

Evaluation of receptor and chemical transport models for PM₁₀ source apportionment

C.A. Belis^{a,*}, D. Pernigotti^a, G. Pirovano^b, O. Favez^c, J.L. Jaffrezo^d, J. Kuenen^e, H. Denier van Der Gon^e, M. Reizer^f, V. Riffault^g, L.Y. Alleman^g, M. Almeida^h, F. Amatoⁱ, A. Angyal^j, G. Argyropoulos^k, S. Bande^l, I. Beslic^m, J.-L. Besombesⁿ, M.C. Bove^{o,1}, P. Brotto^{o,2}, G. Calori^p, D. Cesari^q, C. Colombi^r, D. Contini^q, G. De Gennaro^s, A. Di Gilio^s, E. Diapouli^t, I. El Haddad^u, H. Elbern^v, K. Eleftheriadis^t, J. Ferreira^w, M. Garcia Vivanco^x, S. Gilardoni^y, B. Gollyⁿ, S. Hellebust^z, P.K. Hopke^{aa}, Y. Izadmanesh^{i,ab}, H. Jorquera^{ac}, K. Krajsek^v, R. Kranenburg^e, P. Lazzeri^{ad}, F. Lenartz^{ae}, F. Lucarelli^{af}, K. Maciejewska^f, A. Manders^e, M. Manousakas^t, M. Masiol^{aa}, M. Mircea^{ag}, D. Mooibroek^{ah}, S. Nava^{af}, D. Oliveira^{c,g}, M. Paglione^{y,3}, M. Pandolfiⁱ, M. Perrone^{ai,4}, E. Petralia^{ag}, A. Pietrodangelo^{aj}, S. Pillon^{ak}, P. Pokorna^{aa}, P. Prati^o, D. Salameh^d, C. Samara^k, L. Samek^{al}, D. Saraga^t, S. Sauvage^g, M. Schaap^e, F. Scotto^{am}, K. Segal^m, G. Siour^{an}, R. Taulerⁱ, G. Valli^{ao}, R. Vecchi^{ao}, E. Venturini^{ap}, M. Vestenius^{aq}, A. Waked^{d,g}, E. Yubero^{ar}

^a European Commission, Joint Research Centre, via Fermi, 2748 21027, Ispra, Italy

^b Ricerca sul Sistema Energetico, RSE S.p.A, Via R. Rubattino 54, 20134, Milano, Italy

^c INERIS, Verneuil-en-Halatte, France

^d Univ. Grenoble Alpes, CNRS, IRD, G-INP, IGE, F-38000, Grenoble, France

^e TNO, Princetonlaan 6, NL-3584, CB, Utrecht, the Netherlands

^f Warsaw University of Technology, Faculty of Building Services, Hydro and Environmental Engineering, Nowowiejska 20 Str, 00-653, Warsaw, Poland

^g Département Sciences de l'Atmosphère et Génie de l'Environnement, SAGE, IMT Lille Douai, Université de Lille, 59000, Lille, France

^h C2TN, Centro de Ciências e Tecnologias Nucleares, Instituto Superior Técnico, Universidade de Lisboa, Estrada Nacional 10, Km 139.7, 2695-066, Bobadela LRS, Portugal

ⁱ IDAEA, CSIC, C/Jordi Girona 18-26, 08034, Barcelona, Spain

^j Laboratory of Ion Beam Physics, Institute for Nuclear Research, Hungarian Academy of Sciences (MTA Atomki), H-4026, Debrecen, P. O. Box 51, Hungary

^k Environmental Pollution Control Laboratory, Department of Chemistry, Aristotle University, 541 24, Thessaloniki, Greece

^l ARPA Piemonte, Dipartimento Rischi Naturali ed Ambientali, Via Pio VII 9, 10135, Turin, TO, Italy

^m Institute for Medical Research and Occupational Health, Zagreb, Croatia

ⁿ Université Savoie Mont-Blanc, LCME, F-73 000, Chambéry, France

^o Department of Physics & INFN, University of Genoa, 16146, Genoa, Italy

^p ARIANET Srl, Via Gilino n.9, 20128, Milano, Italy

^q Institute of Atmospheric Sciences and Climate (ISAC-CNR), Division of Lecce, Str. Prv. Lecce-Monteroni km 1.2, 73100, Lecce, Italy

^r Centro Regionale per il Monitoraggio della Qualità dell'Aria, ARPA Lombardia, via Juvara 22, 20129, Milan, Italy

^s Biology Department of University of Bari, via Orabona, 4, 70126, Bari, Italy

^t Institute of Nuclear and Radiological Science and Technology, Energy and Safety, N.C.S.R. Demokritos, Athens, 15341, Greece

^u Paul Scherrer Institut CH-5232 Villigen PSI, Switzerland

^v Forschungszentrum Jülich, Jülich and Rhenish Institute for Environmental Research at the University of Cologne, Cologne, Germany

^w CESAM & Dept Environment and Planning, University of Aveiro, 3810-193, Aveiro, Portugal

^x CIEMAT, Avda. Complutense, 40, Madrid, 28040, Spain

^y Institute of Atmospheric Sciences and Climate, National Research Council, Bologna, I-40129, Italy

^z School of Chemistry & Environmental Research Institute University College Cork, T12 YN60, Cork, Ireland

^{aa} Center for Air Resources Engineering and Science, Clarkson University, Potsdam, NY, 13699, USA

^{ab} Faculty of Chemistry, K. N. Toosi University of Technology, Tehran, Iran

^{ac} Departamento de Ingeniería Química y Bioprocesos, Pontificia Universidad Católica de Chile, Avda. Vicuña Mackenna 4860, Santiago, 7820436, Chile

^{ad} Agenzia Provinciale Protezione Ambiente, via Lidorno, 1, 38123, Trento, Italy

^{ae} ISSeP, Rue du Chêra 200, 4000, Liège, Belgium

^{af} National Institute of Nuclear Physics (INFN) – Florence Section, Sesto F.no (Fi), Italy

* Corresponding author.

^{ag} Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), Via Martiri di Monte Sole, 4, 40129, Bologna, Italy

^{ah} National Institute for Public Health and the Environment, Centre for Environmental Quality (MIL), P.O. Box 1, 3720, BA, Bilthoven, the Netherlands

^{ai} Department of Earth and Environmental Sciences, University of Milano Bicocca, P.zza della Scienza 1, 20126, Milan, Italy

^{aj} National Research Council of Italy, Institute for Atmospheric Pollution Research, Italy

^{ak} Regional Air Observatory Environmental Protection Agency of Veneto Region, Venice, Italy

^{al} Faculty of Physics and Applied Computer Science, AGH UST University of Science and Technology, Av. Mickiewicza 30, 30-059, Krakow, Poland

^{am} Regional Agency for Prevention, Environment and Energy of Emilia, Romagna, Italy

^{an} Laboratoire Inter-Universitaire des Systèmes Atmosphériques (LISA), UMR CNRS 7583, Université Paris Est Créteil et Université de Paris, Institut Pierre Simon Laplace, Créteil, France

^{ao} Department of Physics, Università degli Studi di Milano, & National Institute of Nuclear Physics, INFN-Milan, via Celoria 16, 20133, Milan, Italy

^{ap} University of Bologna, Department of Industrial Chemistry "Toso Montanari", Viale Risorgimento 4, I-40146, Bologna, Italy

^{aq} Finnish Meteorological Institute FMI, Atmospheric Composition Research, Erik Palménin aukio 1, FI-00560, Helsinki, Finland

^{ar} Atmospheric Pollution Laboratory, Department of Applied Physics, Miguel Hernandez University, Av. de la Universidad s/n, 03202, Elche, Spain

ARTICLE INFO

Keywords:

Source apportionment
PM₁₀
Receptor models
Chemical transport models
Intercomparison
Lens

ABSTRACT

In this study, the performance of two types of source apportionment models was evaluated by assessing the results provided by 40 different groups in the framework of an intercomparison organised by FAIRMODE WG3 (Forum for air quality modelling in Europe, Working Group 3). The evaluation was based on two performance indicators: z-scores and the root mean square error weighted by the reference uncertainty (RMSE_u), with pre-established acceptability criteria. By involving models based on completely different and independent input data, such as receptor models (RMs) and chemical transport models (CTMs), the intercomparison provided a unique opportunity for their cross-validation. In addition, comparing the CTM chemical profiles with those measured directly at the source contributed to corroborate the consistency of the tested model results. The most commonly used RM was the US EPA- PMF version 5. RMs showed very good performance for the overall dataset (91% of z-scores accepted) while more difficulties were observed with the source contribution time series (72% of RMSE_u accepted). Industrial activities proved to be the most difficult sources to be quantified by RMs, with high variability in the estimated contributions. In the CTMs, the sum of computed source contributions was lower than the measured gravimetric PM₁₀ mass concentrations. The performance tests pointed out the differences between the two CTM approaches used for source apportionment in this study: brute force (or emission reduction impact) and tagged species methods. The sources meeting the z-score and RMSE_u acceptability criteria tests were 50% and 86%, respectively. The CTM source contributions to PM₁₀ were in the majority of cases lower than the RM averages for the corresponding source. The CTMs and RMs source contributions for the overall dataset were more comparable (83% of the z-scores accepted) than their time series (successful RMSE_u in the range 25% - 34%). The comparability between CTMs and RMs varied depending on the source: traffic/exhaust and industry were the source categories with the best results in the RMSE_u tests while the most critical ones were soil dust and road dust. The differences between RMs and CTMs source reconstructions confirmed the importance of cross validating the results of these two families of models.

1. Introduction

Source Apportionment (SA) encompasses the modelling techniques used to relate the emissions from pollution sources to the concentrations of such pollutants in ambient air. Applications of SA techniques have been reported to (a) determine the causes of pollution levels exceeding legislation thresholds, (b) support the design of air quality plans and action plans, (c) assess the effectiveness of remediation measures, (d) quantify the contribution from different areas within a country, (e) quantify transboundary pollution, (f) quantify natural sources and road salting, and (g) refine emission inventories (Belis et al., 2017).

The abovementioned definition of SA accommodates a wide range of techniques to obtain information about the actual influence that one or more sources have on a specific area over a specific time window. Such techniques may be based on the measured concentrations of pollutants and their components (receptor-oriented models or, more simply, receptor models, RMs) or on chemistry, transport and dispersion models (also known as source-oriented models, SMs). Among the RMs it is possible to distinguish between (a) explorative methods, that rely on empirical coefficients or ratios between species, and (b) receptor models

based on multivariate analysis of all the chemical data at once (Watson et al., 2008; Hopke, 2016). SM techniques are based on air quality models, the most commonly used are Eulerian, Gaussian and Lagrangian ones (Fragkou et al., 2012). Gaussian models are used to describe only the dispersion of a pollutant near the source in a stationary way, while Lagrangian models describe it in a 3D dynamic way. With both types, the chemistry is assumed to be simple and linear or negligible, while Eulerian models (commonly referred to as Chemical Transport Models, CTMs) yield a description of pollutants in terms of motion, transport and dispersion, chemistry and other physical processes, accounting for all sources that are present in a given domain (emission inventories), as well as the influence from more distant sources (boundary conditions).

SA studies aim at quantifying the relationship between the emission from human activities or natural processes (referred to as activity sectors or macrosectors in the emission inventories) and the concentrations of a given pollutant (sectorial source apportionment). Specific SA techniques are designed to allocate the concentration of a pollutant to the geographic areas where it or its precursors were emitted (spatial source apportionment). The present study focuses on the first type of SA output resulting from either RMs or CTMs. These two families of models were selected because a) these are the most used ones in both research studies and air quality management applications in support of the EU Directive 2008/50/EC implementation (Fragkou et al., 2012; Belis et al., 2017), and b) these are techniques aimed at simultaneously describing the behaviour of all the knowable sources (that can be identified within the limitations of each technique).

RMs are used to perform SA by analysing the chemical and physical parameters measured at one or more specific sites (receptors). They are

¹ ARPAL Physical Agents and Air Pollution Sector, 19125 La Spezia, Italy.

² PM_TEN srl, Piazza della Vittoria Genova 7 (GE) - Italy.

³ Institute of Chemical Engineering Sciences (ICE-HT) Foundation for Research and Technology, Hellas (FORTH), GR-26504 Patras, Greece.

⁴ TCR Tecora, V. delle Primule 16, 20815 Cogliate (MB), Italy.

based on the mass conservation principle and identify sources/factors by solving the mass balance equation $\mathbf{X} = \mathbf{G}\mathbf{F} + \mathbf{E}$, where \mathbf{X} is a matrix containing ambient measurements of pollutant properties, typically chemical concentrations of gases and/or particles that include markers for different sources (or source categories), \mathbf{F} is a matrix whose vectors rows represent the profiles of p sources, \mathbf{G} is a matrix whose columns represent the contributions of the p sources and \mathbf{E} is the residual matrix. Those techniques applying weighted least-squares minimisation (or other type of minimisation techniques) fit to the ambient measurements using measured source profiles to solve the equation are referred to as chemical mass balance methods (e.g., CMB), while models which solve the equation without using ‘a priori’ information on source composition are factor analytical models, the most common of which is positive matrix factorisation (PMF). A detailed description of the RMs, is provided e.g. in [Watson et al. \(2008\)](#), [Hopke \(2016\)](#), [Viana et al. \(2008\)](#), and [Belis et al. \(2013\)](#).

