteria : (1) the population density is at least 4 000 inhabitants per square kilometre within a 1-km radius of the station and (2) no major traffic road is located within 300 metres.

This characterization of site typologies can be confirmed by studying the nitrogen monoxide (NO): nitrogen dioxide (NO_2) ratio. NO is known to be a vehicle pollution indicator whereas NO_2 has an important secondary fraction. To consider a station as an urban background site, the ratio R between annual average NO and NO_2 concentrations (NO/NO_2) should be less than 1.5 ppb:ppb, as indicated in a report (Mathé, 2010) written at the national level for regulatory purposes.

We observed a very close NO: NO_2 ratio (expressed as ppb: ppb) between both sites in 2010: 0.40 (LHVP) *versus* 0.38 (Les Halles). The same conclusion can be made for the years 2008 and 2009 (0.45 ± 0.01 for LHVP and 0.48 ± 0.01 ppb: ppb for Les Halles). With NO: NO_2 ratios very similar and less than 1.5, this confirms that these two locations have the same site typology and can be considered as urban background stations.

To prove that it was valid to merge these two specific datasets from different locations, we opted for a comparison of PMF results derived from MEGAPOLI-FRANCIPOL (as performed in the present paper) and FRANCIPOL data files, respectively. PMF modeling simulations were performed using only the FRANCIPOL dataset. A good agreement was found for the majority of the emission sources. The graphs showing the comparison of PMF results between MEGAPOLI-FRANCIPOL and FRANCIPOL are reported in the Supplementary Section S1.

2.3 Experimental set-up

2.3.1 VOC measurements using a Proton Transfer Reaction-Mass Spectrometer (PTR-MS)

Within the MEGAPOLI and FRANCIPOL projects, online high-sensitivity Proton Transfer Reaction-Mass Spectrometers PTR-MS (Ionicon Analytik GmbH, Innsbruck, Austria) were used for real-time (O)VOC measurements. As this instrument has widely been described in recent reviews (Blake et al., 2009 and references therein), only a description of analytical conditions relating to ambient air observations is given here.

During these two intensive field experiments, a PTR-MS was installed in a small room located on the roof of the LHVP site ($15\,\mathrm{m}$ a.g.l.). For the MEGAPOLI winter campaign, PTR-MS measurements performed by the Laboratoire de Chimie Physique (LCP, Marseille, France) have already been described in Dolgorouky et al. (2012). As those performed by the Laboratoire des Sciences du Climat et de l'Environnement (LSCE) during FRANCIPOL have not yet been described elsewhere, more technical details are presented below.

Air samples were drawn up through a Teflon line (0.125-cm inner diameter) fitting into a “DEKABON” tube in order to protect it from light. A Teflon particles filter (0.45-µm pore diameter) was settled at the inlet to avoid aerosols and other fragments from entering the system. The PTR-MS was operating at standard conditions: a drift tube held at 2.2 mbar pressure, 60°C temperature with a drift field of 600 V voltage to maintain an *E/N* ratio of ~ 130 Townsend (Td) [*E*: electrical field strength $[\mathrm{V\,cm^{-1}}]$; *N*: buffer gas number density $[\mathrm{molecule\,cm^{-3}}]$; 1 Td = $10^{-17}\ \mathrm{V/cm^2}$]. First water cluster ions $\mathrm{H_3O^+H_2O}$ (at *m/z* 37.0) and $\mathrm{H_3O^+H_2O\ H_2O}$ (*m/z* 55.0) were also measured as well as $\mathrm{NO^+}$ and $\mathrm{O_2^+}$ masses to indicate any leak into the system and assess the PTR-MS performances.

(O)VOC measurements performed in a full-scan mode were enabled to browse a large range of masses (*m/z* 30.0 – *m/z* 150.0). Eight protonated target masses were considered here: methanol (*m/z* = 33.0), acetonitrile (*m/z* = 42.0), acetaldehyde (*m/z* = 45.0), acetone (*m/z* = 59.0), methylvinylketone (MVK) + methacrolein (MACR) + isoprene hydroxy hydroperoxides (ISOPOOHs) (*m/z* = 71.0), benzene (*m/z* = 79.0), toluene (*m/z* = 93.0) and xylenes (*m*+*p*-, *o*-) + $\mathrm{C_8}$ aromatics (*m/z* = 107.0). With a dwell time of five seconds per mass, a mass spectrum was obtained every two to ten minutes for MEGAPOLI and FRANCIPOL campaigns, respectively. Around 80 % of PTR-MS data were validated. Missing data were partly due to background measurement periods and calibrations. The PTR-MS background for each mass was monitored by sampling zero air through a catalytic converter heated to 250°C to remove chemical species. Daily background values were averaged and subtracted from ambient air measurements.

In order to regularly ensure the analytical stability of the instrument, injections from a standard containing benzene (5.7 ppbv [parts per billion by volume] ± 10 %) and toluene (4.1 ppbv ± 10 %) were performed approximately once a month from March to November 2010. These measurements have shown that the analyzer stability remained stable during the year with variations within ± 10 %. In addition, two full calibrations were performed before and during the intensive field campaign with a Gas Calibration Unit (GCU, Ionicon Analytik GmbH, Innsbruck, Austria). The standard gas mixture provided by Ionicon contained 17 VOCs at 1 ppmv [parts per million by volume]. These calibration procedures consisted of injecting defined concentrations (in the range from 0 to 10 ppbv) of different chemicals (previously diluted with synthetic air) with a relative humidity at 50 %. Gas calibrations allowed to determine the repeatability of measurements, expressed here as a mean coefficient of variation. This coefficient was less than 5 % for most of the masses. Slightly higher coefficients were observed for *m/z* 69.0 (isoprene) and *m/z* 71.0 (crotonaldehyde) with 5.6 % and 5.2 %, respectively. Observed differences between both full calibration procedures were from 1.1 % for methanol to 9.8 % for toluene, thus illustrating good analyzer stability over time.

Detection limits (LoD) were calculated as three times the standard deviation of the normalized background counts when measuring from the catalytically converted zero air. For the MEGAPOLI winter campaign, LoD ranged from 0.020 to 0.317 µg m^{-3} (0.007 – 0.238 ppb) whereas they were estimated between 0.020 to 0.330 µg m^{-3} (0.007 – 0.248 ppb) during the FRANCIPOL intensive campaign. The analytical uncertainty on all data was estimated by taking into account errors on standard gas, calibrations, blanks, reproducibility/repeatability, linearity and relative humidity parameters. The measurement uncertainty was estimated at ± 20 % in agreement with previous studies (Gros et al., 2011; Dolgorouky et al., 2012).

2.3.2 NMHC on-line measurements by Gas Chromatography (GC)

Two different automated Gas Chromatographs equipped with a Flame Ionization Detector (GC-FID) were used in order to continuously measure light (C_2 - C_6) VOCs in ambient air. The AirmoVOC C_2 - C_6 analyzer (Chromatotec, Saint-Antoine, France), provided by LSCE, was installed near the PTR-MS at the LHVP site from January to February 2010 (e.g. MEGAPOLI period). An in-depth description of the analyzer, sampling set up and technical information (sampling flows, preconcentration, desorption-heating times, types of traps and columns . . .) can be found in Gros et al. (2011). For each half-hour analysis, more than 20 VOCs were monitored. A certified standard gas mixture (NPL, National Physical Laboratory, Teddington, Middlesex, UK) containing in average 4.00 ppbv of major C_2 - C_9 NMHCs was used for calibration procedures. The injections of this standard allowed checking compound retention times, testing the repeatability of atmospheric measurements and calculating average response factors to calibrate all the measured ambient hydrocarbons. Detection limits were in the range of 0.013 (*n*-hexane) – 0.060 µg m^{-3} (*iso-/n*-pentanes) (0.004 – 0.020 ppb). The analytical uncertainties on LSCE measurements were estimated from laboratory tests (i.e. memory effects, repeatability, accuracy of the gas standard) to be ± 15 % (Gros et al., 2011).

From March to November 2010 (e.g. FRANCIPOL period), a GC-FID coupled to a thermo-desorption unit was in operation at the "Les Halles" subway station monitored by the regional air quality network AIRPARIF. Air samples were drawn up at $2.7\,\mathrm{m}$ a.g.l.. 29 hydrocarbons from $\mathrm{C_2}$ - $\mathrm{C_9}$ were measured during this experiment. A full calibration was performed once a month with a standard gas mixture containing only propane during 6 hours. As the FID response is proportional to the Effective Carbon Number (ECN) in the molecule, calibration coefficients were calculated for each compound and regularly checked so that they drifted no more than $\pm$ 5 % (tolerance threshold). In addition, a zeroing was carried out every 6 months using a zero air bottle in order to detect any instability or problem with the GC-FID system. LoD were assessed at $0.024\,\mu\mathrm{g\,m^{-3}}$ for all the selected compounds, except for *n*-hexane ($0.013\,\mu\mathrm{g\,m^{-3}}$) (0.022 – 0.004 ppb). A comprehensive calculation of uncertainties could not be performed due a lack of sufficient information from AIRPARIF. Nevertheless, from previous experimental tests, NMHC measurements were provided with an accuracy of $\pm$ 15 %, which happened to be in agreement with the uncertainty value for the LSCE GC-FID.

2.3.3 Additional data available

Some ancillary pollutants and parameters were also measured and used as independent tracers with the aim of strengthening the identification of VOC emission sources derived from the receptor modeling.

Black carbon (BC) was measured using a 7-wavelength (370, 470, 520, 590, 660, 880 and 950 nm) AE31 Aethalometer (Magee Scientific Corporation, Berkeley, CA, USA) with a time resolution of five minutes. BC data were acquired by this instrument from 15 January to 10 September 2010. Raw data were corrected using the algorithm described in Weingartner et al. (2003) and Sciare et al. (2011). BC concentrations issued from fossil fuel and wood burning (BC*ff* and BC*wb*, respectively) were assessed in accordance with their own absorption coefficients using the "Aethalometer" model described by Sandradewi et al. (2008).

Carbon monoxide (CO) measurements were performed using an analyzer based on InfraRed absorption (42i-TL instrument, Thermo Fisher Scientific, Franklin, MA, USA) with a time resolution of five minutes. Nitrogen monoxide and dioxide (NO, $\mathrm{NO_2}$) were measured by chemiluminescence using an AC31M analyzer (Environment SA, Poissy, France) and ozone ($\mathrm{O_3}$) was monitored with an automatic UltraViolet absorption analyzer (41M, Environment SA, Poissy, France). NO, $\mathrm{NO_2}$ and $\mathrm{O_3}$ measurements were provided with a 1-min time resolution by the local air quality network AIRPARIF. In addition, Gas Chromatography-Mass Spectrometry (GC-MS) measurements were performed to measure $\mathrm{C_3}$ - $\mathrm{C_7}$ OVOCs, including aldehydes, ketones, alcohols, ethers and esters during the MEGAPOLI winter campaign. This instrument has been described in details by Roukos et al. (2009). As the measurement frequency was different for each analyzer, a common average time was defined to get all data sets on a similar time step of one hour.

Standard meteorological parameters (such as temperature, relative humidity as well as wind speed and direction) were provided by the French national meteorological service "Météo-France" from continuous measurements recorded at the Paris-Montsouris monitoring station (14th district – $48°49'$ N, $02°20'$ E), located about $2\,\mathrm{km}$ away from the LHVP site.

In order to determine the air masses origin, 5-day back trajectories were calculated every 3 hours from the PC based version of the HYbrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model (Stein et al., 2015) with Global Data Assimilation System (GDAS) meteorological field data. Back trajectories were set to end at Paris coordinates (49°02′ N, 02°53′ E) at 500 m a.g.l..

2.4 Two combined approaches for characterizing VOC emission sources

2.4.1 Bilinear receptor modeling: Positive Matrix Factorization (PMF) tool

Developed about 20 years ago, Positive Matrix Factorization (PMF) is an advanced multivariate factor analysis tool widely used to identify and quantify the main sources of atmospheric pollutants. Concerning VOCs, PMF studies have been conducted in urban (e.g. Brown et al., 2007 – LA, USA; Lanz et al., 2008 – Zürich, Switzerland; Morino et al., 2011 – Tokyo, Japan; Yurdakul et al., 2013 – Ankara, Turkey) and rural areas (e.g. Sauvage et al, 2009 – France; Lanz et al., 2009 – Jungfraujoch, Switzerland; Leuchner et al., 2015 – Hohenpeissenberg, Germany). For this current study, the PMF 5.0 software developed by the EPA (Environmental Protection Agency) was used in the robust mode from ambient air VOC measurements from January to November 2010. A more detailed description of this PMF analysis is given in Appendix A.

PMF mathematical theory was extensively described in Paatero and Tapper (1994). Briefly, this statistical method consists in decomposing an initial chemically speciated dataset into factor profiles and contributions. Equation (1) summarizes this principle in its matrix form:

$$X = G\,F + E \qquad (1)$$

Where X is the input chemical dataset matrix; G is the source contribution matrix; F is the source profiles matrix and E the so-called residual matrix.

The initial chemical database used for this statistical study contains a selection of 19 hydrocarbon species and masses divided into ten compound families: alkanes (ethane, propane, *iso*-butane, *n*-butane, *iso*-pentane, *n*-pentane and *n*-hexane), alkenes (ethylene and propene), alkyne (acetylene), diene (isoprene) aromatics (benzene, toluene, xylenes + C_8 species), alcohol (methanol), nitrile (acetonitrile), aldehyde (acetaldehyde), ketone (acetone) and enones (methylvinylketone, methacrolein and isoprene hydroxy hydroperoxides), which have been measured from 15 January to 22 November 2010 ($n = 6\,445$ with a 1h-time resolution). This combination of hydrocarbon species and masses is similar to that from Gaimoz et al. (2011) except for *iso*-butene. Each missing data point was substituted with the median concentration of the corresponding species over all the measurements and associated with an uncertainty of four times the species-specific median, as suggested in Norris et al. (2014). The proportion of missing values ranges from 19 % (especially for compounds measured by PTR-MS) to 41 % (only for isoprene). This high percentage for isoprene can be mainly explained by analytical problems on GC-FID in July. Despite this limitation, it was decided to take into account this compound as isoprene is a key tracer related to biogenic emissions.

The uncertainty matrix was built upon the procedure described by Norris et al. (2014), adapted from Polissar et al. (1998). This matrix requires the Method Detection Limit (MDL, here in µg m^{-3}) and the analytical uncertainty (u, here in %) for each selected species. MDLs were calculated as 3σ baseline noise and in some cases were homogenized to keep consistency in uncertainty calculations. Species MDLs were ranged from 0.013 to 0.060 µg m^{-3} (0.004 – 0.020 ppb) for NMHCs measured by GC-FID and from 0.020 to 0.330 µg m^{-3} (0.007 – 0.248 ppb) for VOCs measured by PTR-MS. Their analytical uncertainties were respectively estimated at 15 % and 20 % and kept constant over the experiments.

Single species additiona l uncertainties were also calculated using an equation-based on the Signal-to-Noise (S/N) ratio. As a first approach, Paatero and Hopke (2003) suggested categorizing a species as *Bad* if the Signal to Noise Ratio (SNR) was less than 0.2, *Weak* if it was between 0.2 and 2, and *Strong* if it was greater than 2. *Bad* variables are excluded from the dataset, *Weak* variables get their uncertainties tripled while uncertainties of *Strong* variables stay unchanged. Here, all species exhibited a SNR greater than 3, except for isoprene which had a ratio of 1.7 due to its 41 % missing values recorded. To address this lack of isoprene data, several empirical tests (e.g. simulating an averaged seasonal/diurnal cycle of isoprene or increasing the analytical uncertainty of raw data from 15 % up to 30 %) were conducted within PMF simulations with the aim of better modeling the variabilty of this compound. As a consequence of these tests, no significant improvement on the quality of modeling isoprene was observed. Finally, isoprene is still categorizing as *strong* here. In addition to isoprene, acetonitrile exhibited a SNR less than 3 (S/N ratio of 2.7). It was the only VOC to be defined as *weak* because it may be eventually contaminated with local emissions from laboratory exhausts (although visible spikes of acetonitrile were excluded from the initial dataset). Keeping in mind these limitations for isoprene and acetonitrile, it was decided to include these two compounds into the PMF model because they are considered as relevant tracers for biogenic and wood burning activities, respectively. ΣVOC was defined as *Total Variable* and automatically categorized as *weak* to lower its influence in the final PMF results. No optional Extra Modeling Uncertainty was applied here.

All the Q values ($Qtrue$, $Qrobust$, $Qexpected$), scaled residuals, predicted *versus* observed concentrations interpretation and the physical meaning of factor profiles were investigated to determine the optimum number of factors. Although some mathematical indicators pointed towards a 5-factor solution, a source mixing solvents use activities, natural gas and background emissions was detected. In order to split individually each emission source, the 6-factor solution was then investigated and chosen in terms of interpretability and fitting scores. More technical details are reported in Appendix A, Sub-section A2.

PMF output uncertainties were estimated using two Error Estimation methods starting with DISP (dQ-controlled displacement of factor elements) and finally processing to BS (Classical bootstrap). The DISP analysis results were considered validated: no error could be detected and no drop of Q was observed. As no swap occurred, the 6-factor PMF solution was considered sufficiently robust to be used. Bootstrapping was then carried out, executing 100 iterations, using a random seed, a block size of 874 samples and a minimum Pearson correlation coefficient (R-Value) of 0.6. All the modeled factors were well reproduced through this bootstrap technique over at least 88 ± 2 % of runs, hence highlighting their robustness. A low rotational ambiguity of the reconstructed factors was found by testing different degrees of rotations of the solution using the Fpeak parameter (Fpeak = 0.1).

2.4.2 Determination of source profiles from near-field observations

As emphasized in Paatero and Tapper (1994), a Positive Matrix Factorization analysis does not require a priori knowledge on the chemical nature of factor profiles. To help strengthen the identification of VOC emission sources derived from this statistical tool, near-field additional measurements (at limited source points) were worthwhile. These in-situ observations were performed close to specific local emission sources in real conditions as far as possible. They aim at providing chemical fingerprints (considered as reference speciation profiles) of three significant VOC sources representative within the Paris area: road traffic (i), residential wood burning (ii) and domestic natural gas consumption (iii). The speciated profiles of these different anthropogenic sources and their representativity are given here. All the technical details of these experiments are reported in Section S2 in the Supplement.

Highway tunnel experiment

Road-traffic was considered to be one of the most significant sources of primary hydrocarbons in many megacities (Seoul – Na, 2006; Los Angeles – Brown et al., 2007; Zürich – Lanz et al, 2008), including Paris (Gros et al., 2007, 2011; Gaimoz et al., 2011). The accurate characterization of the vehicle fleet footprint is therefore important. Consequently, near-field VOC measurements (within the "PRIMEQUAL – PREQUALIF" program) were conducted inside a highway tunnel located about 20 km southeast of inner Paris centre in autumn 2012.

This first experiment has the advantage of supplying a realistic assessment of the average chemical composition of vehicular emissions, as these in-situ measurements were performed under on-road real driving conditions. Most of VOCs emitted from road-traffic are representative of local primary emissions (due to their relatively short lifetimes). Photochemical reactions leading to changes in the initial composition of the air and to the formation of secondary products can be considered of minor importance.

VOC levels during traffic jam periods (07-09 h – 17:30-19 h, Local Time) were considered as the most representative values of vehicular emissions. In order to omit any local background, nighttime values (as suggested in Ammoura et al., 2014) were subtracted from the peak VOC concentrations. The mass contribution of 19 selected compounds was calculated and reported in Figure 2. Each compound is expressed in terms of weight out of the weight (w/w) of the Total VOC (TVOC) mass. The two predominant species measured inside the highway tunnel were toluene (19.4 %, 27.1 µg m^{-3} (7.1 ppb) on average) and *iso*-pentane (18.6 %; 26.0 µg m^{-3} (8.7 ppb)). The next most abundant VOCs were aromatics (benzene, C_8) and oxygenated compounds (acetaldehyde and methanol – 5.0 - 10 %; 5-9 µg m^{-3} (1.1 - 4.6 ppb)). In addition, significant contributions of light alkenes (ethylene, propene, acetylene; 3.3 % – 4.5 %; 4.6 – 6.2 µg m^{-3}) (4.2 – 5.4 ppb) and alkanes (such as butanes, *n*-pentane, *n*-hexane, methyl alkanes) were also noticed. These observations were found to be consistent with the literature (Na, 2006; Araizaga et al., 2013 concerning NMHCs and Legreid et al., 2007 for OVOCs measurements) and more importantly with the study from Touaty and Bonsang (2000), for which *iso*-pentane, ethene, acetylene, propene and *n*-butane were considered as the major aliphatic compounds observed in the same highway tunnel in August 1996 (Aromatics and OVOCs were not measured during this study).

Fireplace experiment

Residential wood burning activities have been shown to be a significant source of (O)VOCs to local indoor and outdoor air pollution during winter months in urban areas (Evtyugina et al., 2014). Currently, only a few studies about the characterization of VOCs from wood burning have been conducted in Europe (Gustafson et al., 2007; Gaeggeler et al., 2008). After being subject to lively debates within the Île-de-France region, an in-depth investigation of this source would therefore appear necessary to better understand its emission specificities and its potential impacts on atmospheric chemistry.

In order to complete the information on wood burning activities, VOC measurements (within the "CORTEA-CHAMPROBOIS" program) were performed at a fireplace facility located ~ 70 km northeast of the inner Paris centre in March 2013. On-line (PTR-MS) and off-line (sampling flasks analyzed later on at the laboratory with a GC-FID) measurements were performed. These in-situ observations represent a more qualitative (predominant species identification) than quantitative approach as the resulting speciation profile is based on a limited number of data. As illustrated in Figure 3, 19 VOC species could be detected. C_2 hydrocarbons (ethylene, acetylene), C_6 - C_8 aromatics (benzene, xylenes) and oxygenated species (methanol, acetaldehyde and acetone) can be considered as predominant compounds from domestic wood burning. This finding is still consistent with intensive field studies of wood burning performed in Europe (Barrefors and Petersson, 1995; Gaeggeler et al., 2008; Evtyugina et al., 2014; Nalin et al., 2016).