In [Table 1](#) are briefly presented the approaches used to apportion a pollutant to its sources with CTMs: a) brute force (BF) (a terminology widely used in the literature e.g. [Pillon et al., 2011](#); [Burr and Zhang, 2011](#); [Baker et al., 2013](#); [Lin and Wen, 2015](#); [Wang et al., 2015](#)) also

Table 1
Description of CTM source apportionment approaches: tagged species and brute force or emission reduction impact.

Approach	Tagged Species (TS)	Brute force (BF) or Emission reduction impact (ERI)
Description	A specific module tags chemical species to track the contributions of sources in every grid cell to both primary and secondary compounds (reactive tracers).	Estimate the contribution of each source by subtracting from a run with all the sources (base case) a run where the emissions of the investigated source have been reduced by a given percentage.
Kind of approach	Static Depicts the situation corresponding to the input dataset.	Dynamic It is actually a sensitivity analysis. Estimations are obtained by introducing changes in the emissions to make inferences on situations different than those of the input dataset.
Underlying question	What is the actual mass transferred from a pollutant source to its concentration in a given location and period?	What would be the reduction in the pollutant concentrations corresponding to a given reduction in the emissions of its precursors?
Runs	Apportion all the sources in one single run.	Requires a number of runs equal to the number of sources to apportion plus the run with all sources (base case).
Mass conservation	The sum of the contributions of the sources equals the total mass (concentration) of the apportioned pollutant (by definition).	The mass allocated to the sources is obtained from independent runs. Therefore, the sum of source contributions may not match the total pollutant mass obtained in the base case (non-linear behaviour).
Advantages	Provides a picture of the contributions referred to the specific emission inventory and meteorological fields used as input. Useful to attribute the actual impacts of sources on health and vegetation.	Useful to evaluate the impact of abatement measures. Can be used with any model.
Disadvantages	Requires additional coding effort. For non-linear pollutants, the source contributions cannot be extrapolated to situations different than the modelled case.	Requires many runs. The mass is not always conserved (see above). The mass allocated to a source depends on the emission reduction level.

referred to by [Thunis et al. \(2019\)](#) as emission reduction impact (ERI), and b) tagged species (TS).

In the TS approach, a specific module, executed during the model run (online), earmarks the mass of chemical species to track the atmospheric fate of every source category or region (or a combination of both). Since the tagged species (also known as “reactive tracers”) can undergo chemical transformations, the module is also able to track the sources of secondary pollutants. The TS approach fulfils a mass balance equation ensuring the sum of all the source contributions equals the total concentration of the pollutant. TS algorithms are implemented in several CTM systems, such as, CAMx-PSAT ([Yarwood et al., 2004](#); [Wagstrom et al., 2008](#); [ENVIRON, 2014](#)), CMAQ-TSSA ([Wang et al., 2009](#)), CMAQ-ISAM ([Byun and Schere, 2006](#); [Foley et al., 2010](#); [Napelenok et al., 2014](#)), DEHM ([Brandt et al., 2013](#)) and LOTOS-EUROS ([Hendriks et al., 2013](#); [Kranenburg et al., 2013](#)).

The BF/ERI approach estimates the concentration changes (delta) attributable to a source by subtracting from the run where all the source emissions are present (the base case) a run where the emission source to be analysed has been modified (normally reduced) by a given percentage (e.g. [Burr and Zhang, 2011](#); [Wang et al., 2015](#); [Uranishi et al., 2017](#)). In this approach, an independent run is necessary for every source. Unlike TS, there are no specific model modules to execute the BF/ERI approach. It can be applied with any CTM in a post-processing step. In nonlinear situations, the sum of the concentrations allocated to the single sources can differ from the concentration of the pollutant in the base run. This problem may be dealt with in different ways. The more straightforward one is to re-normalise the mass of the sources to the total mass of the pollutant in the base case applying, for instance, multilinear regression or a similar method. A more appropriate technique from the mathematical point of view is to quantify such non-linearities by attributing a mass to all the possible interactions between sources. This is accomplished by computing additional simulations where two or more sources (all the possible combinations) are reduced at once and subsequently applying a factor decomposition ([Stein and Alpert, 1993](#); [Clappier et al., 2017](#)). However, such elaboration requires a considerable number (N) of simulations; where $N = 2^s - 1$ and s denotes the number of sources. Another characteristic of the BF/ERI approach is that the mass attributed to the sources depends on the applied emission reduction factor ([Clappier et al., 2017](#)).

In summary, RMs and CTM-TS approach quantify the mass that is transferred from the source to the receptor. Therefore, they can be grouped under the category mass-transfer (MT) source apportionment. On the contrary, the CTM-BF/ERI approach is actually a sensitivity analysis method that estimates the changes in concentration that would result from a change in emissions and, hence, it should be placed under a different category. Since the results obtained with MT and BF/ERI may differ, we propose to call the output of the former “contribution” and the one of the latter “impact”.

SA techniques have been used for many air pollutants such as NO_2 , O_3 , VOC (e.g. [Baudic et al., 2016](#); [Dunker et al., 2015](#)). However, most of the studies focus on particulate matter (PM). Considering the higher experience of SA on this pollutant and the high number of exceedances in Europe ([EEA, 2018](#)) and in many areas of the world ([Karagulian et al., 2015](#)), this study is focused on SA of PM_{10} .

Assessing the performance of SA model applications is essential to guarantee reliable and harmonised information on the influence of sources on pollution levels. To that end, a series of intercomparisons were performed in the framework of FAIRMODE (Forum on Air Quality Modelling) focused on RMs ([Belis et al., 2015a](#) and [2015b](#)). Moreover, as highlighted in a recent study combining PMF and CAMx-PSAT, a reference methodology to comparing the results of RMs and CTMs is missing ([Bove et al., 2014](#)). In this paper, we introduce a comprehensive approach for the assessment of different source apportionment techniques at once as a step towards an integrated assessment and validation methodology for SA models. The objectives of this study are:

- To assess the differences in the SA results between RMs, between the TS and BF/ERI CTM approaches, and between RMs and CTMs;
- To characterise the performance of SA model applications using pre-established criteria for deciding whether results are comparable or not;
- To assess whether the CTM SA performance is influenced by differences between sites;
- To test the impact of different spatial resolutions and vertical mixing coefficients on the CTM SA performance.

To achieve the abovementioned objectives, a consolidated assessment methodology already applied to RMs was adopted.

2. Methodology

The assessment of the SA models followed multiple steps to evaluate the different aspects of a SA result (see Appendix sections A1 and A2). The focus of this article is on the performance tests used to evaluate every source category separately. Considering that the actual contribution/impact of sources in a given area is unknown, the reference for the performance test is derived from the participants' average and standard deviation. The advantage of this methodology is that every test is associated with acceptability thresholds based on widely recognised indicators that have been tested and used in previous FAIRMODE SA intercomparisons. For a detailed description of the methodology refer to Belis et al. (2015a, 2015b). The same approach was used to evaluate the performances of RMs and CTMs separately. The RMs were assessed, only at the reference site of Lens (see next section for details), while a set of 10 different sites (Fig. 1) was used in the comparison between CTMs. In addition, a cross comparison between RMs and SMs was accomplished for the Lens site, setting the RMs as the reference for methodological reasons without any "a priori" implication about the reliability of the different methods.

This study involved 40 groups, 33 of which delivered 38 RM results

and the remaining seven reported 11 CTM results. The participants delivered results consisting of a set of candidate sources (hereon, candidates) with the estimation of their contributions or impacts (hereon, source contribution or impact estimates, SCIEs) expressed as $\mu\text{g}/\text{m}^3$. In section A3 of the Appendix are listed the most commonly used abbreviations.

2.1. Set up of the RMs intercomparison

The reference site of Lens (France), located in a densely populated area (>500 inhabitants/ km^2), is part of the official monitoring network and is classified as an urban background monitoring station (Fig. 1). According to the EU legislation (2008/50/EC), such locations are places in urban areas where levels are representative of the exposure of the general urban population.

The dataset of measurements used for the RM intercomparison was produced in the framework of the CARA project directed by the French reference laboratory for air-quality monitoring (LCSQA). This dataset contained 116 PM_{10} daily concentrations collected every third day between March 2011 and March 2012. The concentration of 98 chemical species and their uncertainties were provided for every sample including: ions, organic and elemental carbon (OC/EC), trace elements, polycyclic aromatic hydrocarbons (PAHs), anhydrosugars, hopanes, n-alkanes, primary organic aerosol (POA) markers, and the total PM_{10} gravimetric mass concentrations (Table A2 of the Appendix). Details on the analytical techniques can be found in Waked et al. (2014) and Golly et al. (2015). The average PM_{10} concentration and composition in the dataset is shown in Table S1 of the supplementary material. In order to harmonise the nomenclature, the candidates reported by participants were encoded in conformity with the SPECIEUROPE database source categories (Pernigotti et al., 2016). The source chemical profiles from this database are also referred to as 'reference' profiles (Table 2).

In this study the fuel oil source is the combustion of residual (heavy) oil that normally takes place in power plants. Since this fuel is also used

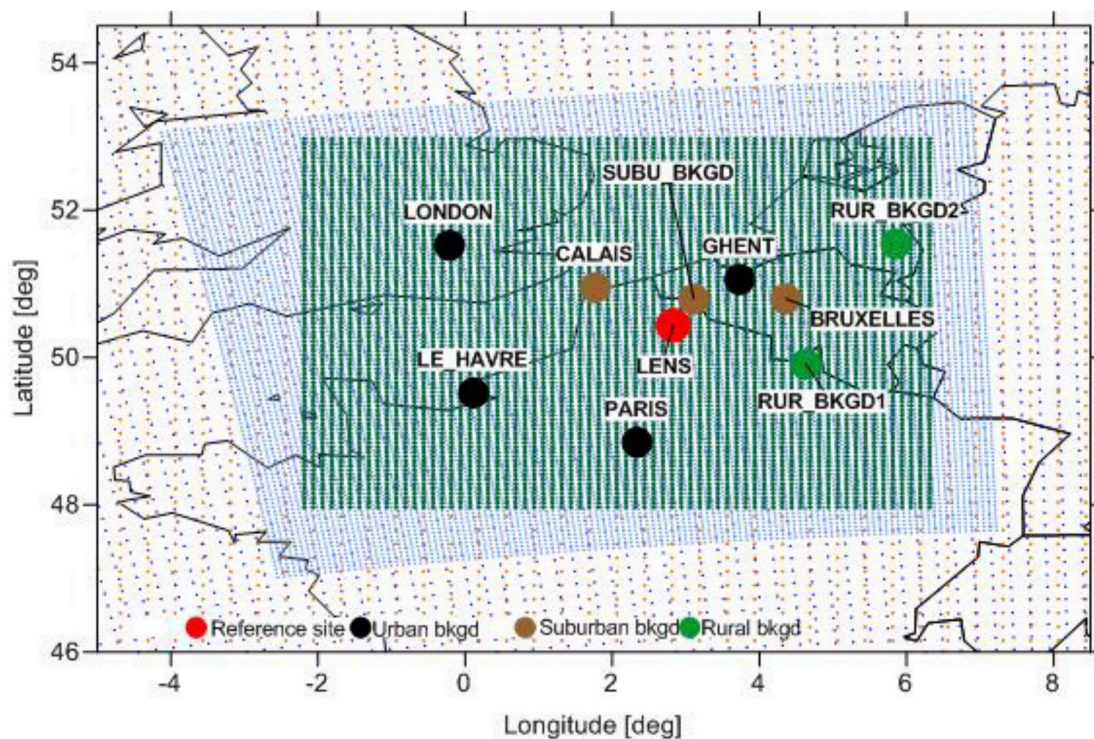


Fig. 1. Map of the Lens domain with indication of the receptors selected for reporting SA results (circles). The receptor colour represents the type of site. FAIRMODE-LENS domain (green dots); WRF-FAIRMODE domains (light blue dots). The lower resolution domains FAIRMODE-EU and WRF-EU domains are also visible in the background (red and blue dots). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 2

Source categories reported in the RM results with the corresponding SPECIEUROPE code. In SPECIEUROPE there are general categories that include others, this is indicated in the hierarchy of categories.