Natural gas experiment

Natural gas is predominately composed of methane (CH_4) accounting for at least 80 % of the total chemical composition. It is also a mixture of other pollutants including lightweight VOCs and lower paraffins (approximately 10 % by volume).

In a first approach to determine the speciation profile from natural gas used in Paris, near-field samplings were performed from a domestic gas flue using three stainless-steel flasks, which have been analyzed by GC-FID at the laboratory. Main results (Figure 4) show a large dominance of alkanes, such as ethane (~80 %), propane (~11 %) and heavier hydrocarbons (like butanes, pentanes) ranging from 4.5 to 0.4 %. Ethane and propane therefore appear as a significant profile signature of natural gas leakages (Passant, 2002).

3 Results and discussions

3.1 Meteorological conditions and air mass origins

Meteorological parameters (e.g. temperature, relative humidity, rainfall, sun exposure, boundary-layer height, wind speed and direction) are known to be key factors governing seasonal and diurnal variations of air pollutant levels.

Air temperatures observed during the campaign were comparable to standard values determined by the French national meteorological service "Météo-France" (available at: http://meteofrance.com), with however an uncommon cold wintertime

(Bressi et al., 2013 - Fig.S1a). Temperatures recorded in January and February 2010 were respectively between -2°C and -3.5°C below normal values (see. Section S3 in the Supplement). Extreme unusual cold-air outbreaks and a few snow flurries affected the Paris region, thus explaining higher temperature anomalies during that period. Levels of hours of sunshine and rainfall were globally consistent with standard values, with however some discrepancies in winter/autumn and spring, respectively (Fig. S3). In addition, atmospheric boundary layers showed seasonal changes with mean heights up to ~ 800 m in winter and up to 1 600 m in spring and summer occurring during the afternoon (see Section S4 in the Supplement). These average seasonal heights are expected to play a key role in pollutant dispersion and consequently impact ambient VOC concentrations.

For the year 2010, Paris was mainly influenced by air masses coming from the west (62 %) and usually associated with clean marine air influences from the Atlantic Ocean (see Figure 5). They are typically representative of local and regional pollution conditions, as already observed in Gros et al., 2011, Gaimoz et al., 2011; Dolgorouky et al., 2012 and Petit et al., 2015. To a lesser extent, Paris can be affected by northeast air masses (26 %) originating from Eastern France, the Benelux area, Northern Germany and indicative of continental imports of long-lived pollutants (Gaimoz et al., 2011). Air masses coming from the west are generally observed in summer and autumn (32 – 41 %) whereas northeast air masses are found to be significant in winter (34 %) and most frequently in spring (ca. 40 %) , due to the stagnation of an anticyclone surrounding the British Isles (Monthly weather report for Paris and its surroundings during April 2010, Météo-France) during that period.

3.2 VOC concentration levels in ambient air

The main results of descriptive statistics for all the measured VOCs (from both GC-FID and PTR-MS instruments) on the whole sample set were summarized in Table 1. The average composition of VOCs was mainly characterized by oxygenated species (0.7 – 5.9 µg m^{-3} (0.4 – 4.4 ppb); 36.5 % of the TVOC mass), alkanes (0.5 – 4.6 µg m^{-3} (0.1 – 3.7 ppb); 39.1 %) followed by aromatics (1.1 – 3.3 µg m^{-3}; 16.9 % (0.3 – 0.9 ppb)) and to a lesser extent by alkenes, alkynes and dienes (0.3 – 1.6 µg m^{-3} (0.1 – 1.3 ppb); 7.5 %). Both alkanes and OVOCs significantly contribute up to 75 % of the TVOC concentrations. With ethane (10.9 %, 4.6 µg m^{-3} (3.7 ppb) on average) being the main alkane, methanol (14.0 %, 5.9 µg m^{-3} (4.4 ppb)) and acetone (11.6 %, 4.9 µg m^{-3} (2.0 ppb)) are considered to be the two major oxygenated compounds measured in this study. This conclusion is in agreement with previous VOC measurements performed in downtown Paris in 2007 (Gros et al., 2011).

The comparison between these average ambient levels and VOC measurements reported in the literature for different urban areas is restricted here to PTR-MS data as they constitute the most original dataset of this study. Most atmospheric studies were indeed conducted in urban metropolitan areas by investigating only NMHC measurements.

Table 2 summarizes PTR-MS data collected during the intensive experiment together with average VOC levels reported in ppb from other cities around the world. For the Paris megacity, a significant decrease in VOC concentrations was observed between spring 2007 and spring 2010 (from - 53.8 % for xylenes and C_8 aromatics to - 25 % for benzene and MVK+MACR+ISOPOOHs), excepted for methanol and acetaldehyde (+ 11.8 % , + 35.7 %, respectively). These differences in aromatics concentration levels are consistent with decreasing downward trends of NMHC recorded during springtime in Paris (Waked et al., 2016). We note that the study of Waked et al. (2016) on VOC trends in Paris only concerns NMHC and not oxygenated species.

As these OVOCs are significantly impacted by biogenic and secondary sources, it is not surprising to observe a different variation between 2007 and 2010.

Among selected species, benzene (as a carcinogenic agent) is one of the few regulated VOCs. According to the Directive 2000/69/EC, the annual mean benzene concentration in ambient air should not exceed 5 µg m^{-3} (1.5 ppb). Background levels of benzene were relatively stable in recent years, with an annual average concentration of 1.1 µg m^{-3} (0.3 ppb) (Airparif, 2015).

Average VOC concentrations were also calculated in line with sampling periods of the other European and global studies over different years (see Table 2). In this study, measured VOC levels were in the range of those found within some European cities (Barcelona, London – from 0.1 to 2.1 ppb concentration differences). However, average VOC levels observed in Paris were significantly lower than those measured in Houston (USA – from 0.1 to 6.9 ppb concentration differences) and more particularly in Beijing (China – from 2.5 to 8.9 ppb), in Mexico City (Mexico – from 0.1 to 27.4 ppb) and in Mohali (India – from 0.9 to 32.7 ppb).

3.3 Seasonal and diurnal variations

The variability in VOC concentration levels is controlled by a combination of factors including source strengths (e.g. emissions), dispersion and dilution processes as well as photochemical reaction rates with OH radicals and other oxidants (Filella and Peñuelas, 2006). Variations of selected trace gases (nitrogen/carbon monoxide, NO/CO – Figure 6) and VOCs illustrating contrasting emission sources and atmospheric lifetimes were analyzed at different time scales. As (O)VOCs measured by PTR-MS constitute the most original data of this study (and representing ~ 37 % of the TVOC mass), a discussion on their variations (Figure 7) and their respective sources is given here. For information, an overview of seasonal and diurnal profiles of lighter hydrocarbons (C_2 - C_6) measured by GC-FID is reported in Section S6 in the Supplement.

Known as combustion tracers (traffic/wood burning), nitrogen monoxide (NO) and carbon monoxide (CO) exhibit higher median concentrations during winter and in late autumn, while lower concentrations appear in summer (Fig. 6a, 6b). These low levels can be explained by greater photochemical reaction rates (linked to higher solar radiation) combined with a stronger vertical atmospheric mixing compared to the other seasons. Another explanation of this variability is the increase in NO and CO emissions due to home heating fuels consumed in winter. NO concentrations are significantly enhanced between 06 and 12 h LT (with its maximum around 09 h). Contrary to NO, the diurnal pattern of CO is characterized by a “double wave” profile with an initial peak at 07-10 h (maximum at 09 h) and a second one at the end of the afternoon between 16 and 21 h. These increases typically correspond to morning and evening rush-hour traffic periods, as previously observed in Ammoura et al. (2014). The evening peak is smaller in magnitude than the morning one partly due to a higher Planetary Boundary-Layer (PBL) height in the afternoon (leading to dispersion and dilution processes) and to more disperse traffic periods. After 21h, CO levels stay quite high (240 – 260 ppb) due to several factors: on-going emissions (traffic and wood-burning activities), lower photochemical reactions and atmospheric dynamics (the shallower boundary layer leads to more accumulation of CO (and other co-emitted species)). The evening event is not observed for NO as during this time ozone (O_3) presents its highest concentrations, leading to the titration of NO.

Good correlations between CO and some alkanes (*iso-/n*-pentane, *n*-hexane), alkenes (ethylene, propene), acetylene and aromatics (benzene, toluene and xylenes + C_8) were found when considering a Pearson's correlation coefficient r greater than 0.6. All these compounds follow a similar seasonal and diurnal pattern, indicating that they share some or almost all common sources related to anthropogenic combustion processes (e.g. road traffic and/or wood burning). These observations are in agreement with the conclusions from Gros et al. (2011) and Gaimoz et al. (2011).

With atmospheric lifetimes from a few hours to several days, oxygenated species (OVOCs) are emitted from primary sources, mainly of biogenic origins, and significant secondary sources related to the oxidation of hydrocarbons. High concentration levels of OVOCs (for instance, methanol) and CO were observed during winter months (the season with coldest temperatures and where wood burning-related activities can play an important role). The low height of the PBL is also a relevant factor to consider as it can lead to the accumulation and the stagnation of VOC species into the troposphere during that season. In addition, significant OVOC levels were observed from April to September. In springtime, elevated baseline levels were measured when the Paris region was mostly influenced by northeast influences (see Fig.5), suggesting that they partly depended on continental imported and already processed air-masses. Biogenic emissions indeed contributed to high OVOC concentrations during this season and in summer.

Methanol is usually released into the atmosphere by vegetation and man-made activities contributing to a relatively high background levels during most of the year. This compound displays a specific diurnal pattern depending on the season and atmospheric dynamics (see Section S4 in the Supplement). In winter, methanol shows a “double wave” profile with two peaks at 10-11 h and 19-20 h (see Fig. 7c, top right), suggesting the influence of anthropogenic activities (e.g. road-traffic, wood burning sources). A slight delay (1-2 hours) is observed for methanol in comparison to other primary species (for instance, aromatics), that highlights the secondary origin of this species. In summer, methanol is characterized by high concentrations during night hours (00 - 06 h), followed by a significant decrease until the early afternoon and another increase from 18 h to midnight (Fig. 7c, bottom right). This night-time maximum of methanol has already been observed in urban environments, however, with no clear explanation (Solomon et al., 2005). This diurnal cycle can possibly be interpreted as the accumulation of species concentrations during the night from a local source under a shallow inversion layer, which is decreasing when the PBL is increasing (as dilution and dispersion processes occurring). However, the corresponding nighttime source has not been yet identified.

With a relatively short lifetime (~ 9 hours), acetaldehyde shows a diurnal cycle fairly comparable to acetone (Fig. 7d, 7e). Lower concentrations were observed during the night and from 18 h. Average levels increase from sunrise to a maximum at noon and slightly decrease in the afternoon. For these two OVOC species, the reduction of concentrations does not occur in the same way. From 12 – 18 h, average acetaldehyde concentrations are linearly decreasing (~1.0 µg m^{-3} or 0.5 ppb) while mean acetone levels show a slower decline rate (~0.5 µg m^{-3} or 0.25 ppb) with a tiny raise at 17h. This finding depends on their emission sources and strengths (e.g. biogenic, solvents use), but also on their respective photochemical reaction rates (1.5

x 10^{-11} cm^3 $\mathrm{molecule}^{-1}$ s^{-1} for acetaldehyde and 1.8 x 10^{-13} cm^3 $\mathrm{molecule}^{-1}$ s^{-1} for acetone) [Atkinson et al., 2006]. As acetone has a relatively long atmospheric lifetime (~ 68 hours), concentration levels are often more homogeneous.

Finally, methylvinylketone, methacrolein and isoprene hydroxy hydroperoxides (MVK+MACR+ISOPOOHs), three secondary products of isoprene photo-oxidation (as good indicators of biogenic activities), exhibit high levels in the late afternoon due to the oxidation of daytime isoprene. The formation of these compounds mostly occurs in summer, but also in winter (Fig. 7f). This fact could eventually be related to anthropogenic activities such as wood burning (see Sub-section 2.4.2, Fig. 3).

3.4 Source apportionment

3.4.1 Motor Vehicle Exhaust factor

The speciation profile of Factor 1 (see Fig. 8a) exhibits high contributions of alkanes, such as pentanes (*iso*-; *n*-) and *n*-hexane with on average ~ 50 % of their variabilities explained by this factor. Aromatic compounds (toluene, xylenes + C_8, benzene; ~ 35 %) and light alkenes (ethylene, propene), which are considered as typical combustion products, are also the predominant species in this factor.

To evaluate the relevance of this factor, a comparison between speciated profiles from tunnel measurements (Fig. 1) and PMF simulations was done and is reported in Figure 9. Traffic profiles are in general coherent and consistent amongst themselves, thus allowing to label this factor as a “Motor Vehicle Exhaust” source. Indeed, a good agreement is observed between these two profiles for the major species such as *iso-/n*-pentane, toluene, ethylene, propene. Instead, significant differences in mass contributions for ethane (almost a factor of 10), acetylene (considered as a key combustion compound emitted from traffic not identified in the PMF profile), isoprene (represented by evaporative sources) and oxygenated species were found. These differences can potentially be explained by several reasons. Firstly, the proportion of VOC emitted from traffic may be different depending upon the types of vehicles/engines/fuels (Montero et al., 2010). VOC emissions can also be dependent on the use of vehicles (age, maintenance), driving situations and thermal conditions (hot soak). Secondly, the vehicle fleet composition is different in the centre of Paris and in a highway tunnel. Although the proportion of passengers cars and Light Duty Vehicles (LDV) accounts for 60–90 % of the total composition of the fleet in circulation in both cases, the share of two-wheelers and heavy goods vehicles can be different. Indeed, heavy vehicles are subject to traffic limitations prohibiting their entry in Paris whereas as they are allowed in the highway tunnel (5 %). The proportion of two-wheelers is significant in Paris (10–20 %) (Airparif, 2013) while they represent less than 2 % of the total vehicles in the tunnel. Finally, PMF artefacts cannot be excluded. We suppose that the contribution of ethane in the “Motor Vehicle Exhaust” factor is overestimated (same as the “Wood Burning” factor) at the expense of the mixed “Natural Gas and Background” source (for which ethane contributions seem to be underestimated).

This factor 1 displays fair correlations with nitrogen monoxide/dioxide (NO/NO_2), carbon monoxide (CO) and black carbon (from its fossil fuel fraction), which are known to be relevant vehicle exhaust markers ($0.53 < r < 0.64$). The average contribution of this factor is rather stable throughout the year (5.8 $\mu g\ m^{-3}$ - see Figure 10, panel 1 and Section S7 in the Supplement). A smaller contribution is found during winter (3.2 $\mu g\ m^{-3}$) whereas the highest emissions from motor vehicle exhaust occur in autumn (8.6 µg m^{-3} on average) with a contribution of up to 10.1 µg m^{-3} in September. This seasonal cycle has already been observed and described in Bressi et al. (2014) for the road-traffic source of fine aerosols in Paris. The diurnal variation of this source is characterized by a “double wave” profile with an initial increase at 07-10 h (LT, Local Time) and a second increase at the end of the afternoon between 16 and 19 h (LT). These increases correspond to morning and evening rush-hour traffic periods. After 21h, the absolute contributions of this factor stay quite high (7.0 – 8.0 µg m^{-3}) due to several factors: on-going emissions (until midnight), lower photochemical reactions and atmospheric dynamics (the shallower boundary layer leads to more accumulation of pollutants at night). Lower contributions are generally displayed on late mornings/early afternoons. This reduction in factor contributions could be mainly explained by dilution and OH oxidation processes of more reactive species, which are not being balanced by additional vehicular emissions. This pronounced cycle has already been reported in previous studies (Gaimoz et al., 2011 and references therein). The temporal source strength variation is usually much more pronounced during weekdays than the weekend.

3.4.2 Evaporative sources factor

The profile of Factor 2 (see Fig. 8b) exhibits a high contribution from propane and *iso-/n*-butanes, with more than 47 % of their variabilities explained by this factor. It was already identified by Gaimoz et al. (2011) and is used here as reference profile from “gasoline evaporation” emissions (including storage, extraction and distribution of gasoline or Liquid Petroleum Gas (LPG)). The generic term "Evaporative sources" is here used to take into account these types of evaporative emissons. Factor 2 also includes a significant proportion of isoprene (20 %). This finding is still consistent with the conclusions of Borbon et al. (2001), which have shown that traffic activities emit a small amount of isoprene. In the same way, oxygenated compounds (acetaldehyde (4 %), acetone (6.6 %)) were found in fugitive evaporative emissions in agreement with what was observed during the highway tunnel experiment (see Fig. 2).

Among independent tracers used, only NO displays a fair agreement with this factor (r = 0.35). A correlation between F2 and F1 can also be noted (r = 0.36), thus indicating that these two factors are related to a common source (e.g. road-traffic). This source is in the range of ~ 3.9 µg m^{-3} over the whole studied period (see Fig. 10, panel 2). The annual trend of F2 seems to be consistent with the motor vehicle exhaust factor (F1), even though its monthly change remains ambiguous. Indeed, lower evaporative contributions are recorded both in winter and in early summer with minimum average contributions in June and July (1.7 µg m^{-3}). This finding was already identified by Frachon H. (2009; pers. communication). However, this value in June is somewhat puzzling as road-traffic emissions are usually significant (4.9 µg m^{-3}). In July, propane and butanes (*iso-/n-*) values were missing due to analytical problems on the operating GC-FID. Consequently, these compounds were simulated by the PMF model (e.g. missing values were virtually substituted by median values) which may underestimate the contribution of this factor during this specific period of time. However, high contributions of this source occur in August (6.4 µg m^{-3}). Although exhaust emissions are not particularly important, this observation could be eventually explained by gasoline storage and distribution sources, which may have increased with higher temperatures during that month. Maxima temperatures have generally been in the range of 16 to 32°C. The source contribution is in average higher in autumn (6.1 µg m^{-3}) with a contribution of up to 6.3 µg m^{-3} in October.

The diurnal variation of this factor contribution is characterized by a nighttime minimum, an increase from 07 to 10 h (consistent with the motor vehicle exhaust factor, F1) and a much slower decrease in emissions during the afternoon than those observed for the vehicle combustion profile. This second factor therefore represents the emissions of less reactive species (OVOCs, propane, butanes), for which concentrations cannot be expected to be consumed photochemically in short transport times. The temporal source strength variation is less pronounced on weekends than weekdays, which is typical of mobile source activity patterns.

According to the Copert IV (European Environment Agency, EEA) program for the calculation of air pollutant emissions from road transport, gasoline evaporation emissions can be explained by the evaporation of VOCs due to temperature, vehicle refueling, running losses, diurnal and “hot soak” reactions (when a hot engine is switched off). It was speculated that hot engines would emit more in the morning than in the evening, considering typical conditions of active inhabitants going to and from their workplace. Fugitive gasoline emissions from the loading of tank trucks, transportation and unloading from tank trucks at service stations and distributions depots can also be likely sources of this factor. In summary, this source depends on several parameters (related to road-traffic conditions, the vehicle fleet composition, economic activities and meteorological observations), which can make the interpretation of its seasonal variability difficult.

3.4.3 Wood Burning factor

In Paris, domestic wood burning represents a non-negligible part (about 5 %) of the energy consumption by fuel used for home heating (Airparif, 2011). The chemical profile of this source (Factor 3), shown in Fig. 8c, is mainly dominated by acetylene with approximately 80 % of its variability explained by this factor. It also includes ethylene (57.4 %), benzene (22.7 %) and oxygenated compounds, such as acetonitrile, acetaldehyde and methanol (with 18.3 %, 12.6 % and 8.2 %, respectively). Acetonitrile is a hydrocarbon commonly used as a marker of biomass burning (Holzinger et al., 1999). All these chemical species typically reflect an anthropogenic source related to wood combustion processes (Lanz et al., 2008; Leuchner et al., 2015) in agreement with the fireplace emission profile (see Sub-section 2.4.2, Fig. 3). No full comparison between both speciation profiles was possible as the fireplace profile was based on a limited number of data. With this mind, only a qualitative approach allowed to identify predominant species emitted from this source and confirm the term "wood burning" assigned to this factor.

Biomass burning emissions are well correlated with black carbon originating from residential wood burning (BC*wb*) and carbon monoxide, a long-lived compound especially emitted from combustion reactions ($0.6 < r < 0.7$). In addition, they well co-vary with napthalene (*m/z* 129.0 measured by PTR-MS) - a known polyaromatic hydrocarbon emitted from combustion processes (industry, tailpipe emissions) including wood burning (Purvis and McCrillis, 2000). As expected, wood burning contributions display a distinct cycle with a winter maximum (20.5 $\mu g\,m^{-3}$ on average) and a summer minimum (3.3 $\mu g\,m^{-3}$). Average contributions of this factor are rather stable in both spring and fall (6.9 and 5.9 $\mu g\,m^{-3}$, respectively).

Wood burning emissions linked to home/building heating are obviously highly dependent on meteorological conditions and particularly on cold temperatures. A clear negative relationship between the wood burning factor and temperature is found (r= - 0.56). The diurnal variation of this source exhibits a “double wave” profile. Average contributions increase from sunrise to a maximum in midmorning and decrease until 16 - 17 h. At the end of the day, a second increase is observed with another maximum contribution at 19 - 21 h. This diel cycle can be explained by domestic behaviors. An important finding is that the diurnal pattern of this source is fairly comparable to that of the motor vehicle exhaust factor. However, the wood burning factor does not display any distinct weekly variation. High contributions are observed all week (without any distinction between weekdays and weekends) compared to motor exhausts, for which vehicular emissions are less pronounced on weekends than weekdays. In addition, it exhibits poor correlations with NO, NO_2 and BC*ff* (r = 0.30, 0.29 and 0.19, respectively), thus indicating that the wood burning factor is completely independent of the motor vehicle exhaust source.