SPECIEUROPE Code	Source categories	Description	Abbr. in graphs	Hierarchy of categories (low to high)	n. of reference profiles	n. of candidate sources reported by participants
1	traffic	all vehicle emissions	tra	1	286	34
2	exhaust	vehicle exhaust only	exh	2_1	130	20
3	diesel	diesel vehicle exhaust	die	3_2_1	49	2
4	gasoline	gasoline vehicle exhaust	gas	4_2_1	10	3
5	road (dust)	traffic resuspension and wear (e.g. tyres, brakes)	roa	5_1	154	16
7	brake	brake wear	bra	7_5_1	2	3
10	soil (dust)	natural dust	soi	10	235	34
12	marine	fresh sea salt	ss	12		39
20	industry	unspecified industrial emissions	ind	20	433	24
30	fuel oil	combustion of residual oil (mainly power plants)	fue	30	88	31
31	coal	combustion of coal	coa	31	47	8
37	ship	emissions from vessels	shi	37	14	9
40	biomass	biomass burning	bib	40	139	39
41	wood	wood burning	woo	41_40	96	8
60	SIA	secondary inorganic aerosol (61 + 62)	sia	60		20
61	ammonium nitrate	secondary inorganic aerosol	amn	61_60	1	20
62	ammonium sulphate	secondary inorganic aerosol	ams	62_60	1	21
69	metallurgy	metallurgic industry	met	69_20	43	2
70	POA	primary organic aerosol	poa	70		33
71	aged sea	secondary sea aerosol	as	71	1	21
72	agriculture	all agricultural emissions	agr	72	20	2

in maritime transport, ship emissions in seaports may contribute up to 10% of the primary PM₁₀ (Amato et al., 2016; Bove et al., 2018). However, due to the considerable distance of the study site from the closest harbour (75–80 km), the influence of this source on the primary PM₁₀ is expected to be small. In this exercise, very few results (11%) univocally identified shipping as a source. This is in line with a study on PM₁₀ sources using receptor models by Waked et al. (2014) in the same location which allocated a 4% of primary PM₁₀ to oil combustion associated with land activities (power generation and industrial emissions) excluding the influence of ship emissions (section 3.1.1).

2.2. Set up of the CTM intercomparison

2.2.1. Domains and time window

CTMs were run over two nested computational domains with different spatial grid resolutions: one covering the whole of Europe and a smaller one around the reference site area (Lens). The FAIRMODE-EU domain lat/long grid (Table S2, supplementary material) was set to be compatible with the spatial resolution of the emission inventory and to avoid any interpolation of the emission data (section 2.2.2). The domain extension was defined to provide suitable regional boundary conditions for the reference area simulations. It includes a portion of North Africa to account for dust emissions, while the northern boundary was set around latitude 65.0 to minimise the spatial distortions at high latitudes when using a lat/long grid. The grid step corresponds to roughly 18–20 km.

The FAIRMODE-LENS domain (Fig. 1) was defined as a subset of the emission inventory grid as well as of EU grid, once again to avoid any interpolation of emission data. The domain is centred over Lens, but it is large enough to allow a reasonable description of the PM fate in the atmosphere, limiting the influence of boundary conditions. The grid step corresponds to roughly 6–7 km.

Two three-month periods for CTM modelling were defined: 1/6 to 31/8/2011 and 15/11/2011 to 15/2/2012. Such time windows were selected to be representative of both the warm and cold seasons and long enough to include both high pollution episodes and low concentration situations. Moreover, because PM chemical composition data were available at Lens only every third day, three-months simulations were

needed to pair at least 30 daily SA results of RMs and CTMs.

2.2.2. Input datasets

Emissions. The anthropogenic emissions used for this intercomparison were derived from the TNO_MACC-III emission data, which is an update of the TNO_MACC-II emission inventory (Kuenen et al., 2014). The inventory is structured according to the Selected Nomenclature for sources of Air Pollution (SNAP, EEA, 1996; EEA, 2000) and is combined with fuel use information by country and by sector derived from the literature and country data. This emission inventory included an enhanced source classification according to fuels for SNAP macrosectors 2 and 7 (Tables 3 and Table A3). Particularly, emissions due to combustion from the residential and small commercial combustion sectors were split between fossil fuels (coal, light liquid, medium liquid, heavy liquid and gas) and solid biomass. Emissions from the road transport sector were split according to three main fuels (gasoline, diesel and LPG/natural gas), while non-exhaust sources included evaporation and tyre and road wear. Among non-road transport emissions, international shipping was specifically accounted for. The reconstruction of the natural emissions (dust resuspension, sea salt and biogenic VOCs) was left to the modelling groups, because in most cases such emission modules were embedded in the modelling chain.

Meteorological fields. Meteorological fields were obtained by applying the WRF model in a nested (one way) configuration in order to provide fields for both the EU and LENS domains. The WRF simulations were performed over a Lambert conformal domain. This choice implied a pre-processing phase of the meteorological fields in order to feed CTMs, however, it should be considered that: (a) most of the CTM model chains needed meteorological fields pre-processing anyhow (e.g. to compute turbulence and other additional parameters), and (b) the use of a lat/long grid over northern Europe may lead to considerable distortion effects.

Both WRF domains covered the corresponding CTM/output domains leaving also a border area. To limit the degradation of the meteorological information during the interpolation phase the WRF domains

adopted the same grid step as the CTM domain (18 km and 6 km, respectively). The WRF physical configuration is provided in Table S3 of the supplementary material. The LOTOS-EUROS was the only model run with its own ECMWF meteorology.

Boundary conditions. Boundary conditions were derived from the MACC global model using the same approach that has been adopted in other European initiatives like EURODELTA-III and AQMEII-3 (<http://aqmeii.jrc.ec.europa.eu/>).

2.2.3. Definition of common sources for RMs and CTMs and CTM receptor sites

RM sources are determined by their chemical fingerprints and time trends, while CTM sources reflect the nomenclature of the emission inventory used as input. In addition, CTMs allocate both primary and secondary PM to its sources, the latter being attributed to the sources of its gaseous precursors. Due to the different definition of their sources, comparing the SA output of RMs and CTMs is not straightforward. Table 3 shows the correspondence between RMs and CTMs sources established in this study to compare their outputs. To that end, the SNAP source classification commonly used in the emission inventories (Table A3) is compared at the macrosector level with the SPECIEUROPE source categories. Since RMs did not provide SCIEs comparable with SNAP macrosectors 5 (extraction and distribution of fossil fuels), 6 (product use), and 9 (waste treatment), the mass apportioned by CTM to these macrosectors was pooled under a generic non-apportioned mass category (99 OTH). For the remaining source categories, two sets with different degree of detail for macrosectors 2 (residential and commercial combustion), 7 (road transport) and 11 (natural sources) were defined (Table 3). The mandatory (MDT) set consisted of 7 sources plus one corresponding to the non-apportioned mass (99 OTH). The more comprehensive optional set (OPT) encompassed 13 sources plus the non-apportioned category.

To assess the spatial variability of the CTM SA performances, a set of receptor sites representing different types of locations (urban, suburban and rural) among the AIRBASE European monitoring network (EEA, 2015) were selected (Table 4). In addition, the site of Ghent was

included because of the availability of detailed PM₁₀ chemical composition from the CHEMKAR PM₁₀ study (VMM, 2013).

2.3. CTM model performance evaluation (MPE)

Each modelling group performed its own model performance evaluation (MPE) by comparing CTM results against observed data for the main chemical species (i.e. sulphate, nitrate, ammonium, elemental carbon and organic aerosol). In order to support the interpretation of the SA performance, a centralised MPE was also accomplished for PM₁₀ and its main chemical components on all the model base case results (Annex 1, supplementary material). The objective was to assess whether the ability of models to reproduce the PM mass and main chemical components influences the model SA output. To that end the modelled mass of PM and main chemical components were compared with measurements. The centralised MPE was performed using a subset of sites selected to cover the different meteorological and emissive features of the Lens domain: Lens (reference site), London and Paris (megacities), Ghent (middle-sized city), Le Havre and Calais (coastal areas), and the rural background station Vredepeel-Vredeweg. Model results were evaluated by means of a few statistical indicators: the mean bias (MB), the normalised mean bias (NMB), the root mean square error (RMSE) and the Pearson correlation.

2.4. CTM source chemical profiles

The differences between CTM SA results were assessed using a similarity test discussed in section 3.2.2. It was accomplished by comparing the chemical profiles of the sources reported by participants (CTM chemical profiles) among each other (ff tests) and with a set of 1160 external (measured and derived) source profiles (fr tests) deriving from the SPECIEUROPE database developed at JRC (Pernigotti et al., 2016) and the SPECIATE database of the US EPA (Hsu et al., 2014).

Table 3

Correspondence of source nomenclature between CTMs and RMs for both the mandatory and optional sets. A more detailed description of the SNAP macrosectors is given in Table A3.

SNAP	Description	CTM mandatory 7 sources	RM correspondence	CTM optional 13 sources	RM correspondence
1	Energy industry	01 ENI	30 fuel oil combustion	01 ENI	30 fuel oil combustion
2.1	Residential & Commercial combustion, fossil fuels	99 OTH	-	02 OTH	74 combustion
2.2	Residential & Commercial combustion, biomass burning	02 BIO	40 biomass burning	02 BIO	40 biomass burning
3 & 4	Industry (combustion & processes)	34 IND	20 industry	34 IND	20 industry
7.1	Road transport, exhaust, gasoline	07 RTR	1 traffic	71 RTG	2 exhaust
7.2	Road transport, exhaust, diesel	07 RTR		72 RTD*	
7.3	Road transport, exhaust, LPG/natural gas	07 RTR		07 RTR	9 other traffic
7.4	Road transport, non-exhaust, evaporation	07 RTR		07 RTR	9 other traffic
7.5	Road transport, non-exhaust, wear	07 RTR		75 RTW	5 road
8	International shipping	08 SHP	37 ship	08 SHP	37 ship
10	Agriculture	10 AGR	72 agr**	10 AGR	72 agr**
11.1	Dust	11 DST	10 soil (crustal)	11 DST	10 soil (crustal)
11.2	Sea salt	99 OTH	-	11 SLT	12 marine or 71 aged sea salt
11.3	Biogenic SOA	99 OTH	73 bso**	11 BSO	73 bso**

*include also the non-road emissions, **reported in very few RM results

Table 4
Receptor sites used to assess the spatial variability of the CTM SA performances.

Station Code	Station Name	Country	Station Type	Area Type	LONG	LAT
LENS	Lens-CARA	France	Background	urban	2.83	50.44
CALAIS	Sangatte	France	Background	suburban	1.77	50.95
LE HAVRE	Le Havre Henri Fabre	France	Background	urban	0.11	49.52
PARIS	Paris 6eme	France	Background	urban	2.34	48.85
LONDON	London N. Kensington	United Kingdom	Background	urban	-0.21	51.52
BRUSSELS	Uccle	Belgium	Background	suburban	4.36	50.80
GHENT	Ghent	Belgium	Background	urban	3.73	51.06
SUBU BKGD	Menen	Belgium	Background	suburban	3.11	50.79
RUR BKGD1	Revin	France	Background	rural	4.63	49.91
RUR BKGD2	Vredepeel-Vredeweg	The Netherlands	Background	rural	5.85	51.54

3. Results and discussion

3.1. Receptor models

A total of 38 different RM results were delivered by 33 participants. The majority of the RM results (31) were obtained using the US-EPA version 5 of the positive matrix factorisation tool (EPA-PMF5; Norris et al., 2014). ME-2 (multilinear engine 2; Paatero, 1999) scripts were used in two cases. Only one result was reported with the following tools: RCMB (robotic chemical mass balance; Argyropoulos and Samara,

2011), MLPCA-MCR-ALS (MLPCA, maximum likelihood principal components analysis-multivariate curve resolution-alternating least squares, Izadmanesh et al., 2017), PMF version 2 (PMF2; Paatero and Tapper, 1993; Paatero, 1997), EPA-PMF version 3 (PMF3; Norris et al., 2008) and EPA-PMF version 4 (PMF4). The RM results were labelled with letters from A to Z and then from *A to *L. In Figs. 2 and 3, the sources are labelled in conformity with the corresponding SPECIEUROPE database source categories (Pernigotti et al., 2016).

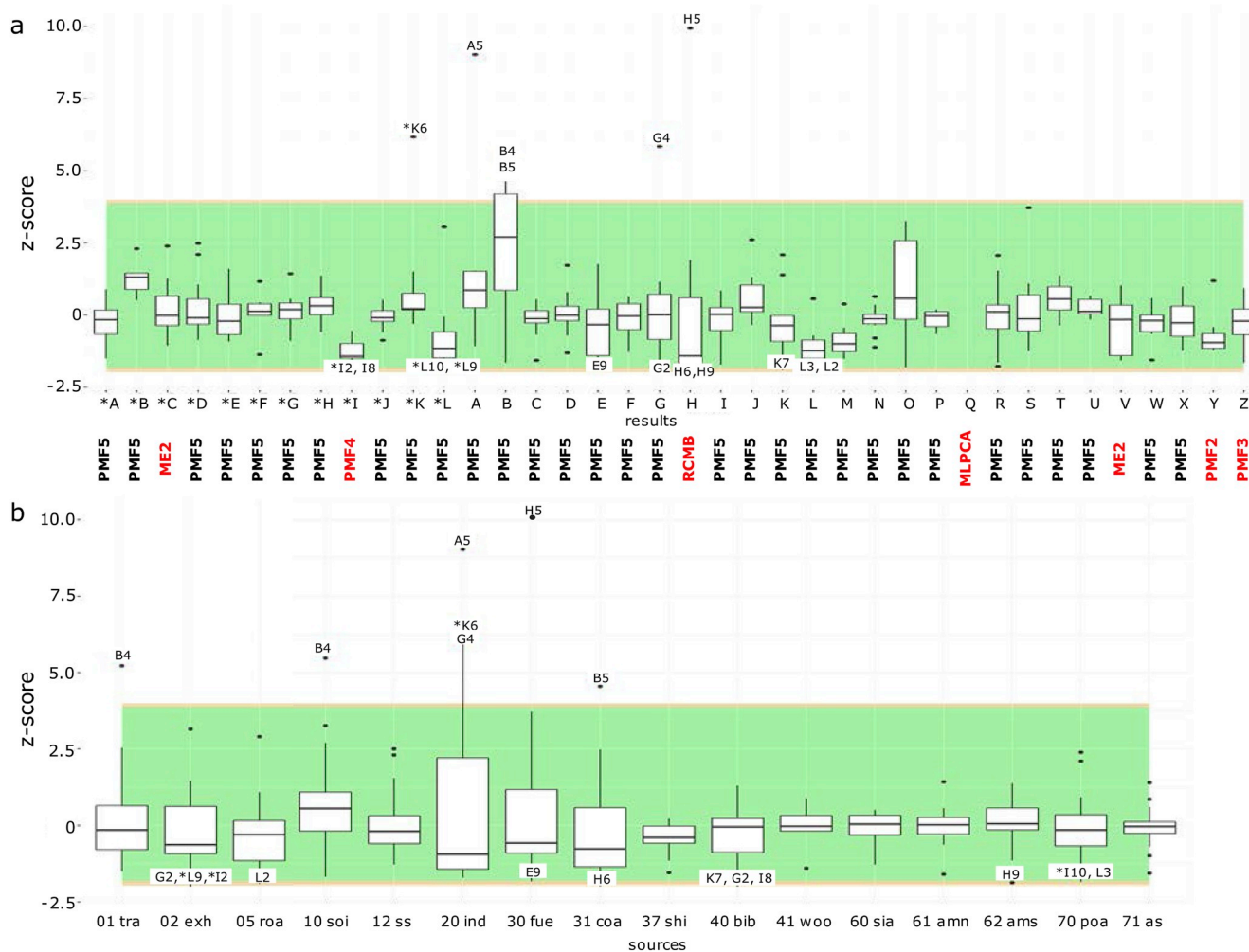


Fig. 2. z-scores performance indicator values arranged by participant (a) and by source category (b). The areas of acceptance and warning are indicated with green and orange background, respectively. Only candidates with warning or bad scores are annotated in the plots (letter: result; sequential number: candidate). In (a) the used RM is indicated under each participant (red for models different than PMF5). The source category codes are given in Table 2. The scores of result Q are not shown because out-of-scale. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

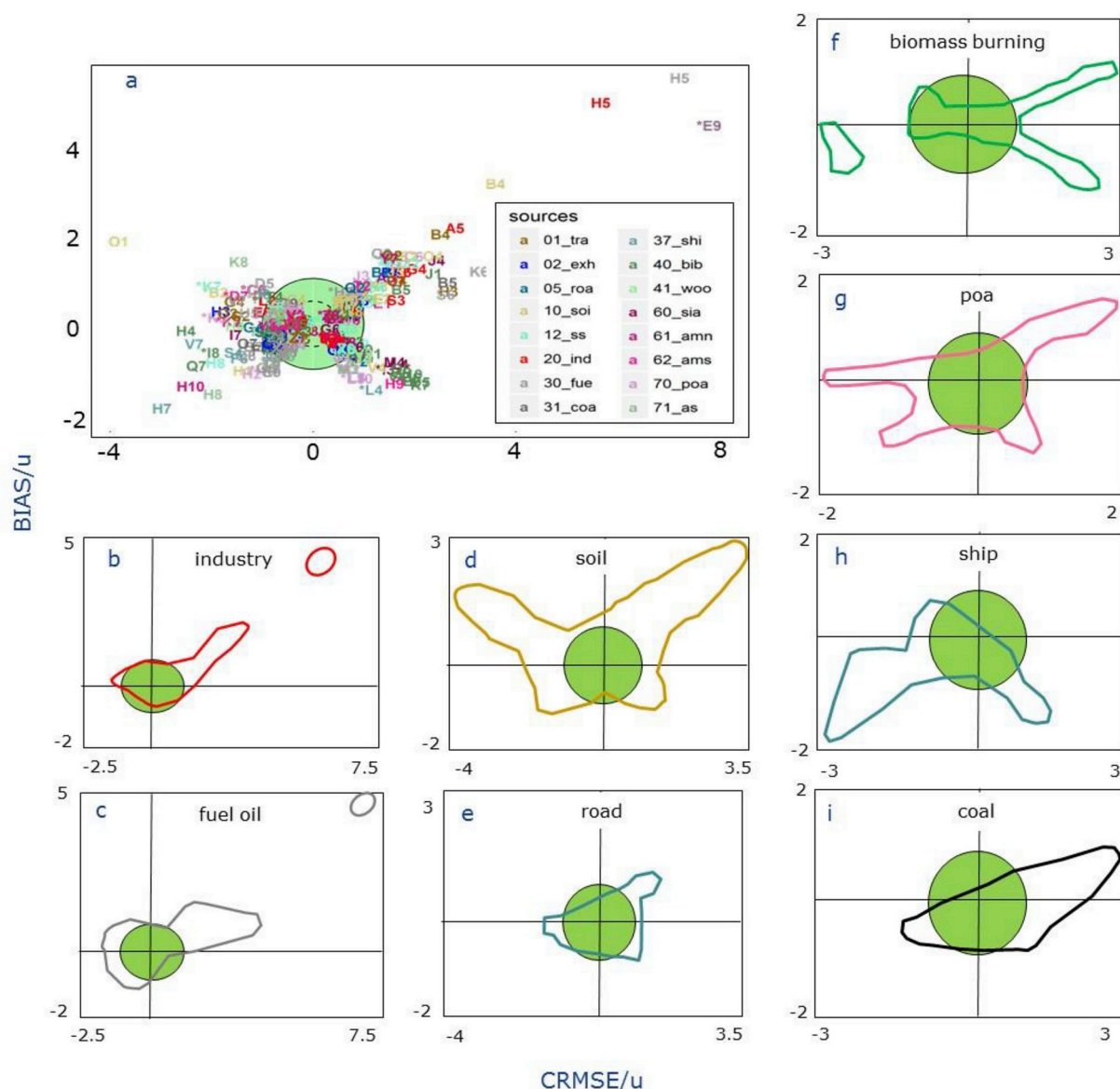


Fig. 3. Target plots of the $RMSE_u$ performance indicator for all the candidate sources (a). The letter of the alphanumeric codes denotes the result while the number is a sequential identifier of the candidates in every result. The scores of selected sources are depicted schematically: (b) industry; (c) fuel-oil; (d) soil; (e) road; (f) biomass burning; (g) primary organic aerosol (poa); (h) ship and (i) coal. Scores inside the green circle (<1) are acceptable. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

3.1.1. Performance tests

The z-scores evaluate the performance of the SCIE for the overall studied time window (one year in this study). In Table 5 are shown the reference SCIE for the sources reported in more than three SA results (diesel, gasoline, brake, metallurgy and agriculture are excluded) with their uncertainty. The time series of the reference and the scatter plots between candidates and references are shown in Figs. S1 and S2 of the supplementary material. If we consider the consensus between participants as an indicator of the relevance of the sources those reported in less than 25% of the results (coal, and ship) could be regarded as non-robust ones (Pernigotti and Belis, 2018). Although biomass burning and wood burning have a slightly different connotation (the former is more inclusive than the latter), in this study the two expressions have apparently been used by participants to represent the same source. The SCIE of a nine source solution reported by Waked et al. (2014) in the same site (Table 5, last two columns) fall within the uncertainty range of the reference (except one). The main differences of the French study with the references of the present study are the absence of industry, the

lack of detail within the traffic source (exhaust, road dust) and the lower contribution of the fuel oil source.

In the performance tests, 91% of the z-scores of RM candidates (Fig. 2) fall within the area of acceptability (green) indicating a general good agreement between the reported results and the reference values with a tolerance of 50% (Fig. 2a). The results Q, B, H, *I and *L are those with the higher number of candidates out of the acceptability zone. The result Q was obtained with the recently developed MLPCA-MRC-ALS technique. It is similar to PMF with the difference that it solves the mass balance equation using the alternating least squares algorithm (Tauler et al., 2009). The out-of-scale z-score of this result has been attributed to a misinterpretation on how to report the average SCIE. This interpretation was confirmed by the analysis of the time series where the $RMSE_u$ are of the same order of magnitude of as the other participants and in some cases passed the test (see below in this section).

Most of the overestimated SCIEs are in the industry source category (20 ind) while some others are in traffic, soil, biomass burning and coal combustion source categories (1 tra, 10 soi, 30 fuel and 31 coa; Fig. 2b).

Table 5

Overall SCIEs and uncertainties of the reference sources used for the RM z-score tests.

SOURCE	ID	results with this source (%)	reference SCIE		uncertainty		Waked et al. (2014)		
			($\mu\text{g}/\text{m}^3$)	(%)	($\mu\text{g}/\text{m}^3$)	(%)	($\mu\text{g}/\text{m}^3$)	(%)	
traffic	1	89	2.7	10	1.8			1.5	7
exhaust	2	53	3.8	15	2.4				
road	5	42	4	15	2.9				
soil	10	89	2.6	10	1.1			2.8	14
marine	12	100	1.7	7	0.8			1.2	6
industry	20	63	1	4	1.1				
fuel oil	30	82	2.2	9	1.1			0.7	4
coal	31	21	1.4	5	1.3				
ship	37	24	3.8	15	0.5				
biomass	40	100	2.8	11	0.7			3.0	15
wood	41	21	2.6	10	0.9				
SIA	60	53	5.5	21	1.3				
ammonium nitrate	61	53	4	15	1.2			4.2	20
ammonium sulphate	62	55	3.5	13	1.4			3.2	16
POA	70	87	2.6	10	1.2			1.2	6
aged sea	71	55	2.9	11	0.8			2.7	13

Conversely, underestimated SCIEs are observed mainly in the exhaust, biomass burning, and primary organic aerosol, source categories (2 exh, 40 bib and 70 poa).

The RMSE_{u} scores of the candidates are displayed in target plots (Fig. 3) where the abscissa represents the uncertainty-weighted centred root mean square error (CRMSE/u) and the ordinate the uncertainty-weighted bias (BIAS/u). The green circle delimits the acceptability area while the dotted and solid lines correspond to the 0.5 and 1.0 levels, respectively. The candidates are indicated with the same codes as used in Fig. 2.

In the RMSE_{u} test, 72% of the candidates fall in the acceptability area (Fig. 3a). The areas with scores of sources showing a high share of rejected candidates are schematically represented in Fig. 3 b – i. Those are: industry (3b), fuel oil (3c), soil (3d), road (3e), biomass burning (3f), primary organic aerosol (3g), ship (3h) and coal (3i). The source categories showing the highest percentage of candidates with poor scores in this test are ship (75%), coal burning (71%), and fuel oil (60%). The first two are likely not well identified sources, considering that they were reported in less than 25% of the results. The low or non-detectable contribution of coal burning is explained by the low share (<5%) in 2011-2012 of this fuel in the French energy mix (<https://www.iea.org>) while contributions from primary ship emissions are unlikely, considering the distance of the site to harbours. Although fuel oil was reported in the majority of the results and the average SCIEs are coherent with the reference (as indicated by the z-score test), the time series of this source in the different results are divergent. In this source, the rejected RMSE_{u} scores are distributed in different quadrants of the target plot. However, those with higher deviations are located in the top right quadrant suggesting overestimation and higher amplitude than the reference. A similar situation is observed in the time series of the industry source category where a 30% of the candidates presented both positive bias and amplitude problems with respect to the reference (i.e. fall in the top right quadrant of the target plot, Fig. 3b). The RMSE_{u} confirms the indications of the z-score test for the industry and coal combustion.