3.4.4 Biogenic factor

The profile of Factor 4 (see Fig. 8d) exhibits a high contribution from isoprene, a known chemical marker of biogenic emissions, with more than 79 % of its variability explained by this factor. In addition, this factor profile includes isoprene's oxidation products (methylvinylketone (MVK), methacrolein (MACR) and isoprene hydroxy hydroperoxides (ISOPOOHs)) with more than 48 %, methanol and acetone - a selection of compounds having a large contribution from biogenic emissions (Kesselmeier and Staudt, 1998; Guenther, 2002). It also accounts a significant contribution of some light alkenes (e.g. ethylene and propene), which can be evenly emitted by plants (Goldstein et al., 1996). Consequently, this factor F4 is termed "biogenic factor". Amounts of light alkanes (butanes, *iso*-pentane, *n*-hexane) and acetonitrile were also found in this profile and could be attributed to a mixing with other temperature-related sources or artefacts from the PMF model (Leuchner et al., 2015).

Biogenic emissions are directly related to solar radiation (Steiner and Goldstein, 2007) and ambient temperature ($r > 0.7$). For that reason, the highest biogenic factor contributions occur in summer (10.5 µg m^{-3} on average) with a contribution of up to 14.3 µg m^{-3} in July. Daily mean contributions gradually increase from 09 h. A slight delay (03 h) is observed in comparison with diurnal temperature/solar radiations variations (for which values increase from sunrise at 06 h). We assume that chemistry affects this source factor as it takes part in the formation of secondary species (MAK+MVK+ISOPOOH, for instance) from the oxidation of primarily emitted compounds (isoprene, OVOC). Diurnal contributions reach their maximum at the end of the day (19 h). Highest nighttime contributions of this source can be explained by the presence of oxygenated species (long-lived compounds already present in the atmosphere and/or secondarily formed from the oxidation of isoprene) in the profile combined with lower photochemical reactions and atmospheric dynamics (a low PBL height) at night.

3.4.5 Solvents use factor

The profile of Factor 5, shown in Fig. 8e, is associated with a large contribution of selected OVOCs (acetaldehyde, methanol and acetone) with on average ~ 33 % of their variabilities explained by this factor. Significant contributions from aromatic compounds (toluene, xylenes + C_8 and benzene) and some alkanes (pentanes, butanes, propane and *n*-hexane) are also observed. Toluene, in addition to road-traffic, is a good marker for solvents originating from an industrial source (Buzcu and Fraser, 2006). Benzene, due to its toxic and carcinogen nature, was regulated in recent years and is strongly limited in solvent formulations. Current standards establish limits in benzene concentrations at 0.1 % in cleaning products. However, PMF results point out the presence of benzene in this factor, suggesting that this compound might potentially still be in use by some manufacturers. Finally, the presence of these aforementioned species illustrates that this profile could be linked to industrial emissions, although a mixing of different sources cannot be excluded.

This factor co-varies well with ethanol, butan-2-one (also called methylethylketone – MEK), isopropyl alcohol or even ethyl acetate (0.68 > r > 0.52, respectively) – four organic compounds that were measured by GC-MS during the MEGAPOLI campaign (January-February 2010). These species are often used as solvents, diluents or cleaning fluids in industrial processes (Zheng et al., 2013). Some manufactories can consume fossil fuels for their activities, which may explain the fairly good correlation between this factor and black carbon originating from fossil fuels (BC*ff*, r = 0.50). Indeed, these fossil fuels could be used by industries as diverse as paints, paintings inks and lacquers (Tsai et al, 2001; Cornelissen and Gustafsson, 2004).

The highest contribution of this factor is observed during winter (14.2 µg m^{-3}) with a contribution of up to 20.9 µg m^{-3} in January. In winter, factor contributions increase at 06 h and reach their maximum between 11 h and 19 h (15 – 20 µg m^{-3}) before a long and gradual decline in the evening (see Fig. 10, panel 5 – top right). Higher contributions in winter can be explained by lower photochemical reactions (combined with weaker OH concentrations/UV radiations) and atmospheric dynamics. Indeed, a shallower PBL (and consequently, less intense vertical dynamics) leads to more accumulation of pollutants and thus to higher source contributions. The daily wintertime variability of this source is in agreement with the diel cycle of independent tracers (ethanol, butan-2-one).

Reconstructed contributions associated with this factor are also significant in summer (12.6 µg m^{-3} in July), which could be mainly explained by the evaporation of solvent inks, paints and other applications during that month due to higher temperatures. In spring/summer/autumn, factor contributions also increase at sunrise, but reach their maximum between 08 and 10 h (typical of anthropogenic activities). They progressively decrease during the afternoon (see Fig.10, panel 5 – bottom right). This gradual decline (not earlier observed in winter) is influenced by greater photochemical reactions and more intense vertical dynamics during these three seasons, leading to dispersion and dilution processes (and consequently, lower source contributions during the afternoon).

The temporal source strength variation is much more pronounced during weekdays than the weekend, except on Saturday morning. These diel and weekly patterns seem to be consistent with industrial source activities.

3.4.6 Natural Gas and Background factor

The profile of Factor 6, shown in Fig. 8f, is mainly dominated by ethane with around 45 % of its variability explained by this factor. It also contains propane (14.7 %) and light alkanes (butanes), which are key long-lived compounds known to be associated with natural gas leakages. Such species have already been identified in the natural gas experiment (see Sub-section 2.4.2, Fig. 4), thus allowing to confirm the identification of this profile. The diel pattern of this factor is mainly based on the diurnal variation of ethane, which is characterized by a nighttime maximum and a mid-afternoon minimum. Mainly due to its low reactivity, the behavior of ethane can be interpreted as homogeneous species levels during the night under a shallow inversion layer, then followed by concentration reductions caused by the increase of the PBL and vertical mixing - leading to dispersion and dilution processes. Average contributions of this factor were significantly higher when the PBL was low (~11.0 – 14.0 µg m^{-3}) and lower when the PBL was high (~ 6.0 µg m^{-3}).

This F6 profile is also characterized by the presence of oxidized pollutants (OVOCs including acetone and methanol) and aromatic compounds (like benzene), which have relatively long atmospheric residence times of respectively 53, 12 and 9 days (assuming OH = $2.0 \text{ x } 10^6$ molecules cm^{-3}) [Atkinson et al., 2000]. Because of their low reactivity, all the species of this factor tend to accumulate in the atmosphere and show significant background levels, especially in the northern hemisphere. The resulting emissions can be considered as a partly aged background air, implying a possible regional background and/or a long-range (intercontinental) transport.

The average contribution of this mixed source (combining both natural gas and background emissions) is in the range of 9.2 µg m^{-3} during the whole studied period. Lowest source contributions were observed in winter, which does not fit with that reported in the literature. As mentioned in the "Motor vehicle exhaust" and "Wood burning" sub-sections (3.4.1 and 3.4.3, respectively), PMF artefacts cannot be ruled out. Indeed, a problem with the distribution of ethane (considered as the key species of the mixed source) within PMF factors was raised. We assumed that higher ethane contributions were partly assigned to the "motor vehicle exhaust" and "wood burning" factors. Consequently, we assumed that the "natural gas and background" factor contributions were underestimated (especially in winter) for the benefits of the "wood burning" factor (another source significantly contributing during this season).

The highest contributions occur in spring (13.3 µg m^{-3}) when the Paris region is mostly influenced by prevailing air masses originating from the north and the northeast parts of Europe passing over Germany and the Benelux area (see Fig. 5). These continental imports constitute background events, which significantly impact baseline levels of ethane and oxygenated species. Slightly lower reconstructed mass contributions of this factor F6 were also observed in autumn. This fact can be explained by the consumption of natural gas (for home heating) during this season as average temperatures are progressively going down. No significant continental influences occur during the fall period as main air masses were coming from the west, south and southeast sectors, thus illustrating the importance of local pollution emissions during this season.

3.5 VOC source contributions

PMF simulations revealed the significant contribution of six VOC emission sources (e.g. five specific factor profiles and a mixed one, for which the natural gas source could not be isolated from background emissions). This Source Apportionment (SA) analysis concluded that the predominant sources at the receptor site were road-traffic-related activities (including motor vehicle exhaust, 15 % of the Total VOC (TVOC) mass on the annual average, and evaporative sources, 10 %), with the remaining sources from natural gas and background (23 %), solvents use (20 %), wood-burning (17 %) and biogenic activities (15 %). Each modeled factor exhibits distinct patterns due to the variations of the different source emissions and meteorological conditions. Monthly averaged contributions (expressed in %) of these factors to TVOC mass are reported in Figure 11. Seasonal variations of the individual sources have already been commented in the previous sections. Therefore, only the most important features are given here.

Road-traffic emissions were identified by PMF simulations to be the main source of VOCs in Paris. The sum of motor vehicle exhaust and evaporative sources accounted for a quarter of the TVOC mass. It showed higher contributions at the end of the year (21 % and 15 %, respectively), that is still consistent with the study from Bressi et al. (2014) and with long-term black carbon measurements (Petit et al., 2015) linked to enhanced traffic during autumn in Paris. Most importantly, it was observed that the wood burning source exhibited a significant contribution in winter months (almost 50 % in January and February), which is still in agreement with wood-burning related-particle emissions (Favez et al., 2009). The biogenic source also displayed a significant contribution (~30 %) in summer (mainly due to the weight of oxygenated species in the factor profile). The solvent use source displayed high contributions during winter months (~33 %, due to a lower PBL height and slower photochemical reactions during that period) and in July (due to the evaporation of solvents controlled by temperature). The source mixing natural gas and background showed a higher proportion in springtime (~34 %) and lower proportions during autumn (~25 %). This conclusion can be explained by pollution events that are both related to air masses imported from continental Europe (see Fig. 5) and/or specific meteorological conditions (low temperatures involving the use of home heating), respectively.

The reactivity of each modeled factor has also been investigated by considering the factor concentration of each species with their OH rate constant ($\mathrm{k_{OH}}$) (Atkinson and Arey, 2003) and is reported in relative (and absolute) contributions in Figure. 11. Among all the emission sources identified by PMF, solvents use and motor vehicle exhaust factors appear as the main reactive sources [26 % ($33\ \mathrm{s^{-1}}$) and 23 % ($40\ \mathrm{s^{-1}}$), respectively]. This can be explained by high constant rates of aromatics and alkenes mainly associated to these two emission sources. The contribution of the biogenic source is surprisingly weak (17 %). Although isoprene is an extremely reactive species, this factor exhibits a high weight of OVOCs for which constant rates can be low. Instead, the relative contribution of the mixed source “natural gas and background” is surprisingly high (16 %) due to the presence of aromatics (toluene and xylenes) in the factor profile. The lower contribution of reactivity is represented by the evaporative sources factor [5 % ($13\ \mathrm{s^{-1}}$)] which contains more stable gases (propane, butanes).

3.6 Comparison with previous Source Apportionment (SA) studies performed in Paris

Based on 1-year daily $\mathrm{PM_{2.5}}$ measurements (September 2009 – September 2010), Bressi et al. (2014) also conducted a Source Apportionment (SA) analysis using the PMF method (EPA PMF 3.0) with the aim of identifying and characterizing major fine aerosols emission sources within the Paris area. Seven factors, namely Ammonium Sulfate (A.S.)-rich factor, Ammonium Nitrate (A.N.)-rich factor, heavy oil combustion, road-traffic, biomass burning, marine aerosols and metal industry were identified. Special attention is paid here to common modeled factor categories.

Primarily of local origin, the road-traffic source (resulting from exhaust and non-exhaust processes) constitutes approximately 14 % of $\mathrm{PM_{2.5}}$ mass (~2.1 µg $\mathrm{m^{-3}}$, on average) over the whole sampling period. Its annual contribution was considered as significant but surprisingly low given the high traffic density in Paris and its surroundings. It exhibits stable averaged contributions throughout the year, with a smaller proportion in winter (6 %, 1.3 µg $\mathrm{m^{-3}}$) and higher in autumn (19 %, 2.5 µg $\mathrm{m^{-3}}$). This temporal source variation is still in agreement with the seasonal cycle of the road-traffic source (combining motor vehicle exhaust and evaporative running losses) issued from our VOC PMF analysis (see Sub-sections 3.4.1 and 3.4.2). The second common wood burning source is estimated for the first time over long periods and contributes to around 12 % (1.8 µg $\mathrm{m^{-3}}$) of the total $\mathrm{PM_{2.5}}$ mass. As expected, higher contributions were significantly observed during winter (22 %, 4.7 µg $\mathrm{m^{-3}}$) and in autumn (18 %, 2.4 µg $\mathrm{m^{-3}}$). This finding is still consistent with the seasonal pattern of the wood burning VOC source. Because of the daily time resolution of filter sampling, no diurnal variation of modeled sources was reported in Bressi et al.

(2014), thus limiting any additional comparison with this study.

Based on 1-month VOC measurements (25 May – 14 June 2007) performed at the LHVP site, Gaimoz et al. (2011) also conducted a PMF analysis with the aim of identifying and apportioning major VOC sources in Paris. Seven factors, namely vehicle exhaust, fuel evaporation, remote industrial sources, natural gas + background, local sources, biogenic and fuel evaporation and wood burning, were found. For an appropriate comparison between this study and our work, special attention is paid here to the modeled speciation profiles and source contributions.

Chemical profiles from Gaimoz et al. (2011) revealed consistent findings with this study. The fuel evaporation factor is mainly composed of butanes, propane and ethane whereas the vehicle exhaust factor includes *iso*-pentane, benzene, toluene, C_8 and C_9-aromatics and in lower proportions ethylene, propene and acetylene. These observations are consistent with modeled evaporative sources and motor vehicle exhaust profiles obtained in this work. A biogenic and fuel evaporation source is also identified and essentially made of isoprene, methanol, acetone and a high proportion of *iso*-pentane, suggesting that this factor is mixing up biotic emissions and road-traffic activities. Highly dependent on (continental) air-mass origins, a remote industrial factor (related to industrial activities and long-range transport of secondary VOCs) is found to exhibit high contributions of OVOCs (methanol, acetone), aromatic species (toluene, C_8-C_9 aromatics) and some light alkanes. Our PMF study emphasized a solvents use source, for which these aforementioned compounds were observed, in addition to benzene. The wood burning source includes only a high contribution of acetonitrile although ethylene, acetylene and benzene are significantly emitted - in accordance with findings from the fireplace experiment (see Sub-section 2.4.2). The mixed natural gas + background source is especially driven by ethane, methanol and acetone. No aromatic species appear in this factor profile. Finally, the local source (LPG – Liquefied Petroleum Gas) including propane and pentanes seems to be associated with fuel evaporation sources. These kinds of emissions have been included in the evaporative sources factor.

During May – June 2007, Gaimoz et al. (2011) concluded that the major VOC sources were related to road-traffic emissions (~39 % of the TVOC mass), with the remaining emissions from wood burning (2 %), biogenic and fuel evaporation (5 %), remote industrial sources (35 %), natural gas and background (13 %) and local sources (7 %) during the whole studied period (Fig. 13, left pie chart). To accurately compare VOC sources proportions between 2007 and 2010 (for a similar combination of hydrocarbons and masses), the contribution of each main factor was recalculated for the specific time period May – June 2010 (Fig. 13, right pie chart).

Significant differences between biogenic and wood burning sources contributions could eventually be both explained by the weight of major OVOCs into speciation profiles (relative proportions of methanol, acetaldehyde, acetone in these factors are higher than those of the comparative study) and the differences in temperatures affecting the Paris region. The temperatures recorded in May – June 2007 / 2010 were 20°C and 16°C, respectively. This would explain a higher home heating consumption and consequently, a higher contribution of the wood burning factor in 2010 (9 % *versus* 2 % for the previous work). Regarding to solvents use source contributions, differences can also be explained by temperatures (as they constitute a relevant indicator in solvent emissions) and by the amount of solvents used in manufactories due to recent regulatory frameworks in place (20 % in 2010 *versus* 35 % in 2007). The difference of natural gas and background source contributions can be due to the importance of air masses coming from the north and northeast parts of Europe between 25 May and 14 June 2010. These air-mass origins were also observed in 2007 and could have affected remote industrial-related emissions and not the mixed source. Slight differences of the motor vehicle exhaust source between 2007 and 2010 (22 % *versus* 14 %) could be explained by densification strategies and technological innovations for reducing car use and emissions. Finally, observed differences for the evaporative sources factor (5 % for 2010 and 17 % in 2007) are related to emissions and high temperatures observed in 2007.

3.7 Comparison with some European SA studies

Yearly average contributions of the modeled VOC sources (see Section 3.5) were also compared with other Source Apportionment (SA) studies performed within urban areas in Europe. From the different European SA studies available, only one is based on a long VOC time series, which is strengthening the novelty and the originality of the current study.

Based on 2-years hourly measurements of C_2 - C_7 NMHCs, Lanz et al. (2008) permitted the identification and characterization of between eight and six emission sources at an urban background site in Zürich (Switzerland) in the years 1993-1994 and 2005-2006. Only measurements from 2005 to 2006 are compared here as they are the most recent observations we have available. Six factors, namely gasoline evaporation, solvents, propane, ethane, wood burning and fuel combustion were determined using the PMF method. This SA study highlighted the importance of vehicular, solvents use, wood burning and gas leakage emissions. The road-traffic-related source included both gasoline evaporation and fuel combustion (motor exhaust) factors. While the first factor is mainly dominated by butanes (*iso-/n-*) and *iso*-pentane, the second one is essentially driven by ethane, ethene, propene, benzene and toluene. These two speciation profiles are still consistent with those obtained from this PMF analysis, except for *iso*-pentane. Considered as a key species of evaporative processes, *iso*-pentane mostly contributed to the motor vehicle exhaust source (Fig. 8a). It was also identified as one of the main compounds emitted in the highway tunnel experiment (Fig. 2), where measured hydrocarbons were representative of fresh emissions (e.g. fuel combustion). This modeled road-traffic source contributed to 26 % of the TVOC mass (13 % for gasoline evaporation and 13 % for fuel combustion factors), which is in the same order of magnitude than that of our vehicle-related source (25 %). The solvents use factor is characterized by pentanes, S-isohexanes (including 2-methylpentane, 3-methylpentane, 2,2-dimethylbutane, 2,3-dimethylbutane) and toluene, in agreement with our solvents use VOC profile which also included oxygenated species (not measured in Lanz et al. (2008)). This industrial factor accounted for 20 % of the TVOC mass. This source contribution is comparable to what we obtained from January to November 2010 (20 %). The wood burning factor is mainly dominated by ethylene, acetylene, ethane, benzene and contributed to 16 % of the TVOC mass for the 2005-2006 sampling period. This finding is fairly in agreement with our annual wood burning contribution. Finally, a natural gas source was also identified and consisted of the combination of two separated factors (“ethane” with “propane”). Its annual contribution is evaluated at 35 % of the TVOC mass whereas our mixed natural gas and background source accounted for 23 %. No biogenic source was detected for this comparative study. To sum up, average contributions of the road-traffic, solvents use and wood burning sources well matched between this SA study and our modeled results although the input chemical matrix and sampling dates are different.

The importance of these three anthropogenic sources was often reported in other existing urban SA studies from short-term measurements performed in Europe. For instance, Niedojadlo et al. (2007) (Wuppertal, Germany) paid particular attention to solvents use and road-traffic sources contributions using the Chemical Mass Balance (CMB) modeling technique. Main results showed that the road-traffic source dominated total VOC emissions (more than 50 % of the total mass) compared to solvents use. In addition, it was considered that the proportion of solvents emissions to TVOC concentrations fell in the range of ~20 % in German cities, which is significantly consistent with Lanz et al. (2008) and with this SA study in Paris.

The consistency in VOC contributions in European urban areas raises the question of their representativity at a larger scale. There are currently many other urban SA studies described in the literature (e.g. Jorquera and Rappenglück, 2004 – Santiago/Chile; Buzcu et al., 2006 – Houston/TX; Brown et al., 2007 – LA; Cai et al., 2010 – Shanghai/China; Morino et al., 2011 – Tokyo/Japan; Yurdakul et al., 2013 – Ankara/Turkey; Zheng et al., 2013 – Mexico). Results of these studies are not detailed here but one common feature for European and global scales is the importance of the road-traffic source (between 30 % and 50 %). One difference concerns the industrial sector which plays (in the investigated European cities) a lower role than in studied urban areas from other continents.

Governmental regulations and standards to control pollutants emissions and economic developments may differ between European countries and the rest of the world. The location of sampling points (or distances from main sources) and meteorological conditions can strongly affect VOC concentrations and their respective emission sources in the considered urban environments.

4 Conclusions

Within the framework of the EU-F7 MEGAPOLI and PRIMEQUAL-FRANCIPOL research programs, a selection of volatile organic compounds (VOCs) were continuously measured in real-time at two background urban sites located in downtown Paris (France) from 15 January to 22 November 2010. Assessed hydrocarbons included alkanes, alkenes-alkynes, isoprene, aromatics and oxygenated compounds (OVOCs). The current study allowed evaluating VOC concentration levels in ambient air and describing their temporal (seasonal and diurnal) time courses over a long period of time in the French megacity. It also showed an innovative methodology to identify, quantify and understand the main VOC emission sources in Paris by combining field experiments (near-field and ambient air measurements) with *source-receptor* statistical modeling. The modeled factor profiles were interpreted with respect to those obtained from literature and from three near-field experiments (inside a highway tunnel, at a fireplace and from a domestic gas flue) performed within the Paris area. These additional measurements helped better characterizing and/or confirming traffic, wood burning and natural gas-related sources among the existing different source profiles, which can be directly derived from a PMF modeling analysis. These source profile studies therefore allowed to check the representativity and the robustness of our conclusions. This PMF analysis successfully reconstructed at least 88 ± 2 % of the measured total VOC mass.

Among the six identified PMF factors, road-traffic activities appeared to be the main VOC source in Paris with an average contribution of 25 % of the TVOC mass at the annual scale. This source both included motor vehicle exhaust (15 %) and evaporative sources (10 %). For the first time, it was also shown that the residential wood burning source exhibited an important contribution in winter (almost 50 %) due to cold temperatures during that season (home heating consumption). The biogenic source also displayed a significant contribution (~30 %) in summer mainly due to the weight of oxygenated species in the factor profile. A solvents source was identified and annually contributed to 20 % of the total VOC mass. Finally, it was also revealed that a source mixing natural gas and background (23 %) could be highly dependent on air-mass origins (especially during continental-influenced periods) and meteorological conditions (temperatures). It exhibited a higher proportion in springtime (34 %, explained by intercontinental imports) and in autumn (25 %, partly for home heating consumption reasons).