Positive bias is also observed in some candidates of soil and traffic while biomass burning and primary organic aerosols (70 poa) present either positive or negative biases. Extreme cases of variance higher than the reference are candidates H5 (industry, fuel oil), *E9 (fuel oil) and B4 (soil). On the contrary, the rejected candidates of the ship source category (37 shi) are mainly located in the lower quadrants suggesting a tendency to underestimate the reference SCIE time series. In a number of sources, the bad scores are distributed in all the quadrants of the target

plot (Fig. 3), therefore, the lack of coherence between them and the reference is probably depending on random variability determined by a combination of factors.

Additionally, the frequency of the candidates in every source category (Table 2) is an indicator of robustness of the source identification (Pernigotti and Belis, 2018). The better identified source categories from this point of view are: biomass burning, marine, traffic, soil, primary organic aerosol, fuel oil, industry, aged sea salt, and the secondary inorganic aerosol (SIA, ammonium nitrate and ammonium sulphate).

3.1.2. Conditions influencing the performance of RM

Although RMSE_{u} and z-score (without sign) are correlated to a certain extent ($R^2 = 0.43$; Fig. S3), they evaluate different aspects of the SA results. As observed in previous intercomparisons (Belis et al., 2015b), the RMSE_{u} test is stricter because, in addition to the bias, it assesses the phase and amplitude of the time series curves (Jolliff et al., 2009). Moreover, this finding is coherent with the concept that the dispersion of the result's average SCIEs around the overall mean (z score) is lower than the one of the single samples' SCIEs in all the results (RMSE_{u}).

The high number of RM results evaluated in this study (38) gives a unique opportunity to analyse the conditions that influence the performance of the results. To that aim, their RMSE_{u} and z-scores were compared with the self-declared experience of the RM practitioners and the number of candidates in the reported results as shown in Fig. S4 of the supplementary material.

In Fig. S4a, the performance of results delivered by practitioners that have conducted 10 or less studies (intermediates) is quite variable while practitioners declaring to have conducted more than 10 studies (experienced); 24% of the total number of participants) always show good performances (low scores). The Mann-Whitney non-parametric test (Palisade StatTools 7.6) confirms that experienced practitioners have significantly better RMSE_{u} performance than intermediate ones ($p=0.0267$) and the Wilcoxon Signed-Rank non-parametric test (Palisade StatTools 7.6) proves that the experienced practitioners have RMSE_{u} lower than 1 ($p=0.001$) while this hypothesis is rejected for intermediate ones ($p=0.7378$). The difference between experienced and intermediates is less evident for the z-scores. Results (solutions) with a number of candidate sources (factors) close to the average (9) present better z-scores than those with a difference of 3 or more sources (Fig. S4b). This relationship is less evident for RMSE_{u} suggesting that failing to retrieve the optimal number of sources impacts mainly on the bias of the result. In multivariate analysis including redundant variables, chemical species with high covariance may degrade the quality of the results and lead to poor model output diagnostics. For that reason, variables that are highly correlated with others may be removed because they do not add meaningful information to the analysis. This is very often the case for families of organic species (PAHs, alkanes, hopanes) in which many or all of the members show very similar time trends. In such cases, replacing them either by one representative of the family or by the sum of all the members of that family leads to better model diagnostics and more physically meaningful results. In this study, it was up to the practitioners to establish the optimal set up of chemical species for their runs. A variable number of the 98 species present in the input dataset were used by practitioners to obtain the reported results. Only seven of them were obtained with more than 80 species while the majority used less than 40 species. Such a low number of species is due to the exclusion of some of them (after selecting the most robust ones) or to the practice of pooling the members of some organic species families (PAHs, hopanes, n-alkanes) and treating them as one single species. The comparison between the number of species and the performance indicators (not shown) did not present any clear relationship between these two parameters.

Constraints were applied in only eight of the RM results (21%): five of those were obtained with EPA-PMF5, two with ME-2 and one with RCMB. No significant influence of using constraints on the SA

performance was observed.

3.2. Chemical transport models

A total of eleven CTM results were reported by seven groups using five different models, as described in section 2.2. These results were encoded with the prefix “c” (to distinguish them from the RM results) and a capital letter corresponding to each model: CAMX-PSAT (Yarwood et al., 2004; A), FARM (Gariazzo et al., 2007; B), LOTOS-EUROS (Kranenburg et al., 2013; D), EURAD (Strunk et al., 2010; E) and CHIMERE (Menut et al., 2013; F). All the groups reported results apportioning the mandatory set of sources, while the optional set of sources (denoted with the suffix “o”) was reported for only three models (A, B, and D). All the models applied specific parameterisation to quantify the PM attributable to dust resuspension while EURAD and LOTOS-EUROS estimated also the road dust. The latter was the only model which quantified the dust resuspension attributable to agriculture. Parameterisation was also used by CAMX-PSAT, FARM and LOTOS-EUROS to compute the PM deriving from sea salt while the former two and EURAD modelled the fraction deriving from biogenic VOCs. Different versions of the BF/ERI approach were used in this study. FARM adopted a 20% emission reduction while EURAD implemented a 100% emission variation. In both cases, the impacts were normalised to correct the effects of non-linearity. The base case total PM mass was used for normalisation in EURAD while FARM adopted the sum of all source deltas for the same purpose. Since nitrate and ammonium were not apportioned in the CHIMERE result, this model result was used to assess whether the testing methodology was able to detect significant differences between the sources where these species are dominant (e.g. road transport and agriculture). In addition to the performance assessment, three sensitivity runs were executed with CAMX-PSAT to analyse the influence of the spatial resolution and the vertical diffusion coefficient on the model SA performance (section 3.2.6).

3.2.1. CTM model performance evaluation

The temporal evolution of the PM_{10} concentration during the winter period was mainly driven by regional scale processes. Therefore, the selected modelling approach adopting a 7 km horizontal resolution was adequate to reproduce it, as well as to perform the SA analysis. According to the MPE, CTM underestimations were more frequent at high concentrations, suggesting that the overall source contributions deriving from this SA analysis are likely more robust than those of the exceedance days that were poorly reproduced by models. The observed strong underestimation of the organic aerosol may have influenced the reliability of the SCIEs, particularly concerning domestic heating during winter, biogenic sources during summer, and road transport for both seasons. The only exception is EURAD that presented a moderate overestimation of OA in summer. This model also showed a marked overestimation of EC. In this MPE, it was observed that FARM, LOTOS-EUROS and EURAD (only summer) models underestimated sulphate and this could have had an influence on the reconstruction of the source contribution from sources where sulphur is an important component (e.g. energy production and shipping). In summer, the overestimation of nitrate observed in CAMX-PSAT, EURAD and partially CHIMERE models gave rise to a corresponding overestimation of the ammonium concentration that may have influenced the estimation of contributions from agriculture.

3.2.2. Chemical profiles of the CTM sources

In CTMs the PM total mass is reconstructed by processing a set of simplified chemical components some of which are pure substances (nitrate, sulphate, ammonium) while others encompass a relatively wide range of species (organic carbon, elemental carbon and other primary aerosol). The concentration of these simplified chemical components in the different sources were reported in the results. Such chemical profiles of the CTM sources (CTM chemical profiles) were assessed by comparing

them among each other (ff tests; in Fig. S5 of the supplementary material) and with the reference profiles of SPECIATE-SPECIEUROPE (fr tests; in Fig. S5). Fig. S5 shows that the CTM chemical profiles of the sources in the different results were in general quite comparable among each other (ff tests). This holds for most of the anthropogenic sources (such as road transport, industry and energy production), because they were reconstructed by all models on the basis of the same emission inventory and speciation profiles. Lower comparability among results was observed for dust, sea salt (not shown in Fig. S5 because in the optional set only) and agriculture. Dust and sea salt are among the sources not included in the emission inventory that were estimated with a different approach in every model.

The similarity with the SPECIATE/SPECIEUROPE reference profiles (fr tests) was relatively good for biomass burning while it was limited for agriculture, dust, industry and energy production. For the latter three sources, this is probably due to the limited number of species in the CTM chemical profiles and the consequent lack of specific markers. Moreover, for point sources such as power plants or factories, it is unlikely that all the reference profiles are coherent with the specific sources affecting a given receptor site. In Fig. S5, a high variability in agriculture with a considerable share of scores in the rejection area is observed in both fr and ff tests. To explain this behaviour, a more detailed investigation of its source profiles including the six considered chemical species was performed (Fig. 4). In this plot, the differences between the CTM approaches become evident. In the TS approach (CAMX-PSAT, LOTOS-EUROS), the profile is dominated by ammonium with a variable contribution of other organic primary anthropogenic aerosol (OPA). This species consists of trace elements and debris resuspension, the latter being an additional emission term present only in the LOTOS-EUROS model. In the BF/ERI approach (FARM, EURAD), the dominant components are nitrate and ammonium. The differences in the CTM chemical profiles depend on the way the two approaches attribute mass to the sources. The TS approach keeps track of the source from which every chemical component derives. When ammonia from agriculture reacts with nitric acid deriving from NO_x emitted by combustion processes (e.g. traffic), the TS approach attributes the mass of ammonium to agriculture and the one of nitrate to the respective combustion source. In contrast, the BF/ERI approach attributes the mass by computing the difference between the base case and a simulation where the emissions of the relevant source are modified. Since changing agricultural emissions of ammonia leads to a variation in the concentration of ammonium nitrate with respect to the base case, both ammonium (NH_4^+) and nitrate (NO_3^-) are attributed to agriculture with this method.

This analysis of the agriculture CTM chemical profiles evidenced the potential differences in the mass attributed to the sources that may derive from the application of different CTM SA approaches. To evaluate the relevance of those differences, the CTM performance tests were performed distinguishing the TS and BF/ERI results.

3.2.3. Performance tests

In order to test the differences between BF/ERI and TS approaches in the apportionment of PM_{10} , the performance tests were calculated using only the results obtained with CTM TS models (i.e. CAMX-PSAT and LOTOS-EUROS) for the construction of the ensemble reference. This approach was chosen in this analysis because its output is comparable with the one of RM (both are classified as MT source apportionment and their outputs are contributions, see section 1). Another advantage of using CTM TS for the reference is that they provide the same values for comparable sources in the optional and mandatory sets. However, the methodological choice of setting one of the CTM approaches as reference is not to be intended as recognition of a higher quality or hierarchy. Therefore, the performance of each model in the different tests has to be intended as a quantitative measure of the distance among the different results. The reference SCIEs for the sources reported in the CTM results for the MDT and OPT set of sources are shown in Table 6. The time series of the references and the comparison between candidates and references

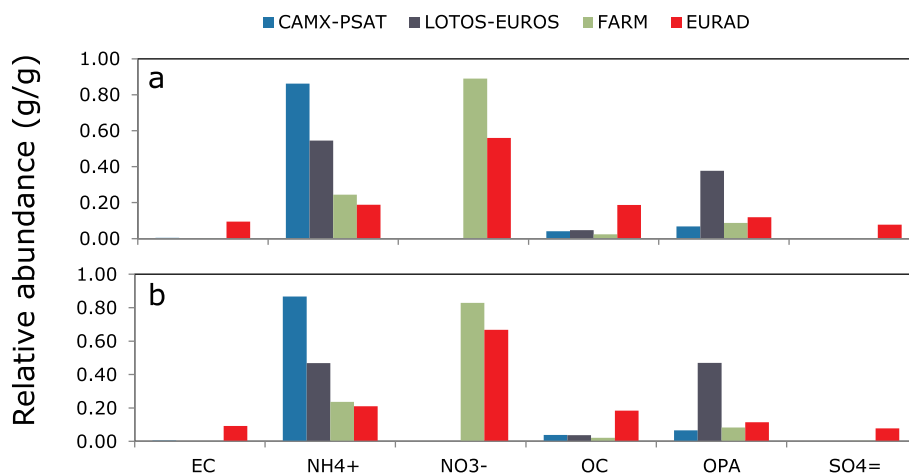


Fig. 4. CTM chemical profiles of the agriculture source reported by the different models expressed as the relative abundance of a species with respect to the source contribution of the source. (a) average of all sites; (b) Lens site.

Table 6

Overall SCIEs and uncertainties of the reference sources ($\mu\text{g}/\text{m}^3$) in the site of Lens used for the z-score tests of the mandatory (MDT) and optional (OPT) set of sources.

source	code	mandatory		optional	
		reference $\mu\text{g}/\text{m}^3$	uncertainty $\mu\text{g}/\text{m}^3$	reference $\mu\text{g}/\text{m}^3$	uncertainty $\mu\text{g}/\text{m}^3$
Energy industry	01 ENI	0.74	0.21	0.74	0.21
Res. & Com. combustion, biomass	02 BIO	0.73	0.29	0.73	0.29
Res. & Com. combustion, fossil	02 OTH			0.34	0.08
Industry	34 IND	1.20	0.56	1.20	0.56
Road transport, exhaust, gasoline	71 RTG			0.10	0.03
Road transport, exhaust, diesel	72 RTD			1.26	0.32
Road transport, non-exhaust wear	75 RTW			0.46	0.21
Road transport, other	07 RTR	1.83	0.35	0.005	0.002
International shipping	08 SHP	0.83	0.27	0.83	0.27
Agriculture	10 AGR	1.77	0.51	1.77	0.51
Dust	11 DST	0.08	0.07	0.08	0.07
Sea salt	11 SLT			2.37	0.68
Biogenic SOA	11 BSO			0.04	n. a.

(scatter plots) are shown in Figs. S6 and S7 of the Supplementary Material, respectively.

As shown in Fig. 5a, b for the reference site of Lens, the majority of the candidates (80% and 87%, respectively) passed the z-score test in the mandatory and optional sets, indicating that the average mass attributed to sources by models was in most cases comparable to the TS reference with a tolerance of 50%. A more detailed evaluation highlighted that candidates of the CHIMERE (cF) result in the power plants, agriculture and ship categories were underestimated. This is explained by the non-apportioned nitrate and ammonium in this result, as discussed above. The biomass burning, dust, and biogenic SOA masses obtained with

FARM (cB) were in the area of overestimation, while the dust source contribution reported by CAMX-PSAT (cA) and the biogenic SOA produced with LOTOS-EUROS (cD) were underestimated with respect to the reference. The explanation is that dust and biogenic aerosol were not present in the common emission inventory and were, therefore, treated differently by every model. In addition, biogenic SOA was not computed by LOTOS-EUROS.

The RMSE_{u} test criteria at the site of Lens were fulfilled by 51% of the candidates of the MDT set and 77% of the OPT set (Fig. 5 c, d). In this test, the distinction between the CTM approaches was very clear. All the candidates obtained with the TS approach were in the area of acceptability and all the candidates outside the target area were produced with the BF/ERI approach. In the MDT set only the candidates for industry and ship of the cB result (FARM) passed the RMSE_{u} test, while the cE (EURAD) and cF (CHIMERE) results presented only one candidate in the target each: industry and biomass burning, respectively (Fig. 5c). A similar situation was observed in the OPT set of sources (Fig. 5d). In this case, the cBo (FARM) were in the target for the ship, industry, diesel and road dust sources.

The results of the performance tests indicate that the differences between the CTM approaches were more evident for the daily SCIE time series (target plots) than for the SCIE averages (z-scores). In order to explore the results more in detail, in Fig. 6 are only shown the scores with the highest differences between TS and BF/ERI (scores outside the target) and the possible causes of the discrepancy are indicated.

In Fig. 6a and b the sources with the highest share of z-scores outside the acceptability area are flagged: dust, biogenic aerosol and biomass. The explanation for the first two sources is that they were treated differently by every model because not present in the common emission inventory (section 2.2.2). In the case of biomass burning, the overestimation of its contribution could be related to the corresponding EC and OC overestimation showed by EURAD when compared against observations and the other models. The analysis of Fig. 6a and b suggests that the differences between the CTM approaches in the average SCIEs (z-scores) are mainly attributable to the different model parameterisation. In the target plots of Fig. 6c and d are shown only the candidates with scores outside the target. Such plots confirm the bias for dust, biomass burning and biogenic aerosol highlighted by the z-score test. Moreover, in these target plots other (anthropogenic) sources obtained with BF/ERI methods appear outside the target: traffic, energy production and agriculture. These three sources are well-known emitters of gaseous precursors (nitrogen oxides, sulphur dioxide and ammonia, respectively), that react in the atmosphere to produce secondary inorganic aerosol (SIA). As already explained in sections 1 and 3.2.2, the two CTM approaches (TS and BF/ERI) allocate SIA in a different way. The

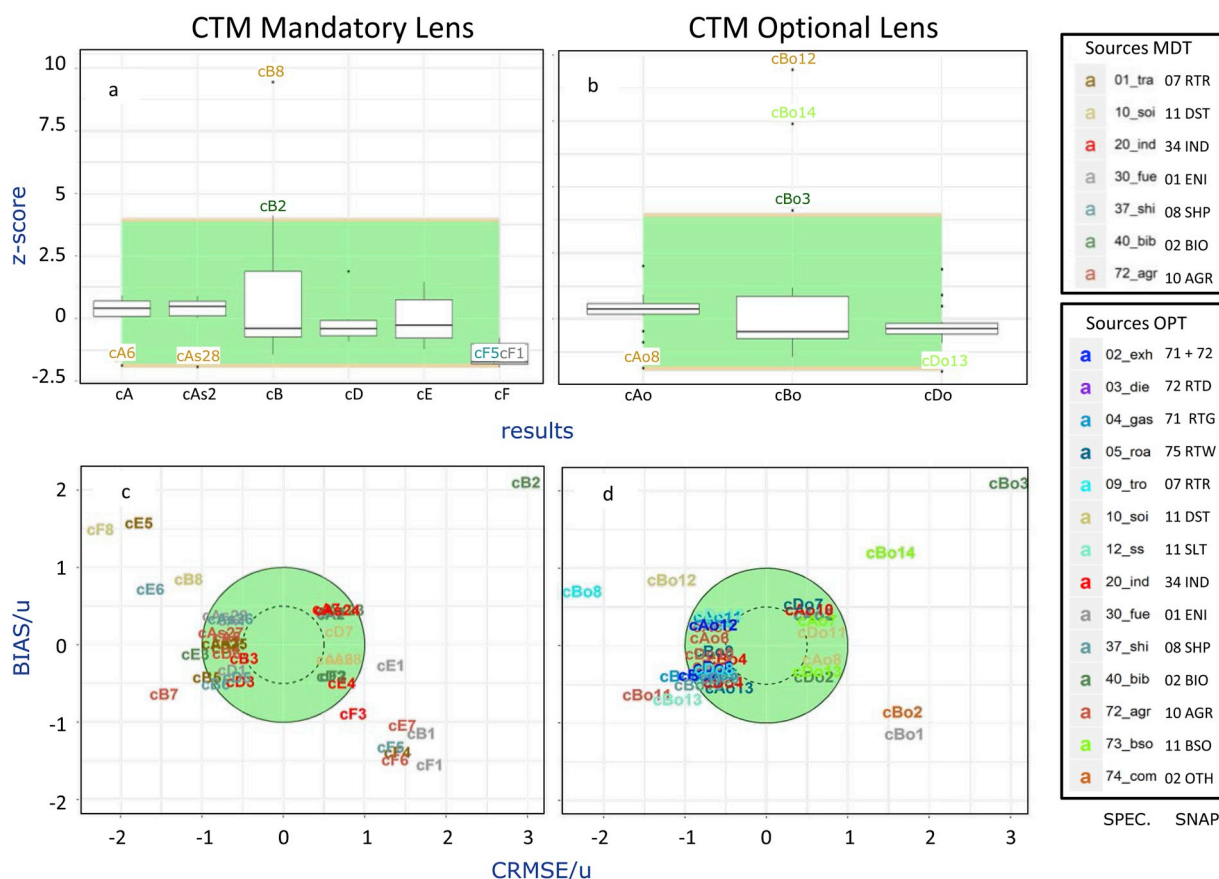


Fig. 5. Performance tests for the CTM apportionment of PM₁₀ at the site of Lens. z-scores for the (a) mandatory (MDT) and (b) optional (OPT) sets arranged by results and target plots for the (c) MDT and (d) OPT sets. The encoding of the candidates is explained in the text (section 3.2). Green background: acceptability area. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

discrepancy between the mass apportioned by TS and BF/ERI to these sources of secondary aerosol suggest that the nonlinearities are not negligible (Clappier et al., 2017). Besides, biomass burning emits, as any combustion process, nitrogen oxides. For that reason, it cannot be excluded that the differences pointed out by the test for this source could be in part attributable to the emission of precursors that are involved in secondary processes. The results of the performance test highlighted in Fig. 6 suggest that the non-linearities causing the differences between TS and BF/ERI are stronger when dealing with SCIEs for short time intervals (daily averages, RMSE_u) than with long-term SCIE averages (six months representing the warm and the cold seasons, z-scores). Another possible explanation for the effect of time resolution on the model SA performance could be the underestimation of CTMs observed during the exceedance days (section 3.2.1).

3.2.4. The case of agriculture

The differences in the CTM chemical profiles of the agriculture PM₁₀ sources detected in the preliminary analysis of the data pointed out a clear influence of the adopted CTM SA approach on the mass attributed to this source (section 3.2.2), highlighting the need for a more thorough analysis. The mass attributed to this source and its time trend varied considerably among the different models. The highest estimation (obtained with CAMX-PSAT; 1.8 µg/m³) was three times higher than the lowest one (EURAD; 0.6 µg/m³) and was associated with the overestimation of NH₄ by CAMx and the effect of normalisation in BF/ERI models. Furthermore, FARM presented the highest levels in the warm period while LOTOS-EUROS and CAMX-PSAT, that presented the most correlated time trends (R=0.6), showed the highest contributions in the cold period (Fig. 7 left).

In the target plot (Fig. 7 right), the TS models (CAMX-PSAT and

LOTOS-EUROS) scored in the acceptability area, while all the BF/ERI models did not. Hence, the RMSE_u test confirms that the results obtained with the two CTM approaches for the agriculture source are not comparable as hypothesised in section 3.2.2.

3.2.5. Inter site variability in the CTM SA

In Fig. 8, the relative average SCIEs for PM₁₀ at every site are displayed for the eight CTM results evaluated in the performance tests. The results of three models (CAMX-PSAT, FARM and LOTOS-EUROS) were available for both the MDT and the OPT sets of sources while for EURAD and CHIMERE results were available only for MDT.

According to the share of sources four groups of sites were identified: the megacities (London and Paris), the medium-sized cities (Brussels, Ghent, Lens), the coastal sites (Calais and Le Havre), and the background sites (Rural backgd 1 and 2, and suburban bkgd). The only exception is EURAD where all the sites presented exactly the same relative split of sources. Therefore, in this result the difference between sites was only modulated by the different PM concentrations. In the megacities, the predominant source was traffic. At the coastal sites, the dominant source was sea-salt (present only in OPT), which was also significant at the other types of sites. In the medium-sized cities, traffic was sometimes dominant (FARM, LOTOS-EUROS) and sometimes was comparable to the other sources. On the contrary, at the background sites there was no clearly dominant source. Traffic, domestic combustion, biomass burning, shipping and industry all contributed in variable but still relatively comparable proportions. The resemblance in the share of sources between medium-sized cities and background sites suggests that due to the high density of urban areas in the domain, the contribution of sources attributed by the models within the cities and outside did not differ significantly. However, this could be also due to limitations in the

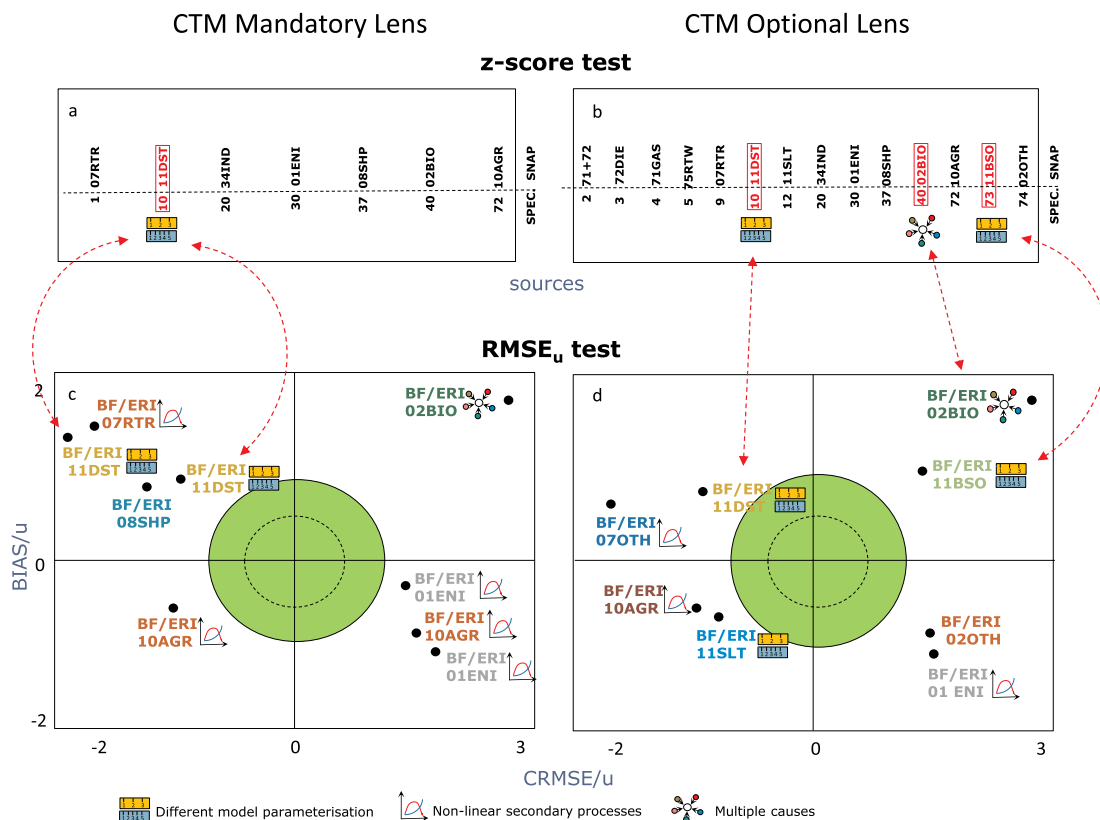


Fig. 6. Performance tests for the CTM apportionment of PM₁₀ at the site of Lens. The sources of the MDT (a) and OPT (b) sets with high share of z-scores out of the acceptability area are flagged (border). Target plots for the MDT (c) and OPT (d) sets display only the candidates outside the target. The icons next to the sources indicate the possible causes of discrepancy. BF/ERI = brute-force/emission reduction impact. Green background: acceptability area. Correspondence between SPECIEUROPE (Table 2) and SNAP nomenclatures (Table 3) is indicated. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