From this initial source apportionment study, natural gas could not be isolated from background emissions by the PMF method, thus leading to a limitation of this analysis. A further work will aim at constraining the reference speciation profile (obtained from domestic gas flue measurements) in order to evaluate the relative contribution of natural gas emissions. Lastly, the quantitative assessment of the contributions from different modeled sources presented in this study will provide an independent evaluation of the quality and the relevance of the corresponding emission inventories. In particular, the comparison will be very valuable with the updated local emission inventory (provided by the regional air quality network AIRPARIF) as some discrepancies had been pointed out with its previous version.

Acknowledgements. Authors would like to thank B. Temime-Roussel and N. Marchand from the Laboratoire Chimie Provence (LCP, University of Provence, Marseille, France) for PTR-MS measurements performed at the LHVP site during the MEGAPOLI winter campaign. We would like to thank also M. Beekmann for the coordination of the EU-F7 MEGAPOLI project. We gratefully acknowledge MM. Squinazzi and Y. Le Moullec for having hosted MEGAPOLI and FRANCIPOL intensive campaigns as well as all colleagues involved in the monitoring process of ambient air measurements, especially Hanitriniala Ravelomanantsoa, Thomas Chaigneau (LHVP) and Laurent Martinon (Laboratoire d'Etude des Particules Inhalées, LEPI). T. Le Priol and J.-F. Petit from the Centre d'Etudes et d'expertise sur les Risques, l'Environnement, la Mobilité et l'Aménagement (CEREMA) are also acknowledged for the logistical assistance in the road-tunnel experiment. This work was supported by the CEA, CNRS, IPSL, ADEME, Île-de-France region funds, the EU-PF7 ANR MEGAPOLI and the French PRIMEQUAL-FRANCIPOL, PREQUALIF and CORTEA-CHAMPROBOIS projects. We are thankful to Sabina Assan for helping with the English version of the manuscript.

References

Airparif: Origine des particules respirées en Île-de-France, Press Pack, http://www.airparif.asso.fr/_pdf/publications/synthese_particules_110914.pdf, 2011.

Airparif: Evolution de la qualité de l'air à Paris entre 2002 et 2012, http://www.airparif.asso.fr/_pdf/publications/rapport-pdp-130703.pdf, 2013.

Airparif: Surveillance et information sur la qualité de l'air en Île-de-France en 2014, available at: http://www.airparif.asso.fr/_pdf/publications/bilan-2014.pdf, 2015.

Ait-Helal, W., Borbon, A., Sauvage, S., de Gouw, J. A., Colomb, A., Gros, V., Freutel, F., Crippa, M., Afif, C., Baltensperger, U., Beekmann, M., Doussin, J.-F., Durand-Jolibois, R., Fronval, I., Grand, N., Leonardis, T., Lopez, M., Michoud, V., Miet, K., Perrier, S., Prévôt, A. S. H., Schneider, J., Siour, G., Zapf, P and Locoge, N.: Volatile and intermediate volatility organic compounds in suburban Paris: variability, origin and importance for SOA formation, Atmos. Chem. Phys.,14, 10439-10464, doi:10.5194/acp-14-10439-2014, 2014.

Ammoura, L., Xueref-Remy, I., Gros, V., Baudic, A., Bonsang, B., Petit, J.- E., Perrussel, O., Bonnaire, N., Sciare, J. and Chevallier, F.: Atmospheric measurements of ratios between $\mathrm{CO_2}$ and co-emitted species from traffic: a tunnel study in the Paris megacity, Atmos. Chem. Phys.,14, 20197-20229, doi:10.5194/acp-14-12871-2014, 2014.

Araizaga, A. E., Mancilla, Y. and Mendoza, A.: Volatile Organic Compound Emissions from Light-Duty Vehicles in Monterrey, Mexico: a Tunnel Study, Int. J. Environ. Res., 7 (2), 277-292, 2013.

Atkinson, R.: Atmospheric chemistry of VOCs and $\mathrm{NO_x}$, Atmos. Environ., 34, 2063-2101, doi:10.1016/S1352-2310(99)00460-4, 2000.

Atkinson, R. and Arey, J.: Atmospheric Degradation of Volatile Organic Compounds, Chem. Rev., 103(12), 4605-4638, doi:10.1021/cr0206420, 2003.

Atkinson, R., Baulch, D. L., Cox, R. A., Crowley, J. N., Hampson, R. F., Hynes, R. G., Jenkin, M. E., Rossi, M. J. and Troe, J.: Evaluated kinetic and photochemical data for atmospheric chemistry: Volume II – Gas phase reactions of organic species, Atmos. Chem. Phys., 6, 3625-4055, doi:10.5194/acp-6-3625-2006, 2006.

Baklanov, A., Lawrence, M., Pandis, S., Mahura, A., Finardi, S., Moussiopoulos, N., Beekmann, M., Laj, P., Gomes, L., Jaffrezo, J.- L., Borbon, A., Coll, I., Gros, V., Sciare, J., Kukkonen, J., Galmarini, S., Giorgi, F., Grimmond, S., Esau, I., Stohl, A., Denby, B., Wagner, T., Butler, T., Baltensperger, U., Builtjes, P., van der Hout, D., van der Gon, H. D., Collins, B., Schluenzen, H., Kulmala, M., Zilitinkevich, S., Sokhi, R., Friedrich, R., Theloke, J., Hummer, U., Jalkinen, L., Halenka, T., Widensholer, A., Pyle, J. and Rossow, W. B.: MEGAPOLI: concept of multi-scale modeling of megacity impact on air quality and climate, Adv. Sci. Res., 4, 115-120, doi:10.5194/asr-4-115-2010, 2010.

Barrefors, G. and Petersson, G.: Volatile hydrocarbons from domestic wood burning, Chemosphere, 30 (8), 1551-1556, doi:10.1016/0045-6535(95)00048-D, 1995.

Beekmann, M., Prévôt, A. S. H., Drewnick, F., Sciare, J., Pandis, S. N., Denier van der Gon, H. A. C., Crippa, M., Freutel, F., Poulain, L., Ghersi, V., Rodriguez, E., Beirle, S., Zotter, P., von der Weiden-Reinmüller, S.- L., Bressi, M., Fountoukis, C., Petetin, H., Szidat, S., Schneider, J., Rosso, A., Haddad, I. El., Megaritis, A., Zhang, Q. J., Michoud, V., Slowik, J. G., Moukhtar, S., Kolmonen, P., Stohl, A., Eckhardt, S., Borbon, A., Gros, V., Marchand, N., Jaffrezo, J. L., Schwarzenboeck, A., Colomb, A., Weidensohler, A., Borrmann, S., Lawrence, M., Baklanov, A. and Baltensperger, U.: In situ, satellite measurement and model evidence on the dominant regional contribution to fine particulate matter levels in the Paris megacity, Atmos. Chem. Phys., 15, 9577-9591, doi:10.5194/acp-15-9577-2015, 2015.

Blake, R. S., Monks, P. S., and Ellis, A. M.: Proton Transfer Reaction Mass Spectrometer, Chem. Rev., 109, 861-896, doi:10.1021/cr800364q, 2009.

Borbon, A., Fontaine, H., Veillerot, M., Locoge, N., Galloo, J. and Guillermo, R.: An investigation into the traffic-related fraction of isoprene at an urban location, Atmos. Environ., 35, 3749-3760, doi:10.1016/S1352-2310(01)00170-4, 2001.

Bressi, M., Sciare, J., Ghersi, V., Bonnaire, N., Nicolas, J. B., Petit, J.- E, Moukhtar, S., Rosso, A., Mihalopoulos, N. and Féron, A.: A one-year comprehensive chemical characterisation of fine aerosol ($PM_{2.5}$) at urban, suburban and rural background sites in the region of Paris (France), Atmos. Chem. Phys., 13, 7825-7844, doi:10.5194/acp-13-7825-2013, 2013.

Bressi, M., Sciare, J., Ghersi, V., Mihalopoulos, N., Petit, J.- E., Nicolas, J. B., Moukhtar, S., Rosso, A., Féron, A., Bonnaire, N., Poulakis, E. and Theodosi, C.: Sources and geographical origins of fine aerosols in Paris (France), Atmos. Chem. Phys., 14, 8813-8839, doi:10.5194/acp-14-8813-2014, 2014.

Brown, S. G., Frankel, A. and Hafner, H. R.: Source apportionment of VOCs in the Los Angeles area using positive matrix factorization, Atmos. Environ., 41, 227-237, doi:10.1016/j.atmosenv.2006.08.021, 2007.

Butler, D.: Megacity project seeks to gauge urban pollution, Nature, 455, 142-143, doi:10.1038/455142b, 2008.

Buzcu, B. and Fraser, M. P.: Source identification and apportionment of volatile organic compounds in Houston, TX, Atmos. Environ., 40, 2385-2400, doi:10.1016/j.atmosenv.2005.12.020, 2006.

Cai, C., Genz, F., Tie, X., Yu, Q. and An, J.: Characteristics and source apportionment of VOC measured in Shanghai, China, Atmos. Environ., 44, 5005-5014, doi:10.1016/j.atmosenv.2010.07.059, 2010.

Carter, W. P. L.: Development of ozone reactivity scales for volatile organic compounds, J. Air Waste Manag. Asso., 44(07), 881-899, doi: 10.1080/1073161X.1994.10467290, 1994.

Chameides, W. L., Fehsenfeld, F., Rodgers, M. O., Cardelino, C., Martinez, J., Parrish, D., Lonneman, W., Lawson, D. R., Rasmussen, R. A., Zimmerman, P., Greenberg, J., Middleton, P. and Wang, T.: Ozone precursor relationships in the ambient atmosphere, J. Geophys. Res. Atmos., 97, D5, 6037-6055, doi:10.1029/91JD03014, 1992.

COPERT 4 (Computer Programme to calculate Emissions from Road Transport), European Environment Agency (EEA), available at: http://www.emisia.com/copert/.

Cornelissen, G. and Gustafsson, Ö.: Sorption of Phenanthrene to Environmental Black Carbon in Sediment with and without Organic Matter and Native Sorbates, Environ. Sci. Technol., 38, 148-155, doi:10.1021/ES034776m, 2004.

Crippa, M., DeCarlo, P. F., Slowik, J. G., Mohr, C., Heringa, M. F., Chirico, R., Poulain, L., Freutel, F., Sciare, J., Cozic, J., Di Marco, C. F., Elsasser, M., Nicolas, J. B., Marchand, N., Abidi, E., Widensohler, A., Drewnick, F., Schneider, J., Borrmann, S., Nemitz, E., Zimmermann, R., Jaffrezo, J.-L., Prévôt, A. S. H. and Baltensperger, U.: Wintertime aerosol composition and source apportionment of the organic fraction in the metropolitan area of Paris, Atmos. Chem. Phys., 13, 961-981, doi:10.5194/acp-13-961-2013, 2013.

Deguillaume, L, Beekmann, M. and Derognat, C.: Uncertainty evaluation of ozone production and its sensitivity to emission changes over the Île-de-France region during summer periods, J. Geophys. Res., 113, D02304, doi:10.1029/2007JD009081, 2008.

Directive 2000/69/EC of the European Parliament and of the Council of 16 November 2000 relating to limit values for benzene and carbon monoxide in ambient air. Official Journal of the European Communities. 13 December 2000.

Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe. Official Journal of the European Communities. 11 June 2008.

Dolgorouky, C., Gros, V., Sarda-Estève, R., Sinha, V., Williams, J., Marchand, N., Sauvage, S., Poulain, L., Sciare, J. and Bonsang, B.: Total OH reactivity measurements in Paris during the 2010 MEGAPOLI winter campaign, Atmos. Chem. Phys.,12, 9593-9612, doi:10.5194/acp-12-9593-2012, 2012.

Eurostat: last updated on July 30th, 2015 at : http://ec.europa.eu/eurostat/fr, 2015.

Evtyugina, M., Alves, C., Calvo, A., Nunes, T., Tarelho, L., Duarte, M., Prozil, S. O., Evtuguin, D. V. and Pio, C.: VOC emissions from residential combustion of Southern and mid-European woods, Atmos. Environ., 83, 90-98, doi:10.1016/j.atmosenv.2013.10.050, 2014.

Favez, O., Cachier, H., Sciare, J. and Moullec, Y. L.: Characterization and contribution to $\mathrm{PM}_{2.5}$ of semi-volatile aerosols in Paris (France), Atmos. Environ., 41, 7969-7976, doi:10.1016/j.atmosenv.2007.09.031, 2007.

Favez, O., Cachier, H., Sciare, J., Sarda-Estève, R. and Martinon, L.: Evidence for a significant contribution of wood burning aerosols to $\mathrm{PM}_{2.5}$ during the winter season in Paris, France, Atmos. Environ., 43, 3640-3644, doi:10.1016/j.atmosenv.2009.04.035, 2009.

Fillela, I. and Peñuelas, J.: Daily, weekly, and seasonal time courses of VOC concentrations in a semi-urban area near Barcelona, Atmos. Environ., 40, 7752-7769, doi:10.1016/j.atmosenv.2006.08.002, 2006.

Fortner, E. C., Zheng, J., Zhang, R., Berk Knighton, W., Volkamer, R. M, Sheehy, P., Molina, L. and André, M.: Measurements of Volatile Organic Compounds using Proton Transfer Reaction - Mass Spectrometry during the MILAGRO 2006 Campaign, Atmos. Chem. Phys., 9, 467-481, doi:10.5194/acp-9-467-2009, 2009.

Frachon, H: Utilisation des mesures de concentration de COV pour la validation des émissions et l'amélioration de la prévision de l'ozone - phase 2, Rapport de stage professionnel effectué en collaboration avec AIRPARIF, 2009.

Gaeggeler, K., Prévôt, A. S. H., Dommen, J., Legreid, G., Reimann, S. and Baltensperger, U.: Residential wood burning in an Alpine valley as a source for oxygenated volatile organic compounds, hydrocarbons and organic acids, Atmos. Environ., 42, 8278-8287, doi:10.1016/j.atmosenv.2008.07.038, 2008.

Gaimoz, C., Sauvage, S., Gros, V., Herrmann, F., Williams, J., Locoge, N., Perrussel, O., Bonsang, B., d'Argouges, O., Sarda-Estève, R. and Sciare, J.: Volatile Organic Compounds sources in Paris in spring 2007. Part II: source apportionment using Positive Matrix Factorization, Environ. Chem., 8, 91-103, doi:10.1071/EN10067, 2011.

Goldstein, A. H., Fan, S. M., Goulden, M. L., Munger, J. W. and Wofsy, S. C.: Emissions of ethene, propene and 1-butene by a midlatitude forest, Journal of Geophysical Research, 101, 9149-9157, doi:10.1029/96jd00334, 1996.

Gros, V., Sciare, J. and Yu, T.: Air-quality measurements in megacities: focus on gaseous organic and particulate pollutants and comparison between two contrasted cities, Paris and Beijing, C. R. Geoscience, 339, 764-774, doi:10.1016/j.crte.2007.08.007, 2007.

Gros, V., Gaimoz, C., Herrmann, F., Custer, T., Williams, J., Bonsang, B., Sauvage, S., Locoge, N., d'Argouges, O., Sarda-Estève, R. and Sciare, J.: Volatile Organic Compounds sources in Paris in spring 2007. Part I: qualitative analysis, Environ. Chem., 8, 74-90, doi:10.1071/EN10068, 2011.

Guenther, A.: The contribution of reactive carbon emissions from vegetation to the carbon balance of terrestrial ecosystems, Chemosphere, 49 (8), 837-844, doi:10.1016/S0045-6535(02)00384-3, 2002.

Gurjar, B.: Air quality in megacities. Retrieved from http://www.eoearth.org/view/article/51cbece17896bb431f68e326, 2014.

Gustafson, P., Barregard, L., Strandberg, B. and Sallsten, G.: The impact of domestic wood burning on personal, indoor and outdoor levels of 1.3-butadiene, benzene, formaldehyde and acetaldehyde, J. Environ. Monit., 9 (1), 23-32, doi:10.1039/B614142K, 2007.

Holzinger, R., Warneke, C., Hansel, A., Jordan, A., Lindinger, W., Scharffe, H., Schade, G. and Crutzen, P. J.: Biomass burning as a source of formaldehyde, acetaldehyde, acetone, acetonitrile and hydrogen cyanide, Geophys. Res. Lett., 26, 1161-1164, doi:10.1029/1999GL900156, 1999.

IARC: Outdoor air pollution a leading environmental cause of cancer deaths, Press Release n°221, 2013.

Jorquera, H. and Rappenglück, B.: Receptor modeling of ambient VOC at Santiago, Chile, Atmos. Environ., 38, 4243-4263, doi:10.1016/j.atmosenv.2004.04.03, 2004.

Karl, T., Jobson, T., Kuster, W. C., Williams, E., Stutz, J., Shetter, R., Hall, S. R., Goldan, P., Fehsenfeld, F. and Lindinger, W.: Use of proton-transfer-reaction mass spectrometry to characterize volatile organic compound sources at the La Porte super site during the Texas Air Quality Study 2000, J. Geophys. Res. Atmos., 235, 4508, doi:10.1029/2002JD003333, 2003.

Kesselmeier, J. and Staudt, M.: Biogenic Volatile Organic Compounds (VOC): An overview on Emission, Physiology and Ecology, Atmos. Chem., 33, 23-88, 1999.

Langford, B., Nemitz, E., House, E., Phillips, G. J., Famulari, D., Davison, B., Hopkins, J. R., Lewis, A. C. and Hewitt, C. N.: Fluxes and concentrations of volatile organic compounds above central London, UK, Atmos. Chem. Phys., 10, 627-645, doi:10.5194/acp-10-627-2010, 2010.

Lanz, V. A., Hueglin, C., Buchmann, B., Hill, M., Locher, R., Staehelin, J. and Reimann, S.: Receptor modeling of C_2 - C_7 hydrocarbon sources at an urban background site in Zurich, Switzerland: changes between 1993-1994 and 2005-2006, Atmos. Chem. Phys., 8, 2313-2332, doi:10.5194/acp-8-2313-2008, 2008.

Lanz, V. A., Henne, S., Staehelin, J., Hueglin, C., Vollmer, M. K., Steinbacher, M., Buchmann, B. and Reimann, S.: Statistical analysis of anthropogenic non-methane VOC variability at a European background location (Jungfraujoch, Switzerland), Atmos. Chem. Phys., 9, 3445-3459, doi:10.5194/acp-9-3445-2009, 2009.

Legreid, G., Reimann, S., Steinbacher, M., Staehelin, J., Young, D. and Stemmler, K.: Measurements of OVOCs and NMHCs in a Swiss Highway Tunnel for Estimation of Road Transport Emissions, Environ. Sci. Technol., 41 (20), 7060-7066, doi:10.1021/es062309+, 2007.

Lelieved, J., Evans, J. S., Fnais, M., Giannadaki, D. and Pozzer, A.: The contribution of outdoor air pollution sources to premature mortality on a global scale, Nature, 525, 367-371, doi:10.1038/nature15371, 2015.

Leuchner, M, Gubo, S, Schunk, C, Wastl, C, Kirchner, M., Menzel, A. and Plass-Dülmer, C.: Can positive matrix factorization help to understand patterns of organic trace gases at the continental Global Atmosphere Watch site Hohenpeissenberg ?, Atmos. Chem. Phys., 15, 1221-1226, doi:10.5194/acp-15-1221-2015, 2015.

Liu, S. C., Trainer, M., Fehsenfeld, F. C., Parrish, D. D., Williams, E. J., Fahey, D. W., Hübler, G. and Murphy, P. C.: Ozone production in the rural troposphere and the implications for regional and global ozone distributions, J. Geophys. Res. Atmos., 92, D4, 4191-4207, doi:10.1029/JD092Id04P04191, 1987.

Logan, J. A., Prather, M. J., Wofsy, S. C. and McElroy, M. B.: Tropospheric chemistry: A global perspective, J. Geophys. Res., 86, C8, 7210-7254, doi:10.1029/JC086Ic08P07210, 1981.

Mathé, F.: Evolution de la classification et des critères d'implantation des stations de mesure de la qualité de l'air - Participation à la réactualisation du guide de classification des stations, LCSQA-Mines Douai Report, 2010.

Météo-France, Bulletins Climatologiques Mensuels pour Paris et sa petite couronne de Janvier à Novembre 2010, available at: https://donneespubliques.meteofrance.fr (last access: July 2015).

Montero, L., Duane, M.,Manfredi, U., Astorga, C., Martini, G., Carriero, M., Krasenbrink, A. and Larsen, B.-R.: Hydrocarbon emission fingerprints from contemporary vehicle/engine technologies with conventional and new fuels, Atmos. Environ., 44, 2167-2175, 2010.

Morino, Y., Ohara, T., Yokouchi, Y. and Ooki, A.: Comprehensive source apportionment of volatile organic compounds using observational data, two receptor models, and an emission inventory in Tokyo metropolitan area, J. Geophys. Res., 116, D0211, doi:10.1029/2010JD014762, 2011.

Na, K.: Determination of VOC source signature of vehicle exhaust in a traffic tunnel, Journal of Environmental Management, 81, 392-398, doi:10.1016/j.jenvman.2005.11.004, 2006.

Nalin, F., Golly, B., Besombes, J.-L., Pelletier, C., Aujay, R., Verlhac, S., Dermigny, A., Fievet, A., Karoski, N., Dubois, P., Collet, S., Favez, O. and Albinet, A.: Fast oxidation processes from emission to ambient air introduction of aerosol emitted by residential log wood stoves, Atmos. Env., accepted for publication, 2016.

Ng, N. L, Kroll, J. H, Chan, A. W. H, Chhabra, P. S., Flagan, R. C. and Seinfeld, J. H.: Secondary organic aerosol formation from m-xylene, toluene, and benzene, Atmos. Chem. Phys., 7, 3909-3922,doi:10.5194/acp-7-3909-2007, 2007.

Niedojadlo, A., Heinz Becker, K., Kurtenbach, R. and Wiesen, P.: The contribution of traffic and solvent use to the total NMVOC emission in a German city derived from measurements and CMB modelling, Atmos. Environ., 41, 7108-7126, doi:10.1016/j.atmosenv.2007.04.056, 2007.

Norris, G., Duvall, R., Brown, S. and Bai, S.: EPA Positive Matrix Factorization (PMF) 5.0: Fundamentals & User Guide. Prepared for the US. Environmental Protection Agency (EPA), Washington, DC, by the National Exposure Research Laboratory, Research Triangle Park; Sonoma Technology, Inc., Petaluma, 2014.

Paatero, P. and Tapper, U.: Positive Matrix Factorization: A non-negative factor model with optimal utilization of error estimates of data values, Environmetrics, 5(2), 111-126, doi:10.1002/env.3170050203, 1994.

Paatero, P. and Hopke, P. K.: Discarding or downweighting high-noise variables in factor analytic models, Anal. Chim. Acta, 490, 277-289, doi:10.1016/S0003-2670(02)01643-4, 2003.

Paatero, P., Eberly, S., Brown, S. G. and Norris, G. A.: Methods for estimating uncertainty in factor analytic solutions, Atmos. Meas. Tech., 7, 781-797, doi:10.5194/amt-7-781-2014, 2014.

Passant, N. R.: Speciation of UK emissions of non-methane volatile organic compounds, AEA Technology, report N° AEAT/ENV/R/0545, Issue 1.

Petit, J.- E., Favez, O., Sciare, J., Crenn, V., Sarda-Estève, R., Bonnaire, N., Mocnik, G., Dupont, J.- C., Haeffelin, M. and Leoz-Garziandia, E. : Two years of near real-time chemical composition of submicron aerosols in the region of Paris using an Aerosol Chemical Speciation Monitor (ACSM) and a multi-wavelength Aethalometer, Atmos. Chem. Phys., 15, 2985-3005, doi:10.5194/acp-15-2985-2015, 2015.

Polissar, A. V., Hopke, P. K., Paatero, P., Malm, W. C. and Sisler, J. F.: Atmospheric aerosol over Alaska, 2. Elemental Composition and Sources, J. Geophys. Res. Atmos., 103, 19045-19057, doi:10.1029/98JD01212, 1998.

Politis, D. N. and White, H.: Automatic block-length selection for the dependent bootstrap, Econometrics Reviews, 23, 53-70, doi:10.1081/ETC-120028836, 2004.

Purvis, C. R. and McCrillis, R. C.: Fine Particulate Matter (PM) and Organic Speciation of Fireplace Emissions, Environ. Sci. Technol., 34(9), 1653-1658, doi10.1021/es981006f, 2000.

Rivera-Rios, J. C., Nguyen, T. B., Crounse, J. D., Jud, W., St. Clair, J. M., Mikoviny, T., Gilman, J. B., Lerner, B. M., Kaiser, J. B., de Gouw, J.,Wisthaler, A, Hansel, A., Wennberg, P. O., Seinfeld, J. H. and Keutsch, F. N.: Conversion of hydroperoxides to carbonyls in field and laboratory instrumentation: Observational biais in diagnosing pristine versus anthropogenically controlled atmospheric chemistry, Geophys. Res. Lett., 41, 8645-8651, doi:10.1002/2014GL061919, 2014.

Roukos, J., Plaisance, H., Leonardis, T., Bates, M. and Locoge, N.: Development and validation of an automated monitoring system for oxygenated volatile organic compounds and nitrile compounds in ambient air, J. Chromatogr. A., 1216(49), 8642-8651, doi:10.1016/j.chroma.2009.10.018, 2009.

Sandradewi, J., Prévôt, A. S., Szidat, S., Perron, N., Alfarra, M. R., Lanz, V. A., Weingartner, E. and Baltensperger, U.: Using aerosol light absorption measurements for the quantitative determination of wood burning and traffic emission contributors to particulate matter, Environ. Sci. Technol., 42 (9), 3316-3323, doi:10.1021/es702253m, 2008.

Sauvage, S., Plaisance, H., Locoge, N., Wroblewski, A., Coddeville, P., Galloo, J.- C.: Long term measurement and source apportionment of non-methane hydrocarbons in three French rural areas, Atmos. Environ., 43, 2430-2441, doi:10.1016/j.atmosenv.2009.02.001, 2009.

Sciare, J., d'Argouges, O., Zhang, Q. J., Sarda-Estève, R., Gaimoz, C., Gros, V., Beekmann, M. and Sanchez, O.: Comparison between simulated and observed chemical composition of fine aerosols in Paris (France) during springtime: contribution of regional versus continental emissions, Atmos. Chem. Phys., 10, 11987-12004, doi:10.5194/acp-10-11987-2010, 2010.

Sciare, J., d'Argouges, O., Sarda-Estève, R., Gaimoz, C., Dolgorouky, C., Bonnaire, N., Favez, O., Bonsang, B. and Gros, V.: Large contribution of water-insoluble secondary organic aerosols in the region of Paris (France) during wintertime, J. Geophys. Res. Atmospheres, 116(D22),n/a–n/a, doi:10.1029/2011JD015756, 2011.

Seco, R., Peñuelas, J., Filella, I., Llusia, J., Schallhart, S., Metzger, A., Müller, M. and Hansel, A.: Volatile organic compounds in the western Mediterranean basin: urban and rural winter measurements during the DAURE campaign, Atmos. Chem. Phys., 13, 4291-4306, doi:10.5194/acp-13-4291-2013, 2013.

Shao, M., Lu, S., Liu, Y., Xie, X., Chang, C., Huang, S. and Chen, Z.: Volatile organic compounds measured in summer in Beijing and their role in ground-level ozone formation, J. Geophys. Res.-Atmos., 114, D00G06, doi:10.1029/2008JD010863, 2009.

Sillman, S.: The relation between ozone, NO_x and hydrocarbons in urban and polluted rural environments, Atmos. Environ., 33, 1821-1845, doi:10.1016/S1352-2310(98)00345-8, 1999.

Skyllakou, K., Murphy, B. N., Megaritis, A. G., Fountoukis, C. and Pandis, S. N.: Contributions of local and regional sources to fine PM in the megacity of Paris, Atmos. Chem. Phys., 14, 2343-2352, doi:10.5194/acp-14-2343-2014, 2014.

Solomon, S. J., Custer, T., Schade, G., Dias, A. P. S. and Burrows, J.: Atmospheric methanol measurement using selective catalytic methanol to formaldehyde conversion, Atmos. Chem. Phys., 5, 2787-2796, doi:10.5194/acp-5-2787-2005, 2005.

Stein, A. F., Draxler, R. R., Rolph, G. D., Stunder, B. J. B., Cohen, M. D. and Ngan, F.: NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System, Bull. Am. Meteorol. Soc., 96(12), 2059-2077, doi:10.1175/BAMS-D-14-00110.1, 2015.

Steiner, A. H. and Goldstein, A. L.: Biogenic VOCs. Volatile Organic Compounds in the Atmosphere, 82-128, 2007.

Touaty, M. and Bonsang, B.: Hydrocarbon emissions in a highway tunnel in the Paris area, Atmos. Environ., 34, 985-996, doi:10.1016/S1352-2310(99)00195-8, 2000.

Tsai, P.- J., Shieh, H.- Y., Lee, W.- J. and Lai, S.- O.: Health-risk assessment for workers exposed to Polycyclic Aromatic Hydrocarbons (PAHs) in a carbon black manufacturing industry, Science of the Total Environment, 278, 137-150, doi:10.1016/S0048-9697(01)00643-X, 2001.

Tsigaridis, K. and Kanakidou, M.: Global modelling of secondary organic aerosol in the troposphere: a sensitivity analysis, Atmos. Chem. Phys., 3, 1849-1869, doi:10.5194/acp-3-1849-2003, 2003.

United Nations, World Urbanization Prospects: The 2014 Revision, United Nations publication, ST/ESA/SER.A/352, available at: http://esa.un.org/unpd/wup/Highlights/WUP2014-Highlights.pdf, (last access: October 2015), 2014.

Waked, A., Sauvage, S, Borbon, A., Gauduin, J., Pallares, C., Vagnot, M.-P., Léonardis, T. and Locoge, N.: Multi-year levels and trends of non-methane hydrocarbon concentrations observed in ambient air in France, Atmos. Environ., 141, 263-275, doi:10.1016/j.atmosenv.2016.06.59, 2016.

Weingartner, E., Saathoff, H., Schnaiter, M., Streit, N., Bitnar, B. and Baltensperger, U.: Absorption of light by soot particles: determination of the absorption coefficient by means of aethalometers, J. Aerosol Sci., 34(10), 1445-1463, doi:10.1016/S0021-8502(03)00359-8, 2003.

Yurdakul, S., Civan, M. and Gürdal, T.: Volatile organic compounds in suburban Ankara atmosphere, Turkey: Sources and variability, Atmospheric Research, 120-121, 298-311, doi:10.1016/j.atmosres.2012.09.015, 2013.

Zheng, J., Garzon, J. P., Huertas, M. E., Zhang, R., Levy, M., Ma, Y., Huertas, J. I., Jardon, R. T., Ruiz, L. G., Tan, H. and Molina, L. T.: Volatile organic compounds in Tijuana during the Cal-Mex 2010 campaign: Measurements and source apportionment, Atmos. Environ., 70, 521-531, doi:10.1016/j.atmosenv.2012.11.030, 2013.

Table 1. Statistical summaries (µg m^{-3}) of selected VOC concentrations measured at urban background sites. Statistics were calculated based on hourly mean data, initially obtained every 30 mn (ethane > isoprene) and every 5 to 10 mn (for aromatics and OVOCs). These measurements were undertaken from 15 January to 22 November 2010 (~ 10 months). A conversion factor is provided here to convert VOC concentrations (µg m^{-3}) into (ppb) mixing ratios.

	Species	Conversion Factor	Minimum	25th Percentile	Median	Mean	75th Percentile	Maximum	σ
ALKANES	Ethane	1.25	0.83	3.07	4.14	4.56	5.42	26.31	2.26
	Propane	1.83	0.23	1.63	2.44	2.78	3.45	25.64	1.80
	Iso-butane	2.42	0.23	1.12	1.58	1.96	2.30	23.52	1.51
	N-butane	2.42	0.40	1.88	2.69	3.35	3.93	56.10	2.88
	Iso-pentane	3.00	0.25	1.25	1.82	2.24	2.68	25.81	1.65
	N-pentane	3.00	0.10	0.58	0.85	1.04	1.28	12.04	0.76
	N-hexane	3.58	0.06	0.27	0.40	0.49	0.59	4.25	0.34
ALKENES	Ethylene	1.17	0.04	0.81	1.25	1.55	1.92	14.04	1.14
	Propene	1.75	0.09	0.37	0.53	0.64	0.78	5.93)	0.44
ALKYNE	Acetylene	1.08	0.04	0.29	0.48	0.68	0.81	7.39	0.64
DIENE	Isoprene	2.83	0.08	0.13	0.19	0.26	0.31	1.74	0.22
AROMATICS	Benzene	3.25	0.04	0.62	0.89	1.05	1.26	7.60	0.66
	Toluene	3.83	0.12	1.79	2.46	3.29	3.68	34.56	2.86
	Xylenes + C_8	4.42	0.26	1.58	2.19	2.76	3.25	21.84	1.97
ALCOHOL	Methanol	1.33	0.86	3.66	4.83	5.89	6.83	39.29	3.89
NITRILE	Acetonitrile	1.71	0.10	0.30	0.40	0.71	0.67	31.87	1.09
ALDEHYDE	Acetaldehyde	1.83	0.54	2.02	2.71	3.17	3.71	15.04	1.83
KETONE	Acetone	2.42	0.73	3.05	4.33	4.87	5.79	22.24	2.64
ENONE	MVK + MACR	2.92	0.05	0.30	0.48	0.65	0.77	6.27	0.59

Table 2. Comparison of mean concentrations of selected VOCs (measured by PTR-MS) with ambient levels observed in the literature from different urban atmospheres. All average values are reported in ppb.

VOCs measured by PTR-MS (*m/z*)	Paris[a] (*P.*) Jan.-Nov. (*Spring*) 2010	Paris[b] Spring 2007	Barcelona[c,1] Winter 2009	London[d] (*P.*) October 2006 (*2010*)	Mohali[e] (*P.*) May 2012 (*2010*)	Mexico City[f] (*P.*) March 2006 (*2010*)	Beijing[g] (*P.*) August 2005 (*2010*)	Houston[h] (*P.*) Aug.-Sept. 2000 (*2010*)
Methanol (*33.0*)	4.5 (*6.6*)	5.9	NA	NA (*3.3*)	38 (*5.3*)	NA (*1.6*)	11.7 (*2.8*)	10.8 (*3.9*)
Acetonitrile (*42.0*)	0.7 (*1.2*)	0.4	0.2–0.5	0.3 (*0.2*)	1.4 (*0.5*)	0.3–1.4 (*0.2*)	NA (*0.3*)	0.5 (*0.5*)
Acetaldehyde (*45.0*)	1.8 (*1.9*)	1.4	0.8–1.7	3.6 (*1.5*)	6.7 (*1.7*)	3.0–12.0 (*1.1*)	3.6 (*1.1*)	3.4 (*1.5*)
Acetone (*59.0*)	2.1 (*2.5*)	3.0	1.1–1.6	1.6 (*2.2*)	5.9 (*2.1*)	NA (*1.7*)	4.4 (*1.6*)	4.0 (*2.1*)
MVK + MACR + ISOPOOHs (*71.0*)	0.2 (*0.2*)	0.3	0.07–0.12	NA (*0.2*)	NA (*0.1*)	NA (*0.1*)	0.3–0.6 (*0.2*)	0.8 (*0.2*)
Benzene (*79.0*)	0.3 (*0.3*)	0.4	0.2–0.6	0.1 (*0.4*)	1.7 (*0.3*)	NA (*0.3*)	NA (*0.2*)	0.6 (*0.3*)
Toluene (*93.0*)	0.9 (*0.9*)	1.4	0.8–2.7	1.9 (*0.9*)	2.7 (*0.8*)	3.0–28.0 (*0.6*)	1.0–4.0 (*0.5*)	0.8 (*0.9*)
Xylenes + C_8 (*107.0*)	0.6 (*0.6*)	1.3	0.9–3.4	0.2 (*0.7*)	2.0 (*0.7*)	NA (*0.4*)	NA (*0.4*)	0.6 (*0.6*)

[a] This study (*Values in brackets from VOC measurements performed during the same sampling period of the other urban studies are given for comparison*).

[b] Gros et al. (2011)

[c] Seco et al. (2013)

[d] Langford et al. (2010)

[e] Sinha et al. (2014)

[f] Fortner et al. (2009) – Values estimated from graphs

[g] Shao et al. (2009)

[h] Karl et al. (2003)

[1] A full comparison was not possible because no data was available between 16 February and 24 March 2010.

NA – Non Available data.

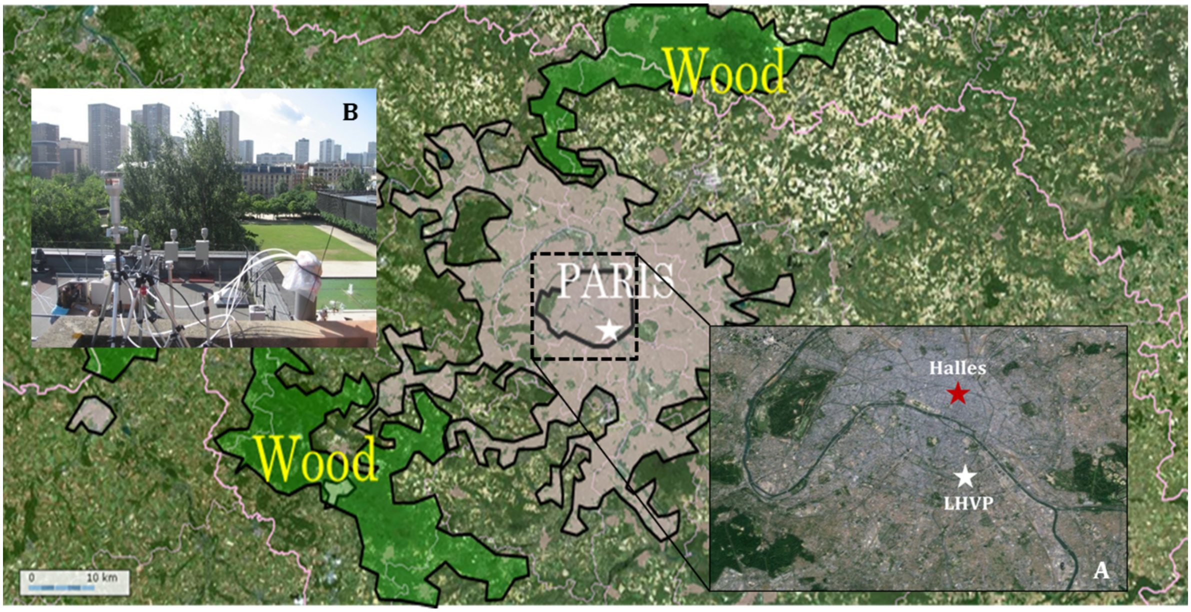


Figure 1. Maps of Paris and Île-de-France region. The picture A shows the location of the two main sampling sites in downtown Paris. The white and red stars locate the position of the LHVP laboratory and the AIRPARIF site, respectively. The picture B shows the terrace roof of LHVP.

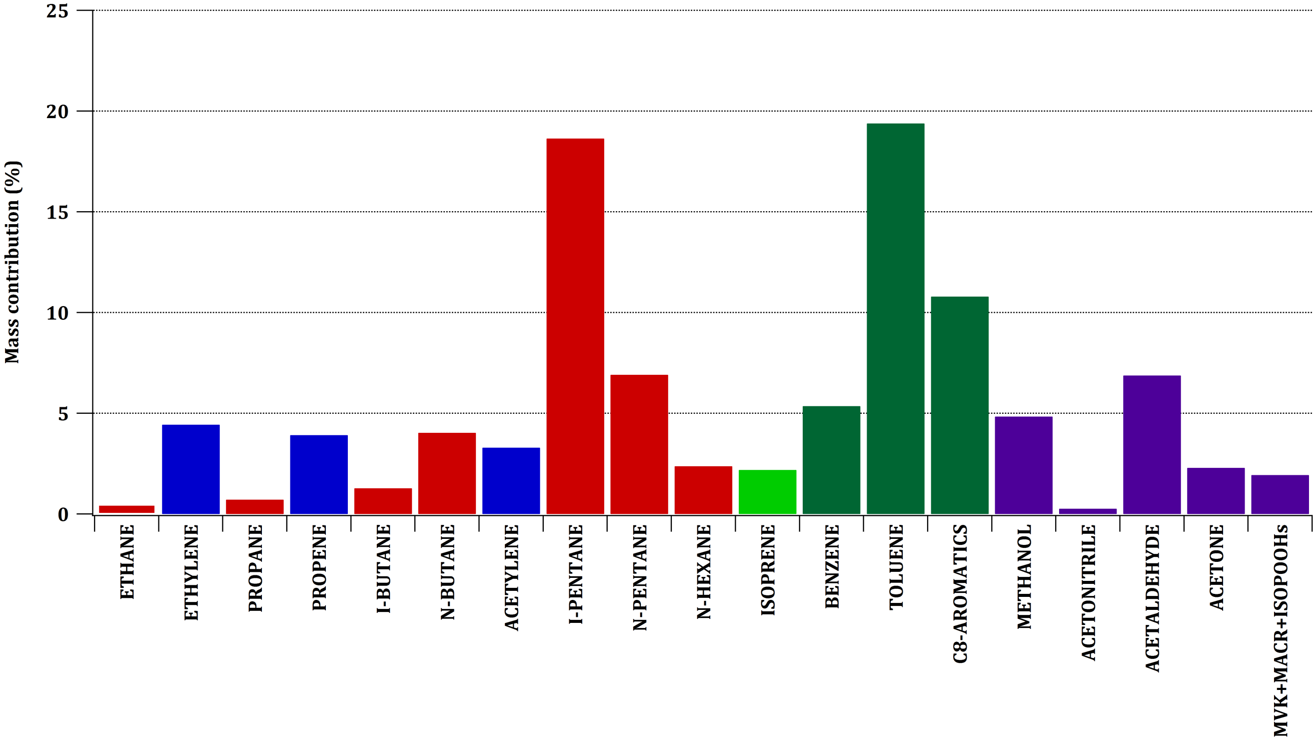


Figure 2. Average highway tunnel profile (in mass contribution, %) assessed from traffic peaks concentrations and subtracted from nighttime values. Red, blue, light/dark green and purple bars correspond to alkanes, alkenes-alkynes, isoprene/aromatics and oxygenated species, respectively.