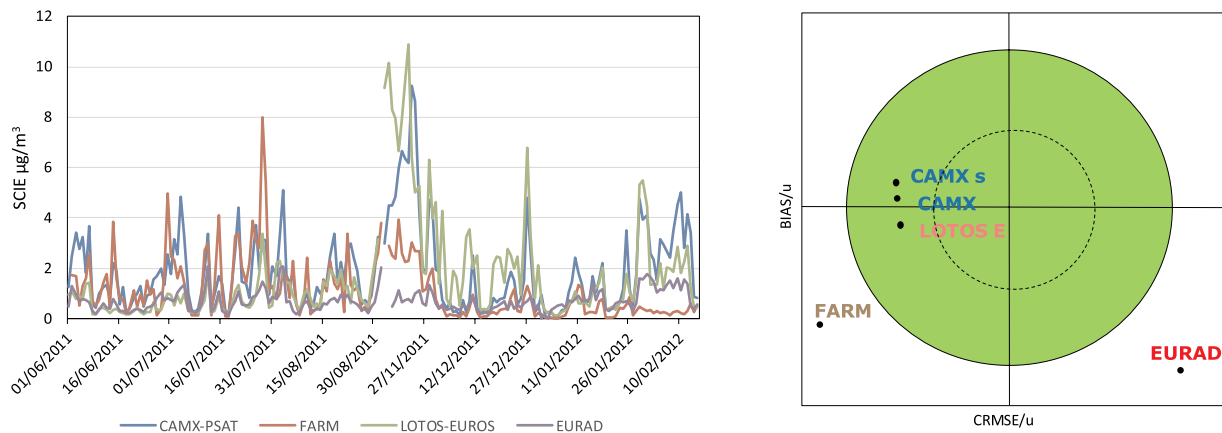


Fig. 7. Time trend of the agriculture SCIEs for PM₁₀ reported by every model and their corresponding RMSE_u scores in the target plot. In the target plot the names of some models are abbreviated: LOTOS E: LOTOS-EUROS, CAMX: CAMX-PSAT and CAMXs: sensitivity run of CAMX-PSAT. Green background: acceptability area. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

emission inventory and to the spatial resolution of the simulations that curbed the identification of hot spots associated with local sources.

In the OPT results, it is possible to observe the split between the different components of the traffic. Diesel exhaust is pointed out as the most important component by FARM and CAMX-PSAT while road dust is put forward by LOTOS-EUROS. In the latter model, however, the OPT traffic source (sum of 7.1, 2, 3, 4 and 5; Table 3) did not match the total traffic obtained by the MDT set. This difference could be due to the

specific approach used in this model to estimate the road dust component that may have been set up differently in the OPT and the MDT runs. As mentioned in section 3.2, the low values of CHIMERE are due to the non-apportioned nitrate and ammonium mass. This model attributes the highest share to dust, while CAMX-PSAT and EURAD apportioned very little or no mass to this source.

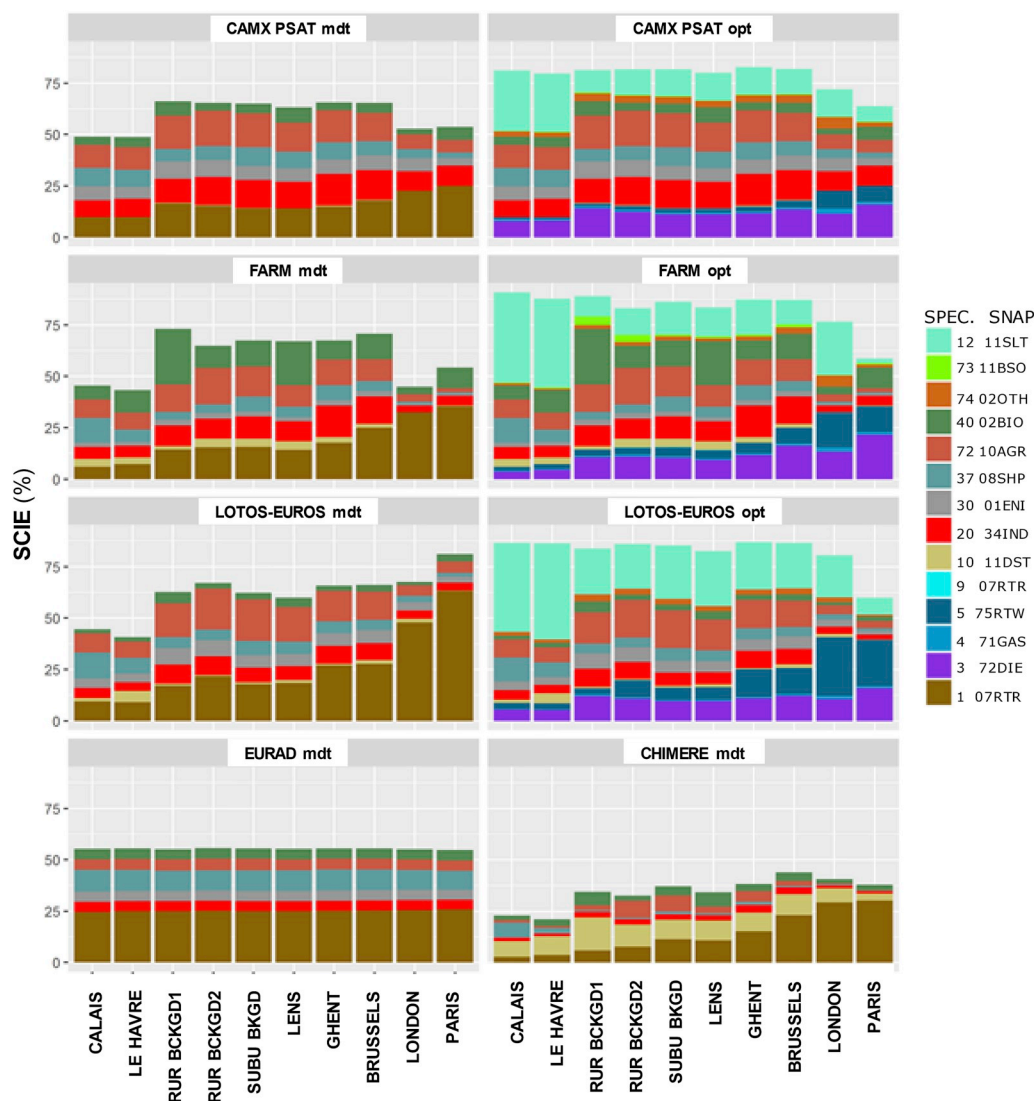


Fig. 8. Relative average SCIEs for all the PM_{10} sources at every site reported in the eight CTM results used for the performance evaluation. The mass of the “non-apportioned” source (99 OTH) is not shown.

3.2.6. CTM sensitivity analysis

Three configurations were set up to perform a sensitivity analysis using the TS approach model CAMX-PSAT: the base case and two additional runs. The first sensitivity run (VD) included computations of the vertical diffusion coefficients used in the CMAQ scheme (Byun, 1999) which are different from the YSU scheme (Hong et al., 2006), adopted in the base case. The second sensitivity run (DSR) was performed switching off the Lens inner domain (7 km grid step), thus running CAMX-PSAT only over the EU wide domain at 20 km grid step. The goal of the first run was to assess the how vertical diffusion may impact on the ability of the model to quantify different sources (e.g. local sources versus long range transport) while the second run focused on the influence of horizontal resolution on the SCIEs of local sources.

The analysis was performed at three receptor sites: Lens, Paris, and RUR BCKGD2. Lens was the reference site, while Paris and RUR BCKGD2 were selected because they were influenced by very different emission strengths. Paris is a megacity, where the influence of local emissions (e.g. transport and domestic heating) are higher than in the rural site which is mainly subject to the influence of agriculture and secondary pollution. The results presented here concern only the OPT set. The VD run was performed only for the winter period because it is characterised by a stronger influence of vertical diffusion processes on pollutant

concentrations.

The increased model grid cell dimension in an area close to localised primary emissions (traffic) like the Paris site caused a reduction in the concentrations of primary pollutants associated with that source (Fig. 9, left). A PM_{10} concentration decrease ($-4 \mu\text{g}/\text{m}^3$) for DSR matched a decrease of -32% in elemental carbon (EC), -38% in primary organic aerosol (POA), and -29% in other primary anthropogenic aerosol in the PM_{10} fraction (OPA-10) compared to the base case.

When comparing the performances of CAMX-PSAT using two different grid steps (with all models in the reference), it was also observed that the contribution of traffic was underestimated when using a lower spatial resolution (Fig. 9, right). No significant changes were observed in the other tested sources (industry, energy production, biomass burning and agriculture).

CAMX-PSAT proved to be less sensitive to the variation of vertical diffusion coefficients. The CMAQ algorithm used in the VD run gives rise to stronger vertical diffusivity with respect to YSU scheme used in the base case, thus inducing a decrease in ground level concentrations. In the VD run, the most relevant variations took place in the primary species at Paris. The concentration reductions were, however, modest: PM_{10} - $0.5 \mu\text{g}/\text{m}^3$, OPA-10 - 3%, EC - 7%, and POA - 4% (Fig. 9, left).

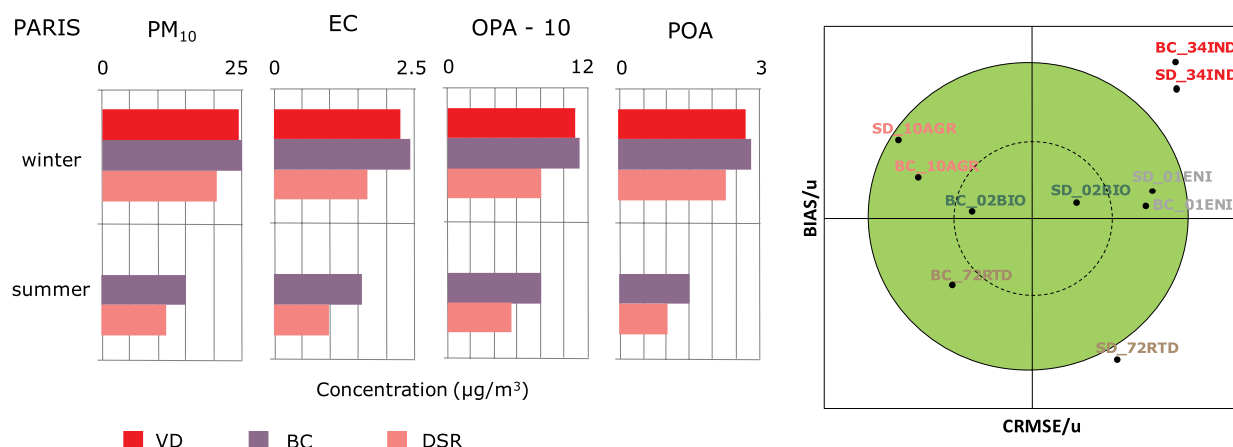


Fig. 9. Variation in concentrations of primary chemical species (left) between the base case (BC) and the sensitivity runs VD (vertical diffusion) and DSR (decreased spatial resolution). The impact of the spatial resolution on the model performance is also shown (right). Green background: acceptability area. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

3.3. CTMs vs RMs

The comparison between the two families of models (RMs and CTMs) was accomplished only for the reference site of Lens. For the interpretation of the results, it was necessary to consider that the estimation for RMs refers to the specific site where the dataset of measured PM₁₀ was collected, while the CTMs provide an average estimation for the grid cell, roughly corresponding to a 6–7 km grid step, containing the monitoring site. According to our understanding, the representativeness of the site (background) is coherent with the dimension of the grid cell. It is also worth mentioning that RMs output time resolution is that of the samples (daily in our case with samples collected every third day), while CTMs generate results with an hourly time resolution, which were averaged to match the RM data.