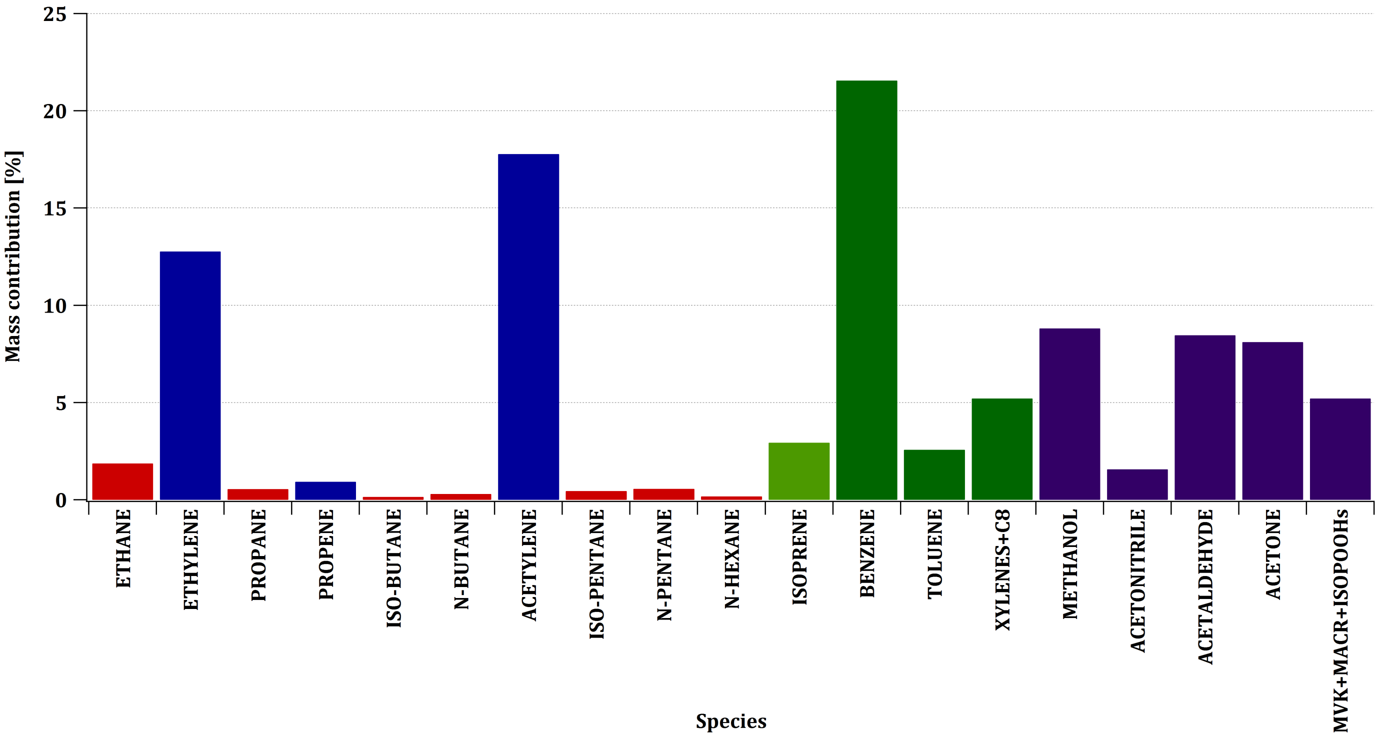


Figure 3. Average VOC fingerprint (in mass contribution, %) from domestic biomass burning obtained during the fireplace experiment. Red, blue, light/dark green and purple bars correspond to alkanes, alkenes-alkynes, isoprene/aromatics and oxygenated species, respectively.

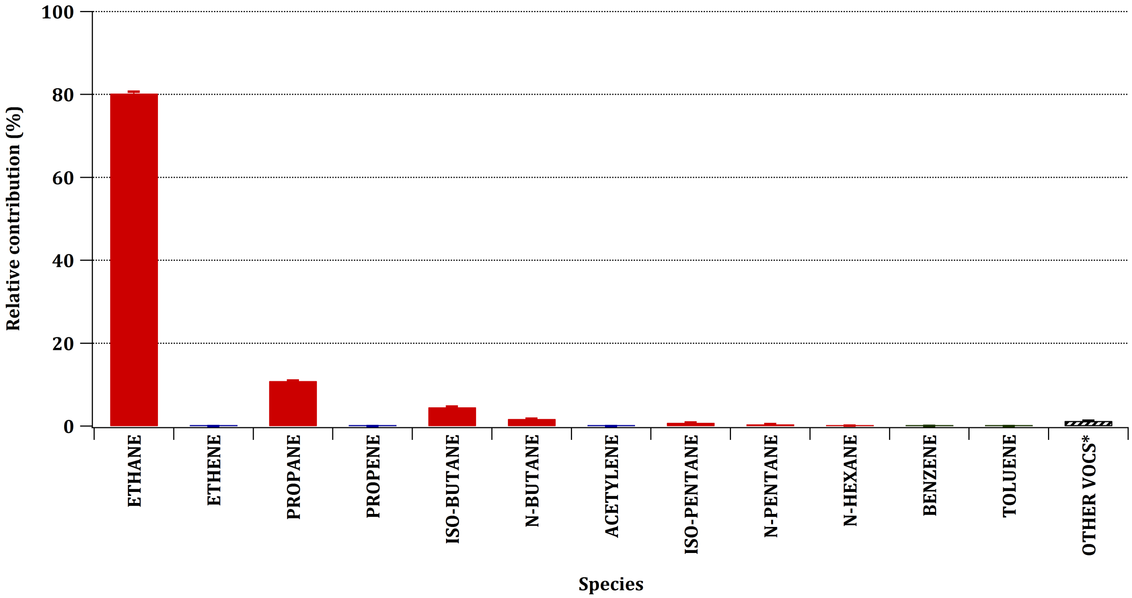
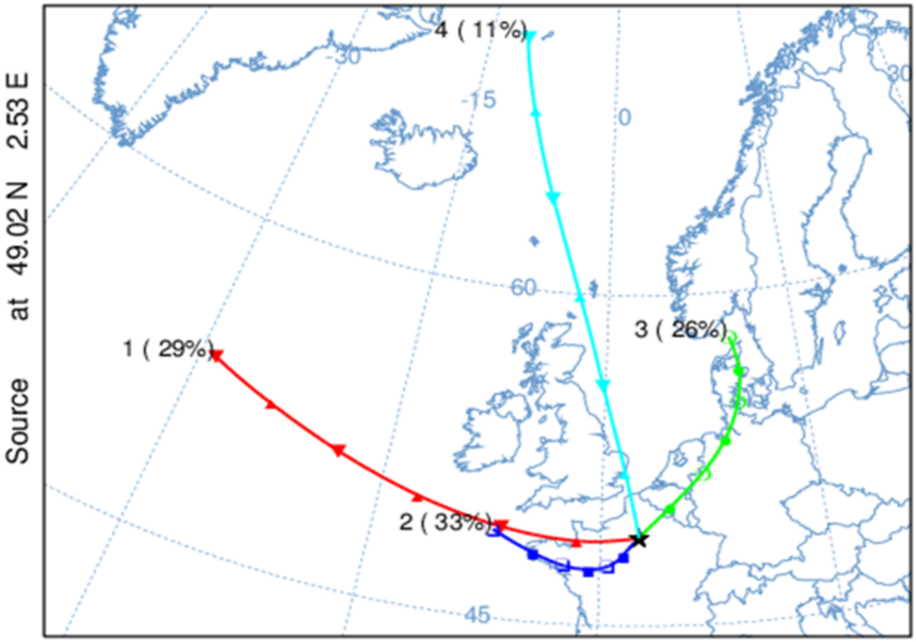


Figure 4. Average chemical composition of natural gas used in Paris. *Other VOCs include heavier alkanes (e.g. cyclopentane/hexane, dimethyl butanes) and butenes in lower proportions. Whiskers correspond to error bars (1σ).



Cluster CL#	Air mass origins	Year (J-N)	Winter	Spring	Summer	Autumn
CL#1	LRT (West)	29 %	24 %	25 %	32 %	32 %
CL#2	SRT (West)	33 %	29 %	25 %	41 %	33 %
CL#3	Continental influences (NE)	26 %	34 %	39 %	19 %	22 %
CL#4	North	11 %	13 %	11 %	8 %	12 %

LRT: Long Range Transport
SRT: Short Range Transport

Figure 5. Average trajectories obtained after clustering analysis and the relative proportion of clusters (%) over the year and per season.

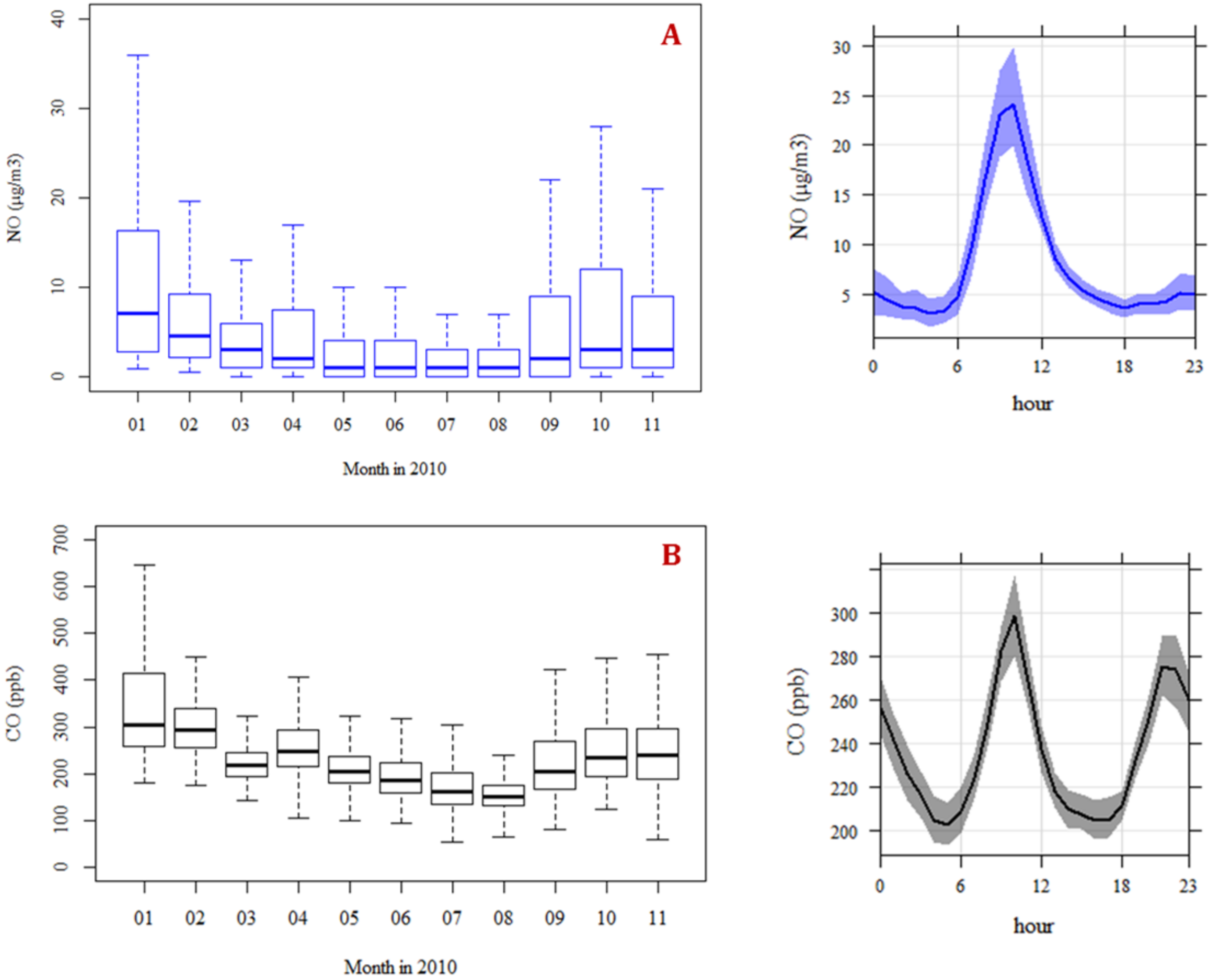
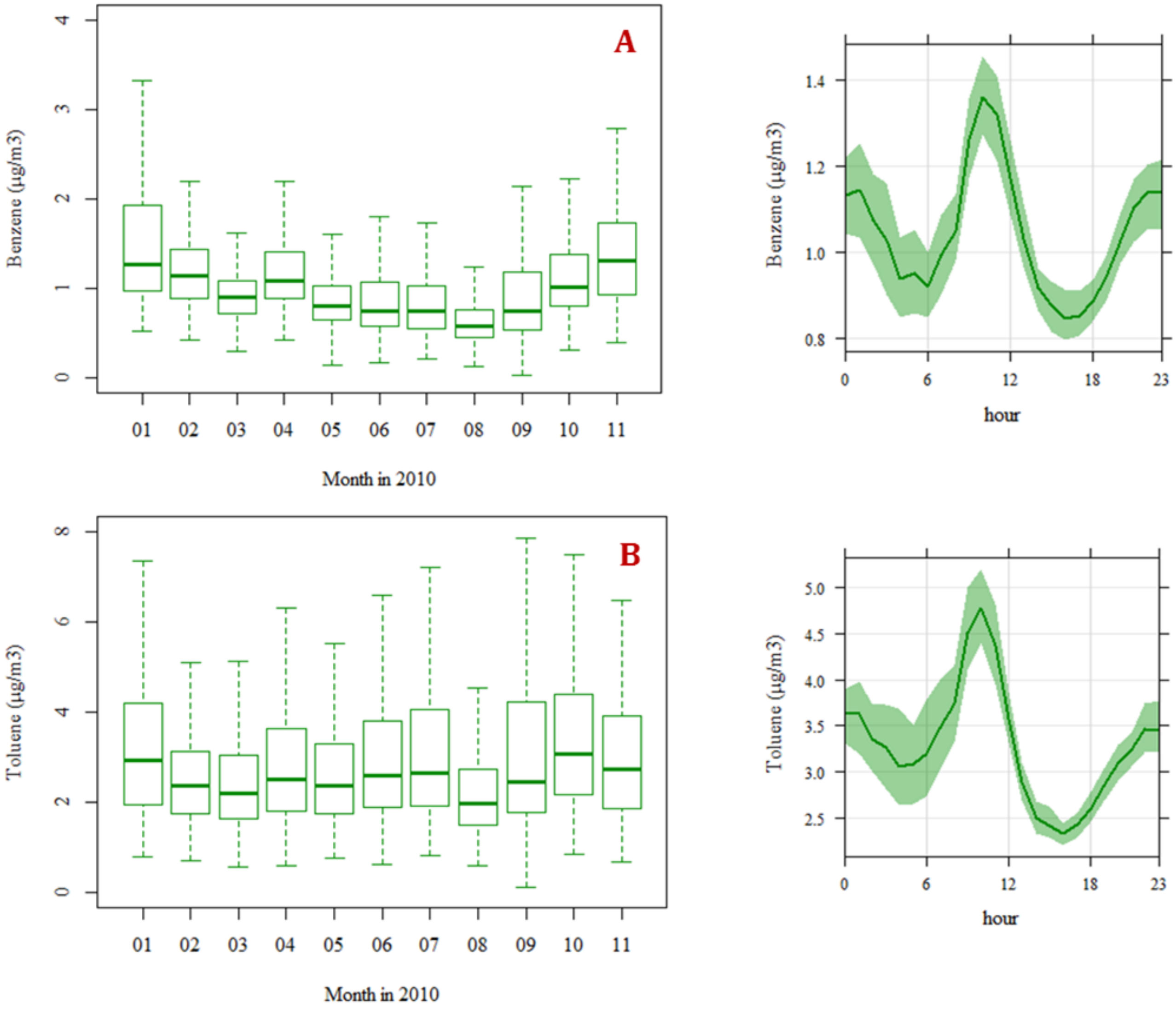
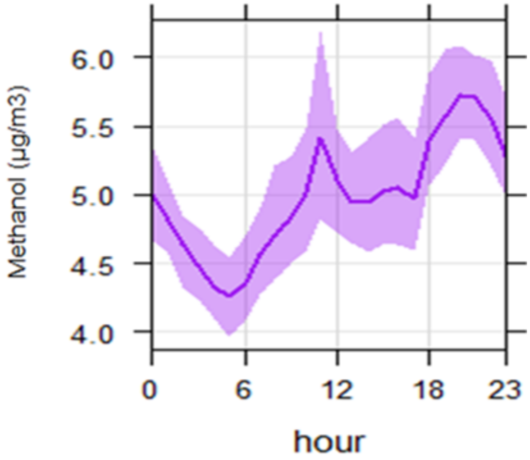


Figure 6. (Left) Monthly box and whisker plots of NO (A) and CO (B) expressed in µg m^{-3} and ppb, respectively. Solid lines represent the median concentration and the box shows the InterQuartile Range (IQR). The bottom and top of the box depict the 25th (the first quartile) and the 75th (the third quartile) percentile. The ends of the whiskers correspond to the lowest and highest data still within 1.5 times the IQR of Q1 and Q3, respectively.

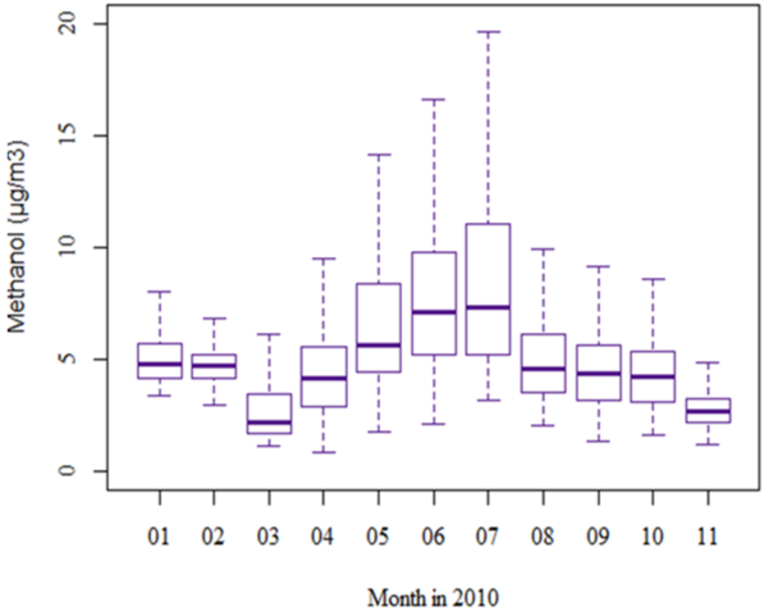
(Right) Diurnal variations of NO and CO averaged over the whole sampling period. Time is given as Local Time. Lines correspond to hourly means and shaded areas indicate the 95 % confidence intervals of the mean.

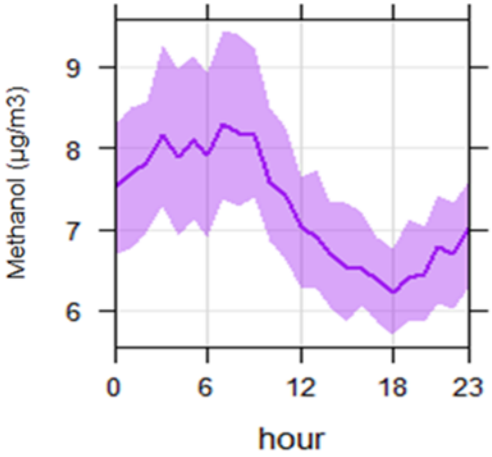
Figure 7. (Left) Monthly box and whisker plots of benzene (A), toluene (B), methanol (C), acetaldehyde (D), acetone (E) and MVK + MACR + ISOPOOHs (F) expressed as µg m^{-3}. Solid lines represent the median concentration and the box shows the InterQuartile Range (IQR). The bottom and top of the box depict the 25th (the first quartile) and the 75th (the third quartile) percentile. The ends of the whiskers correspond to the lowest and highest data still within 1.5 times the IQR of Q1 and Q3, respectively. (Right) Diurnal variations of (O)VOCs averaged over the whole sampling period. Time is given as Local Time. Lines correspond to hourly means and shaded areas indicate the 95 % confidence intervals of the mean.



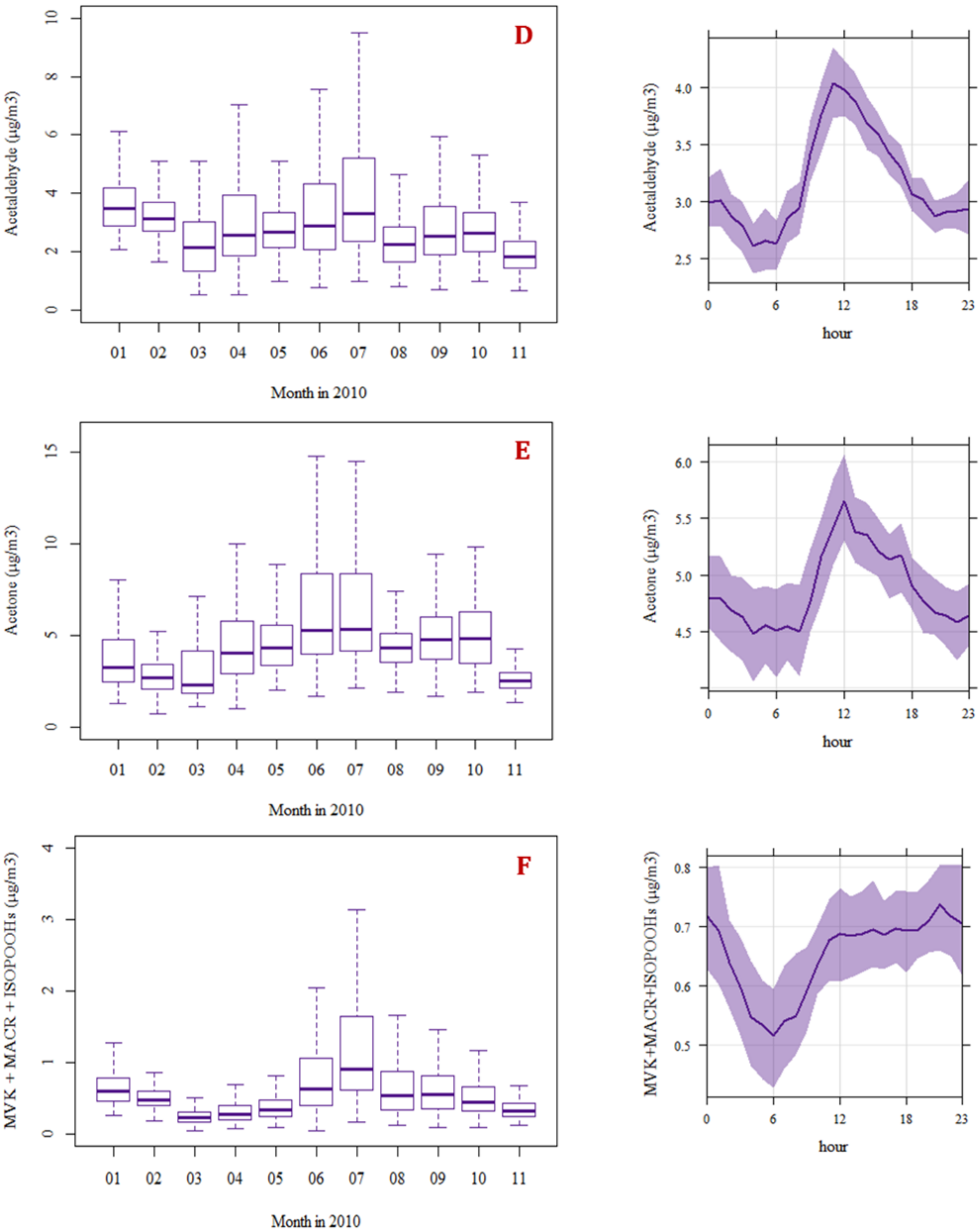


WINTER
(January – February)





SUMMER
(June – July - August)


D
Acetaldehyde (µg/m3)
0
2
4
6
8
10
01
02
03
04
05
06
07
08
09
10
11
Month in 2010
Acetaldehyde (µg/m3)
2.5
3.0
3.5
4.0
0
6
12
18
23
hour
E
Acetone (µg/m3)
0
5
10
15
01
02
03
04
05
06
07
08
09
10
11
Month in 2010
Acetone (µg/m3)
4.5
5.0
5.5
6.0
0
6
12
18
23
hour
F
MVK + MACR + ISOPOOHs (µg/m3)
0
1
2
3
4
01
02
03
04
05
06
07
08
09
10
11
Month in 2010
MVK+MACR+ISOPOOHs (µg/m3)
0.5
0.6
0.7
0.8
0
6
12
18
23
hour

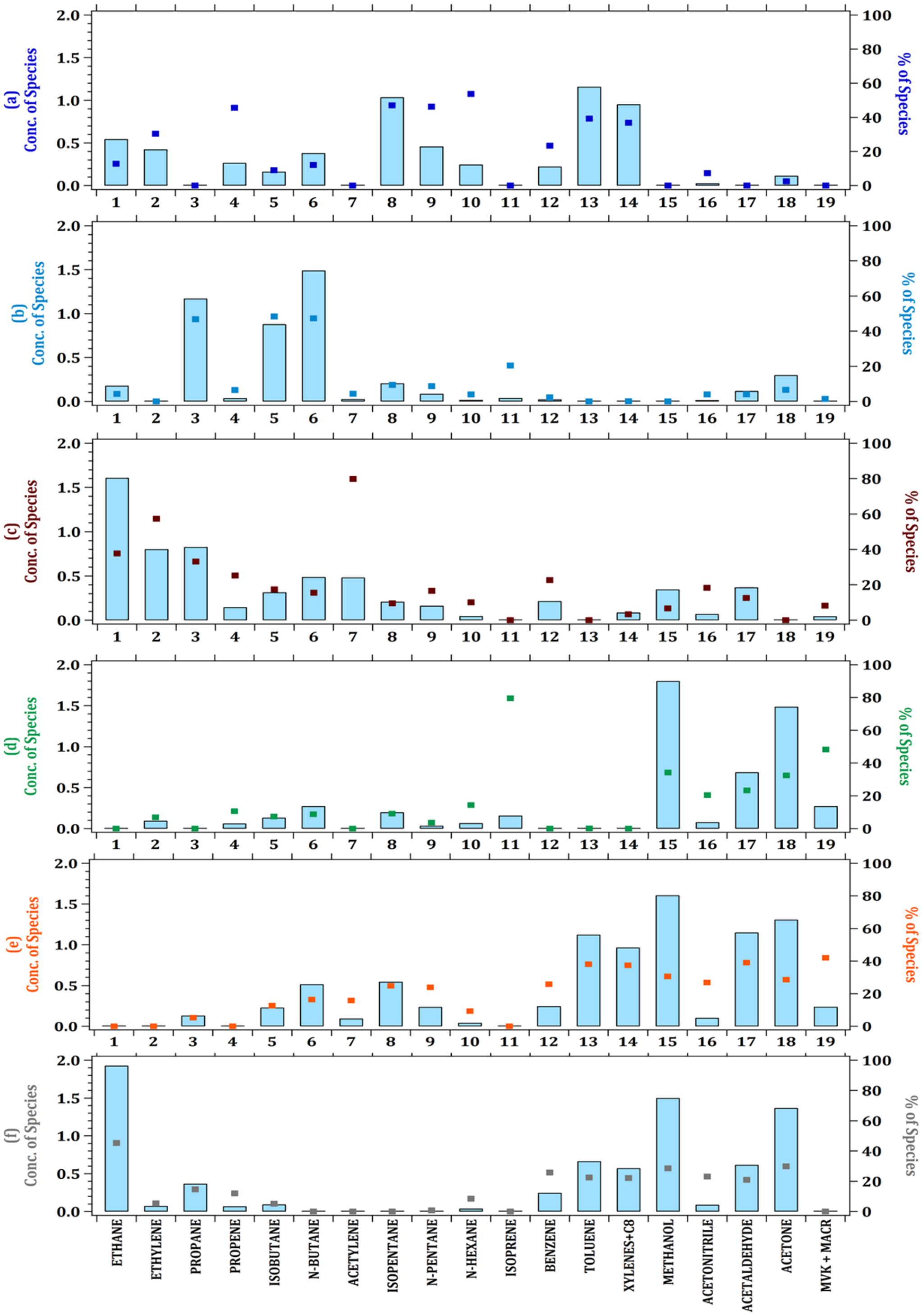


Figure 8. Source composition profiles of the 6-factor PMF solution. The concentrations (μg m^{-3}) and the percent of each species apportioned to the factor are displayed as a pale blue bar and a color box, respectively. **(a)** F1 – Motor Vehicle Exhaust ; **(b)** F2 – Evaporative sources ; **(c)** F3 – Wood Burning; **(d)** F4 – Biogenic ; **(e)** F5 – Solvents use ; **(f)** F6 – Natural Gas and Background.

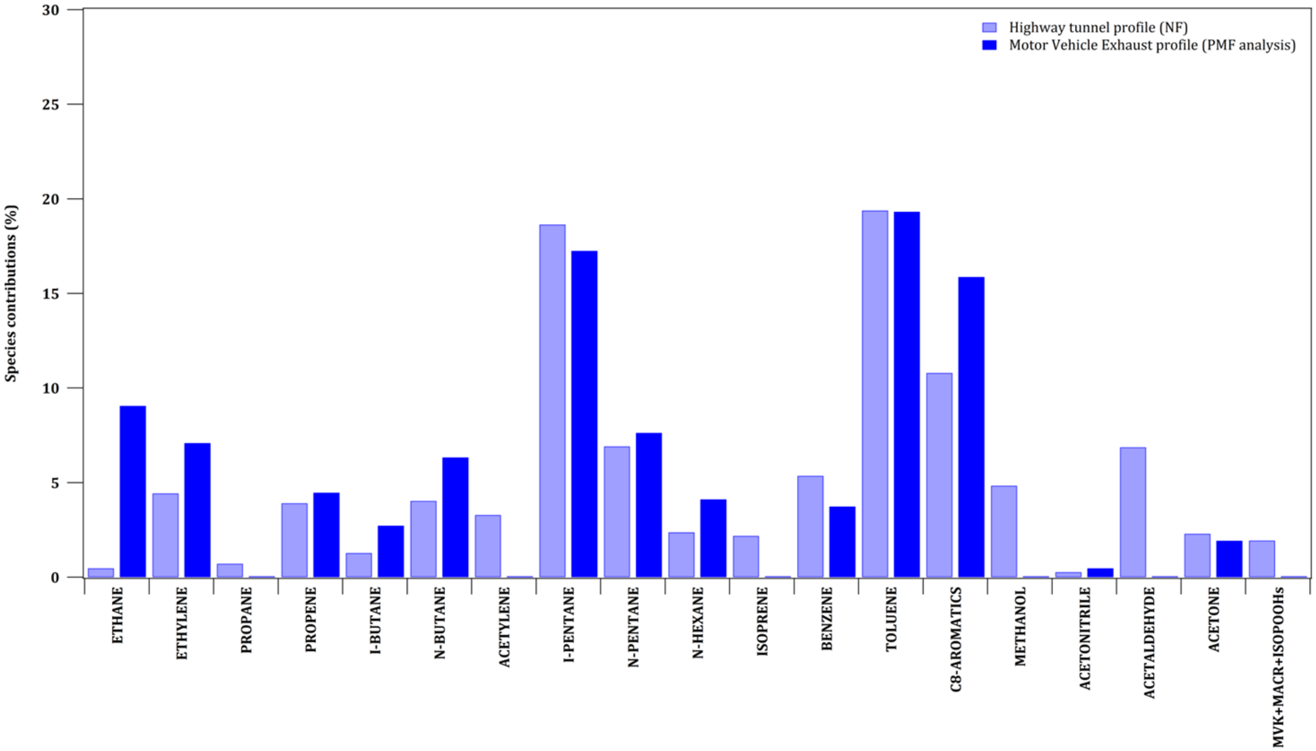
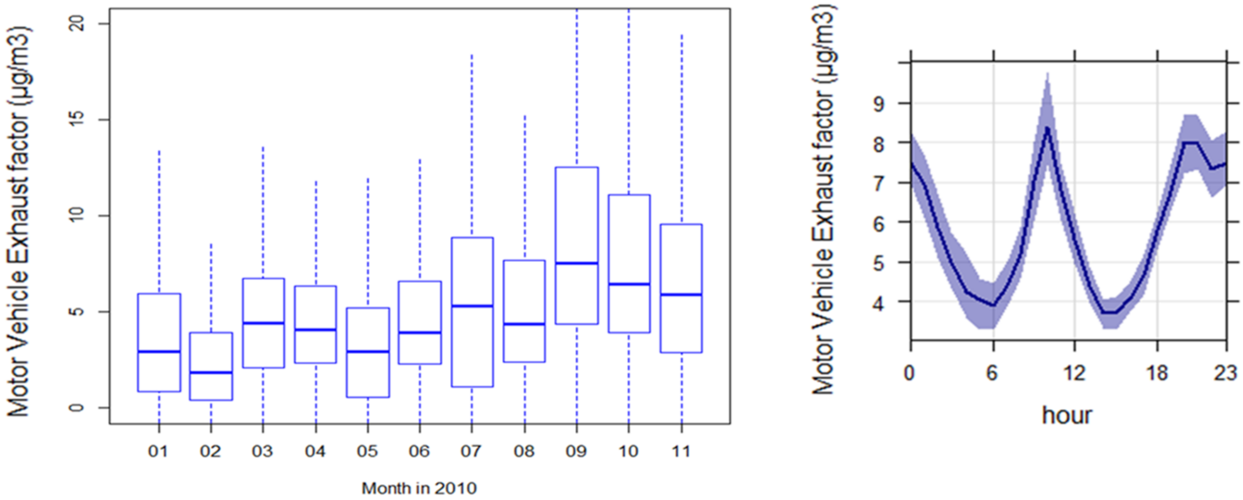
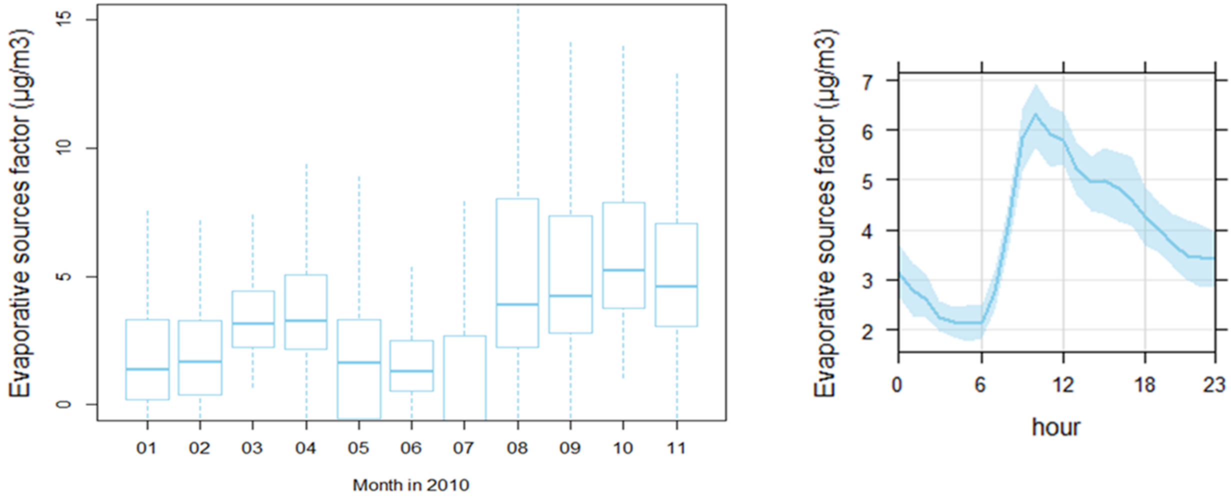


Figure 9. Comparison of speciated profiles issued from the highway tunnel experiment and PMF simulations (F1 - Motor Vehicle Exhaust). The species contributions are expressed in %. NF = Near-Field.

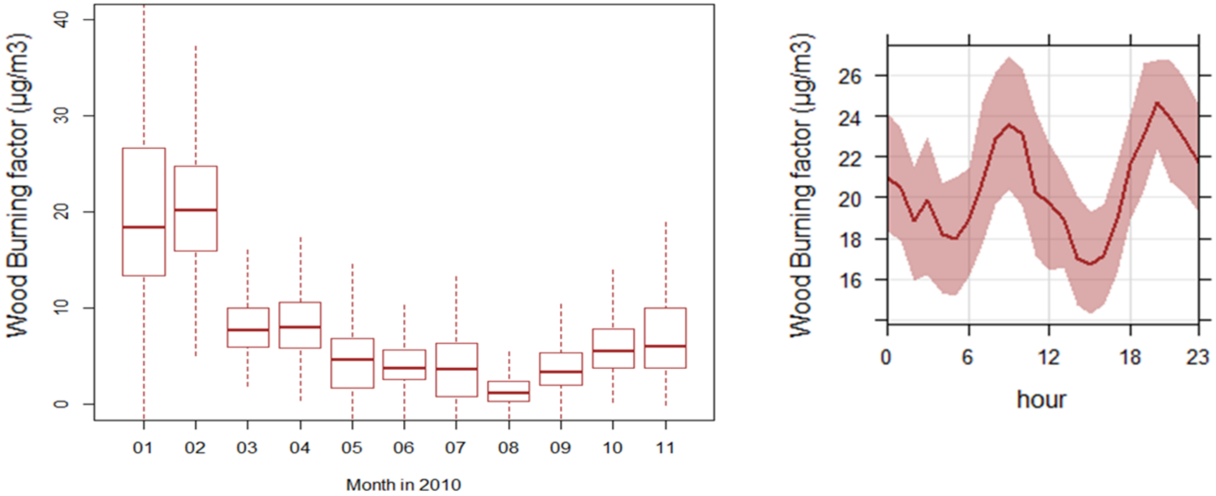
PANEL 1



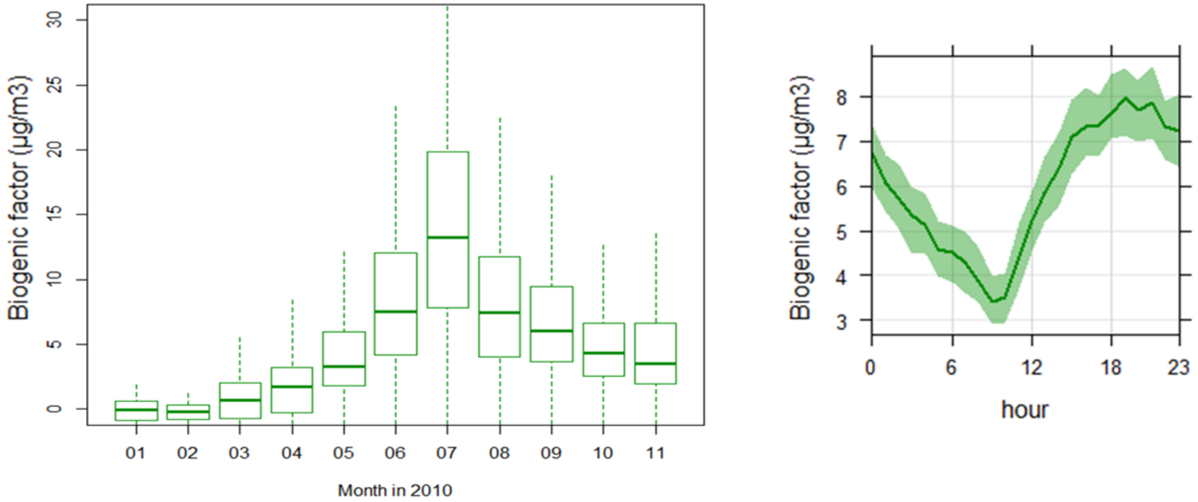
PANEL 2



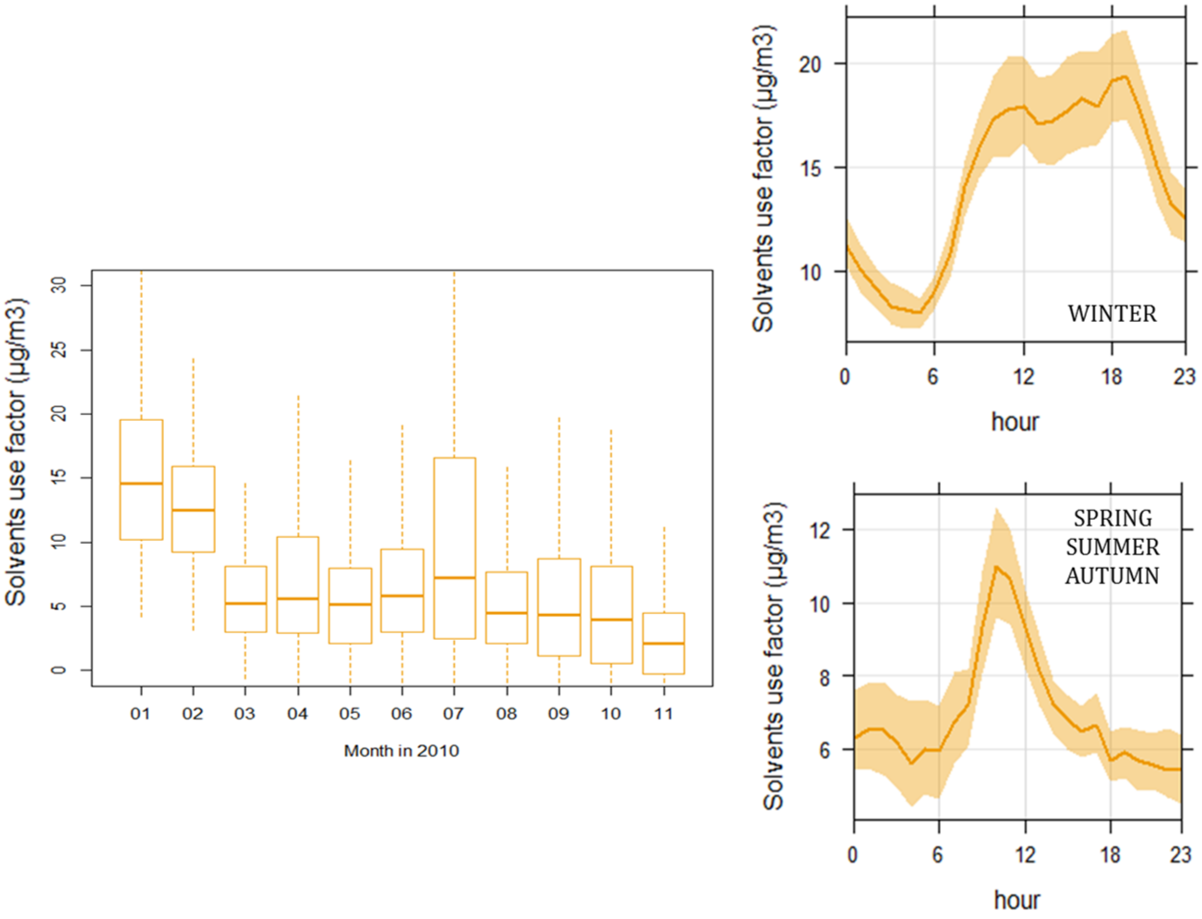
PANEL 3



PANEL 4



PANEL 5


Solvents use factor (µg/m3)
0
5
10
15
20
25
30
01
02
03
04
05
06
07
08
09
10
11
Month in 2010
Solvents use factor (µg/m3)
10
15
20
0
6
12
18
23
hour
WINTER
Solvents use factor (µg/m3)
6
8
10
12
0
6
12
18
23
hour
SPRING
SUMMER
AUTUMN

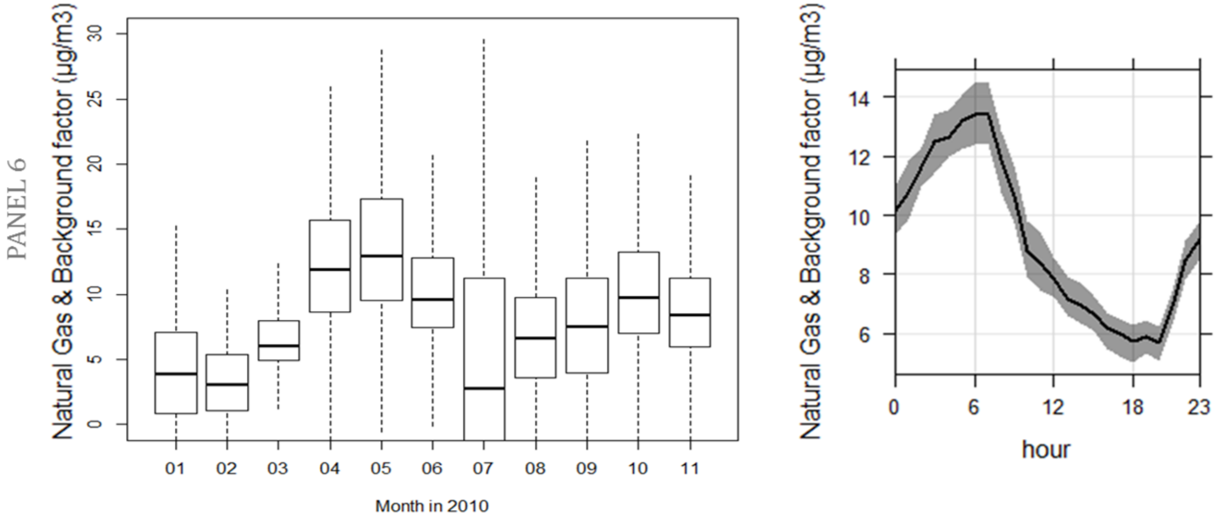


Figure 10. (Left) Monthly box and whisker plots of modeled sources from the 6-factor solution. Concentration levels are expressed in µg m^{-3}. Solid lines represent the median concentrations and the box shows the InterQuartile Range (IQR). The bottom and top of the box depict the 25th (the first quartile) and the 75th (the third quartile) percentile. The ends of the whiskers correspond to the lowest and highest data still within 1.5 times the IQR of Q1 and Q3, respectively. (Right) Diurnal variations of the resolved PMF factors. Time is given in Local Time. Lines correspond to hourly means and shaded areas indicate the 95 % confidence intervals of the mean.

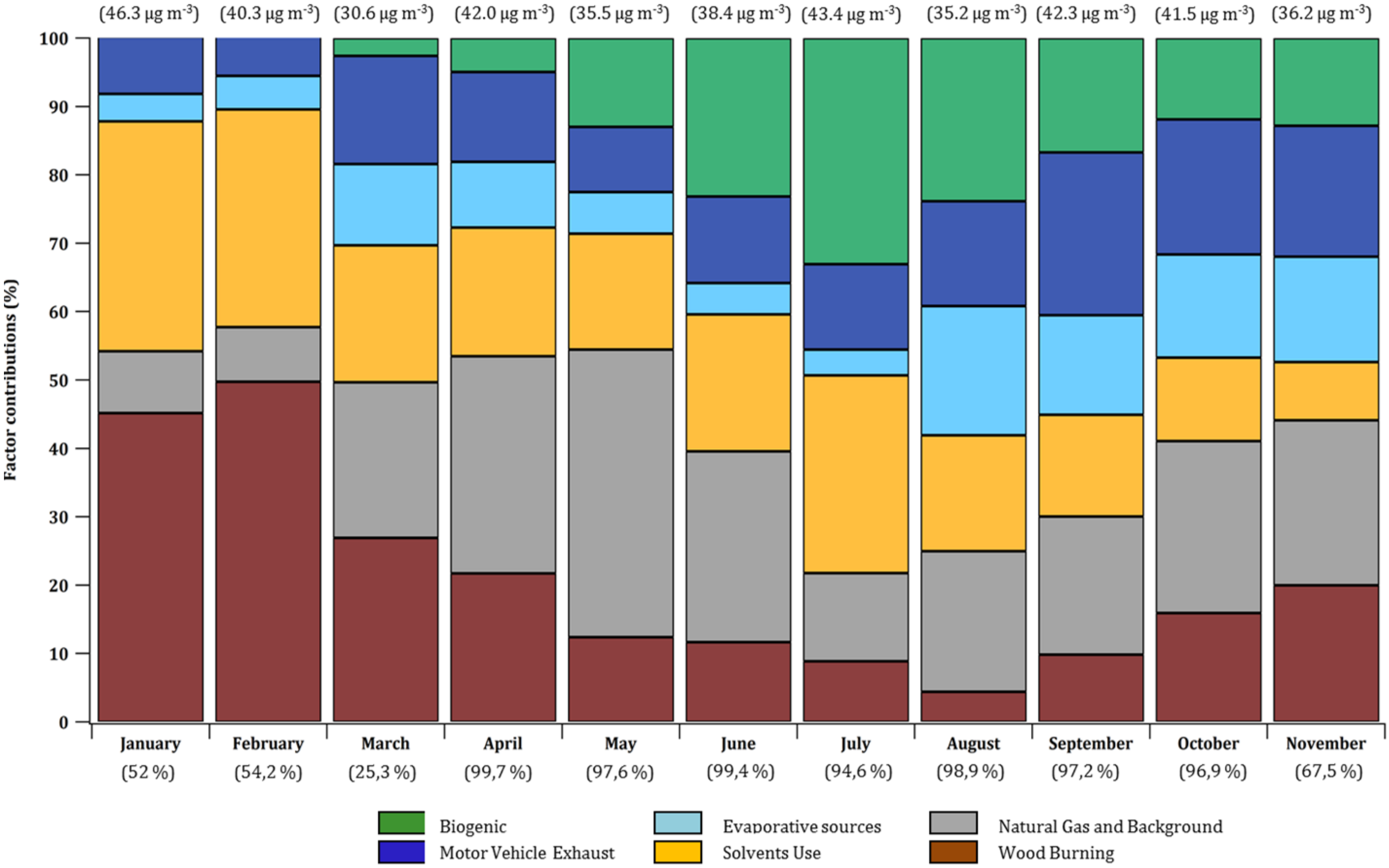


Figure 11. Variations of monthly averaged contributions of the six modeled VOC sources (expressed in %); (Top) - Average predicted VOC concentration levels per month (µg m^{-3}); (Bottom) - Completeness of the data per month (%).

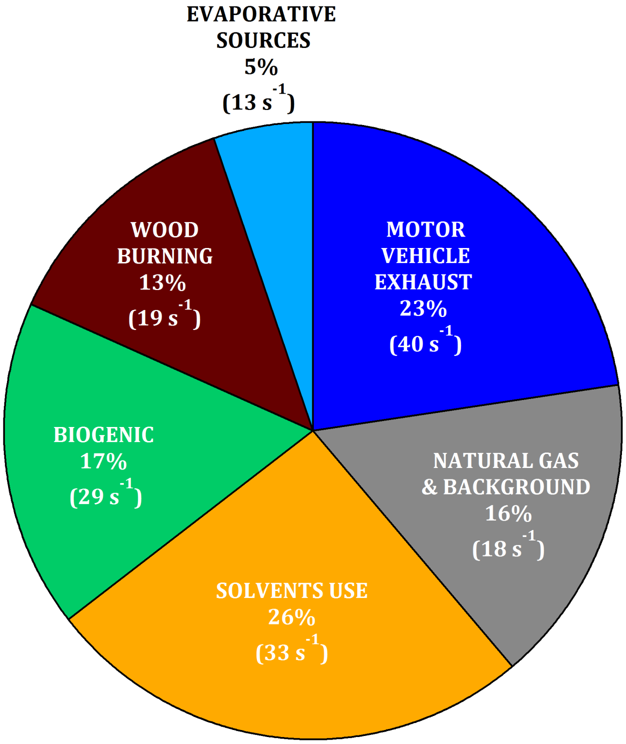


Figure 12. Relative and absolute contributions of reactivity of each PMF factors (% and s^{-1}, respectively).

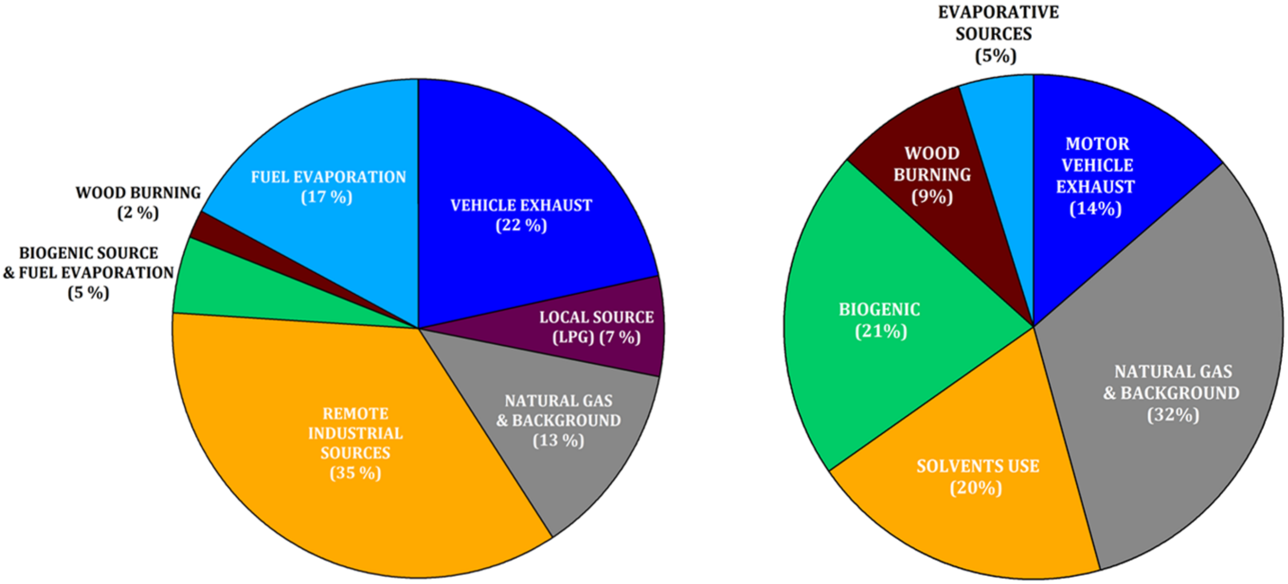


Figure 13. Relative contributions to the TVOC mass of 7 and 6 PMF sources identified from 25 May to 14 June 2007 (Gaimoz et al., 2011 - left pie chart) and 2010 (this study - right pie chart), respectively.

Appendix A: Application of the Positive Matrix Factorization (PMF) approach in source apportionment of VOCs in Paris.

A1 Data preparation

Initially, the EPA PMF 5.0 model requires two input datasets: one with the chemical species atmospheric concentrations for each observation point and another with either uncertainties values or parameters for calculating the associated uncertainty.

The initial chemical dataset contains a selection of 19 hydrocarbon species and masses (for a detailed overview, see list of compounds in the Sub-section 2.4.1) measured from 15 January to 22 November 2010. Non methane hydrocarbons (NMHCs) and aromatic/oxygenated species (OVOCs) were respectively monitored with GC-FID and PTR-MS instruments belonging to different partners involved during the MEGAPOLI and FRANCIPOL intensive field campaigns. Unfortunately, no intercomparison between these instruments was possible because there was approximately a one-month delay between both experiments. However, preliminary PMF modeling simulations were performed using only the FRANCIPOL dataset (24 March – 22 November). The results have shown similar source profiles (see Section S1), as those already described in this paper. Consequently, two datasets (corresponding to MEGAPOLI and FRANCIPOL ones, respectively) were considered to form a single one and use it as an input unified database for the final PMF analysis.

The uncertainty dataset was built upon the equation-based method described by Norris et al. (2014). It requires both Method Detection Limit (MDL, here in µg m^{-3}) and the analytical uncertainty (u, here in %) for each considered species. Two sets of MDLs were used, one for each measurement campaign. Slight differences among species MDLs were found for *n*-hexane, aromatics, acetaldehyde and MVK+MACR+ISOPOOHs between both experiments. Of these VOCs, MDLs from the FRANCIPOL campaign were chosen for representativity reasons (as the corresponding dataset represents ~ 88 % of the total data matrix) to keep consistency in uncertainty calculations. The analytical uncertainties were respectively estimated at 15 % and 20 % and kept constant over the experiments.

The PMF uncertainty (σ) is therefore calculated as follows:

$$If\ X_{ij} \leq MDL,\ \forall j;\ X_{ij} = \frac{MDL}{2}\ and\ \sigma_{ij} = \frac{5}{6} \times MDL \tag{A1}$$

$$If\ X_{ij} \geq MDL,\ \forall j;\ X_{ij}\ does\ not\ change\ and\ \sigma_{ij} = \sqrt{(ErrorFraction\ u\ \times\ X_{ij})^2 + (0.5 \times MDL)^2} \tag{A2}$$

MDLs and analytical uncertainties (u) for each VOC are reported in Table A.1.

A2 Estimation of the number of PMF factors (p)

The accurate number of PMF factors (p values) in models must be ultimately estimated by the user using several exploratory means. Specific parameters were used to determine the appropriate p value such as the assessment of Q values, scaled residuals, predicted *versus* observed concentrations interpretation and the physical meaning of factor profiles.

Eight different modeling conditions were examined with p values ranging from 3 to 10, each simulation being randomly conducted 20 times. The reviewing of the IS (the maximum individual standard deviation) parameter highlighted a slope failure for $p = 5$ whereas the IM (the maximum individual column mean) indicator reported another slope failure for $p = 6$. Choosing less factors, $p < 6$, concatenated three source profiles (attributed to solvents use, natural gas and background emissions, respectively) into a factor, whereas choosing $p = 6$ allowed splitting one of them. Opting for $p > 6$ did not provide any supplemental physical meaningfulness to existing profiles. The investigation of r^2 from ten modeled solutions also reported a slope failure for $p = 6$. In addition, only *Qtrue* / *Qexpected* value for $p = 6$ was closer than 1.0 (e.g. 0.94 in comparison with 1.12 for $p = 5$ and 0.8 for $p = 7$), thus suggesting that the 6-factor configuration is supposed to be the most optimum solution for this PMF analysis. Finally, this configuration was investigated over all the details. Usually, PMF identifies the best solution by the lowest *Qrobust* value (e.g. the mimimum Q). Within this analysis, its corresponding PMF solution was not considered due to a lack of physical significance for one factor profile (e.g. solvents). Therefore, another PMF solution closest to the selected *Qrobust* value was subsequently examined and chosen in terms of interpretability and fitting scores.

A3 Robustness of PMF results

Further technical and mathematical indicators regarding to the 6-factor configuration are reported here to assess the robustness and the quality of the final PMF solution. Firstly, the ratio between *Qrobust* and *Qtrue* reached around 1.0, thus indicating that the modeled results were not biased by peak events. Almost 100 % of the scaled residuals were within $\pm 3\sigma$ and were normally distributed for all species. In addition, the Kolmogorov-Smirnoff (KS) test granted a KS p-value very close to zero, thus illustrating a statistically significant test with a α risk of 5 %. The correlation between total VOC reconstructed concentrations from all the factors with total VOC observed concentrations is depicted in Figure A.1. With R^2 very close to 0.9, almost all variance in the total concentration of the 19 VOCs can be explained by the PMF model.

Almost all the chemical species also displayed good determination coefficients (r^2 higher than 0.6 for 15 compounds) between predicted and observed concentrations, with the exception of propane and n-hexane showing a fairly reasonable coefficient between 0.5 and 0.6 (due to their 31 % and 36 % missing values, respectively). Isoprene and acetonitrile exhibited bad r^2 values (0.29 and 0.06, respectively) due to either a relatively high number of missing values or a “weak” additional error, for which sample uncertainties were tripled. Slopes were close to 1.0 for most species (higher than 0.6 for 17 VOCs), excepted for isoprene (0.5) and acetonitrile (0.02). The limitations of the PMF model to simulate isoprene and acetonitrile have therefore been kept in mind within the reconstructed results description and discussions.

A4 Estimation of model prediction uncertainties

PMF output uncertainties can be estimated using the Error Estimation options starting with DISP (dQ-controlled displacement of factor elements) and processing to BS (Classical bootstrap). These two uncertainty methods are designed to provide key information on the stability and the precision of the chosen PMF solution (Paatero et al., 2014).

The DISP, Base Model Displacement Error Estimation, assesses the rotational ambiguity of the PMF solution by exploring intervals (minimum and maximum) of source profile values. During the DISP, a minimum Q value is newly calculated (based on the adjustment up and down in factor profile values) and compared with the unadjusted solution Q value. The difference between the initial Q value and the modified Q value (so-called dQ) should be lower than dQmax value, for which four levels (values: 4, 8, 15 and 25) were taken into account. For each dQmax value, 120 intervals were estimated. The DISP analysis results were considered validated: no error could be detected and no drop of Q was observed. As no swap occurred, the PMF solution was considered sufficiently robust to be used.

The BS, Base Model Bootstrap Error Estimation, is also used to evaluate the reproducibility of the PMF solution, with a specific focus on the original sub-matrix F. A further description on the bootstrapping technique is presented in Norris et al. (2014) and in Paatero et al. (2014). A base model bootstrap method was then carried out, executing one hundred iterations, using a random seed, a block size of 874 samples (calculated according to the methodology of Politis and White (2004)) and a minimum Pearson correlation coefficient (R-Value) of 0.6. All factors were well reproduced through this technique over at least 88 % of runs, thus indicating that BS uncertainties can be interpreted and the number of factors may be appropriate. Consequently, 12 % of runs were redistributed into the different existing factors. No runs were unmapped. Finally, around 91 % of species with the base run profile value were identified within the InterQuartile Range (IQR, e.g. 25th - 75th percentile of bootstrap runs) for all factors considered.

Finally, the rotational ambiguity of this 6-factor PMF configuration was also investigated using the Fpeak parameter. Different Fpeak values from -5 to 5 were used to generate a more realistic PMF solution. The results from the nonzero Fpeak values were generally consistent with the runs associated with the zero Fpeak value (e.g. Base Model run), thus illustrating a low rotational ambiguity of the final PMF solution.

Table A.1. Method Detection Limits (MDLs) and Analytical Uncertainties (*u*) for each species used in PMF modeling simulations.

Species	MDL–MEGAPOLI µg m^{-3} (ppb)	MDL–FRANCIPOL µg m^{-3} (ppb)	*u* (%)
Ethane*	0.025 (0.020)	0.024 (0.019)	15
Ethylene*	0.023 (0.020)	0.024 (0.021)	15
Propane*	0.037 (0.020)	0.024 (0.013)	15
Propene*	0.035 (0.020)	0.024 (0.014)	15
Iso-butane*	0.048 (0.020)	0.024 (0.010)	15
N-butane*	0.048 (0.020)	0.024 (0.010)	15
Acetylene*	0.022 (0.020)	0.024 (0.022)	15
Iso-pentane*	0.060 (0.020)	0.024 (0.008)	15
N-pentane*	0.060 (0.020)	0.024 (0.008)	15
N-hexane*	0.013 (0.004)	0.013 (0.004)	15
Isoprene*	0.024 (0.008)	0.024 (0.008)	20
Benzene**	0.071 (0.022)	0.071 (0.022)	20
Toluene**	0.240 (0.063)	0.240 (0.063)	20
Xylenes + C_8**	0.259 (0.059)	0.259 (0.059)	20
Methanol**	0.317 (0.238)	0.330 (0.248)	20
Acetonitrile**	0.068 (0.040)	0.084 (0.049)	20
Acetaldehyde**	0.167 (0.091)	0.167 (0.091)	20
Acetone**	0.092 (0.038)	0.118 (0.049)	20
MVK[1] + MACR[2] + ISOPOOHs[3]**	0.020 (0.007)	0.020 (0.007)	20
Σ VOC	1.629 (0.749)	1.542 (0.723)	20

[1] MVK = Methylvinylketone

[2] MACR = Methacrolein

[3] ISOPOOHs = Isoprene hydroxy hydroperoxides (Rivera-Rios et al., 2014)

* Hydrocarbons measured using a GC-FID by LSCE (MEGAPOLI, LHVP) and AIRPARIF (FRANCIPOL, “Les Halles” subway station).

** Masses measured using a PTR-MS by LCP (MEGAPOLI, LHVP) and LSCE (FRANCIPOL, LHVP).

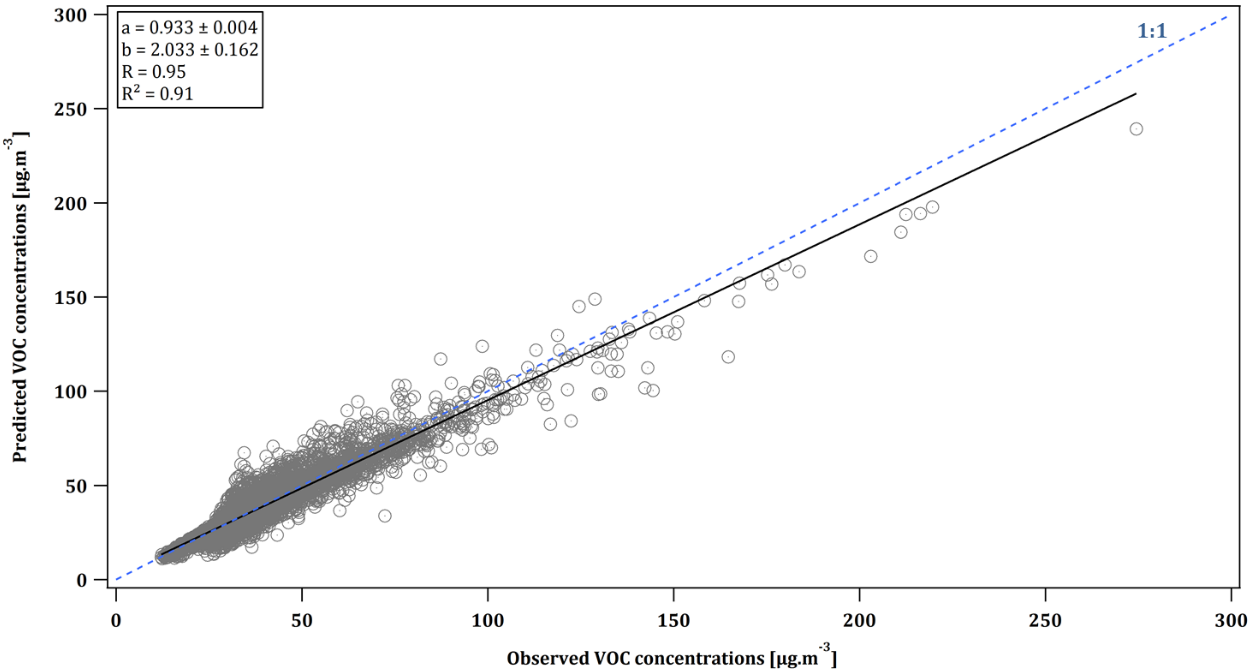


Figure A.1. Agreement between total predicted and observed VOC concentrations based on the 6-factor PMF solution.