Another aspect to consider when comparing RMs and CTMs is the different level of detail about the chemical composition of the PM managed in the two families of models. While RMs are based on a relatively detailed information about the chemical composition (in general, 20 or more species derived from chemical analyses), there is a limited number of chemical families in the CTMs (typically 6 or 7), depending also on the completeness of the emission inventories.

Finally, when matching the two series of sources, it should be considered that most of the inorganic ions, ammonium, sulphate and nitrate are attributed by RMs to one (SIA) or two secondary aerosol categories (ammonium nitrate and ammonium sulphate), while the CTMs allocate these chemical species to the corresponding precursor sources. To account for the different way in which these inorganic ions are handled by the two families of models the comparison between them was first accomplished subtracting these inorganic ions from the CTMs categories and allocating them in a fictitious CTM category named SIA. This correction was discarded for two main reasons. Firstly, despite RMs allocate most of the ammonium, sulphate and nitrate to the SIA category or subcategories, these ions are also present in many other sources. Secondly, when subtracting these ions from the CTM original sources, the contributions of a number of them fell down to values near zero or even negative in some BF/ERI results, as a consequence of the non-linearity problem (see section 1; Table 1). For instance, when applying this correction to the CTM category energy industry, its SCIEs decreased on average by 97% and 100% in the MDT and OPT sets, respectively. Other heavily altered categories were ship (-89%, -93%), exhaust (-87%), agriculture (-79%, -77%) and industry (-58%, -52%). To avoid introducing distortions, the comparison between RM and CTM was then performed on the SCIEs as reported by participants and the different handling of secondary inorganic ions by the two families of models

considered in the interpretation of the results.

In Fig. 10 the average source contributions estimated by RMs and CTMs using both TS (CAMX-PSAT, LOTOS-EUROS) and BF/ERI (FARM, EURAD) approaches for the MDT and OPT sets of sources are compared. To support the interpretation of results, in Fig. 10 are also plotted sources for which the comparison was not possible: agriculture (only CTMs), primary organic (only RMs) and secondary inorganic aerosol (ammonium nitrate and ammonium sulphate, RMs).

The primary sources with the highest SCIEs in RMs are: exhaust, soil (dust), ship, road dust and biomass burning. The coefficient of variation for the RMs varied between 16% and 77% with the exception of industry where it reached 107%. This latter source has been identified as the most critical one in the evaluation of RMs due to the dispersion of the reported SCIEs (see section 3.1.1). Traffic, agriculture, industry and biomass burning are the most important sources in the CTM MDT set while marine, agriculture, exhaust, biomass burning and industry show the highest PM shares in the CTM OPT set.

The CTM SCIEs are in the majority of cases lower than the RM averages for the corresponding source (Fig. 10). Nevertheless, there are cases where CTMs are comparable to or even higher than RMs. In the MDT set the SCIE of traffic reported with EURAD was 40% above the one of the RM average while the other models were 40% to 60% below such reference. Similarly, in industry and biomass burning there was one CTM result reporting SCIEs within 10% of the RMs average (CAMX-PSAT and FARM, respectively) while the others reported much lower values. In the OPT set, the marine source is the one where the CTM SCIEs are the closest to the RM average. FARM and CAMX-PSAT were less than 15% from the RM average, likely due to their parameterisation based on observations. On the contrary, LOTOS-EUROS presented values 70% higher than RMs, probably due to a parameterisation based on observations in areas where the contribution of this source is different than in Lens.

An explanation for the high levels of EURAD compared to the other CTMs and RMs in traffic PM may be the overestimation of EC and organic aerosol in the base case. However, another possible explanation is associated with the used BF/ERI approach. As explained in section 1, the non-linear behaviour of sources leads to inconsistencies between the base case PM mass and the sum of the masses of all the sources. Such inconsistencies increase with the percentage of emission variation used to compute the impact of the sources (Thunis et al., 2018). In the BF/ERI results reported for this study, the impact of the single sources was normalised to reconcile their sum with the total PM₁₀ mass. However, this method does not correct the relative influence of the sources because they are affected by non-linearity to different extents. The high

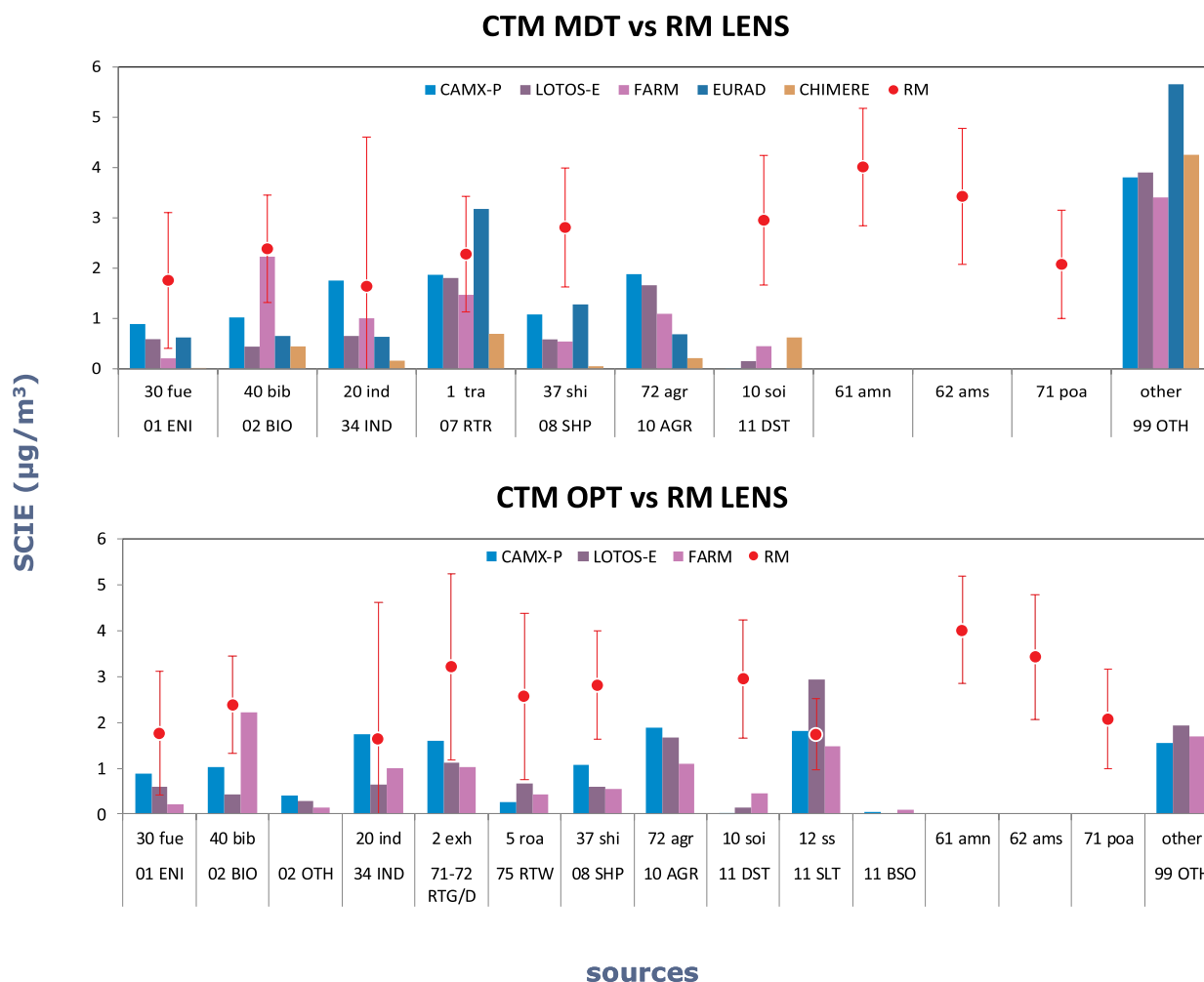


Fig. 10. Average SCIEs of the sources reported in the present intercomparison obtained with RMs compared with those deriving from CTMs for both the MDT (top) and OPT (bottom) sets of sources. The sources are encoded according to both the SPECIEUROPE (Table 2) and SNAP (Table 3) nomenclatures. The error bars indicate the standard deviation of the RM SCIEs.

level of emission variation (100%) and the limited effectiveness of the normalisation may explain the high traffic SCIEs of EURAD compared with the other models (also observed in the 99 OTH category). The same could be argued for the biomass burning SCIEs of FARM. In this model, however, the non-linearity is expected to be lower because the emission variation was only 20%. In addition, its average SCIE is the only one comparable with those of RMs. The higher TS shares of agriculture with respect to those of BF/ERI models may be partially explained by the normalisation that did not correct the effects of non-linearity and by the overestimation of NH_4 by CAMX-PSAT. In addition, the TS models have allocated higher mass of OC and in one case (LOTOS-EUROS) of OPA because the dust resuspension deriving from the agricultural activities was allocated to this source (section 3.2.2).

Although most of the main SIA mass is not allocated by RMs to their sources, their SCIEs are higher than those attributed by CTMs even for sources where these ions have a dominant contribution, such as energy production, shipping and industry. A possible explanation is that the CTM underestimation of the gravimetric mass overcompensates the fact that they allocate also the secondary inorganic fraction to these sources. This hypothesis is coherent with the outcome of the MPE suggesting that the PM mass underestimation by CTMs is partly due to problems in reconstructing the particulate organic aerosol (section 3.2.1).

The plots in Fig. 10 provide a preliminary understanding about the bias between RMs and CTMs in the different sources. However, to

establish whether two SCIEs are comparable or not, the plots in Fig. 10 are not enough and distance indicators with acceptability criteria are required. To that end, the comparison between RMs and CTMs was performed using the methodology described in sections 2 and A.1 of the Appendix. The performance tests were performed setting the CTM sources as candidates and the SCIE of the corresponding RM source category as reference value. In Fig. S8 of the supplementary material the z-score tests for the MDT and OPT sets are displayed. Even though all the CTMs tend to underestimate the SCIEs when compared with the RM values, the majority of the candidate sources fall in the area of acceptance.

In Fig. S9 of the supplementary material the target plots split by source set (MDT and OPT) and by CTM approach (TS and BF/ERI) are shown. The dispersion of the sources in the target plot is relatively uniform between the two sets of sources (for those present in both of them) and between TS and BF/ERI. This is explained by the fact that the differences between CTMs and RMs are too large compared to those between CTMs. Therefore, in this test where the acceptability criteria are dependent on the reference SCIE (RMs in this case) the differences between CTMs are less evident. Another aspect to take into consideration is the variability between the CTMs within each approach as shown in Fig. 10.

In the z-score tests (both MDT and OPT sets), 83% of the candidates rank in the acceptability area while in the RMSE_u test only 34% and 25%

of the candidates of the MDT and OPT sets, respectively, pass the test. To summarise the main findings of the performance tests, a synthesis of the comparison between RMs and CTMs for both MDT and OPT sets is presented in Fig. 11.

From the z-score plot, three groups of sources can be distinguished (Fig. 11, left). The first group includes the sources for which the CTM SCIEs are comparable with those of RM (marine and industry) or show little underestimation (traffic). Although their CTM SCIEs are lower than those of RMs, the second group includes sources whose differences fall within the acceptability range of this test (exhaust, energy, ship and biomass burning). The third group includes sources with part of the SCIEs outside the acceptability area (road dust, dust). These conclusions are confirmed and complemented by the $RMSE_u$ test, where industry and traffic (first group) fall completely inside the acceptability area (Fig. 11, right). Despite no observed bias, marine (first group) falls outside the acceptability area due to the lower correlation and amplitude of the time series with respect to the reference. Exhaust and road dust in this test show a lower bias compared to the z-score test, likely due to the effect of the relatively high RMs dispersion (the $RMSE_u$ scores are normalised by the uncertainty of the reference). The remaining sources present, in addition to poor correlation with the RM time series, different extents of negative biases as already pointed out in the z-score test. In addition, the time series of road dust, dust, and biomass burning reported by CTMs also present differences of amplitude with respect to the RM.

The good performances of the industry and exhaust sources could be partially due to the relatively high uncertainty of the RM SCIEs for these categories. The poor performance of marine in this test is indicative that, despite average SCIEs being quite similar in RMs and CTMs, their time trends are poorly correlated and have different amplitudes. In this exercise, dust was handled differently by every CTM model (section 2.2.2) because not represented in the emission inventory. This is likely the cause of the sizeable bias component between RMs and CTMs. A possible explanation for the observed bias in road dust could be an underestimation in the input emission inventory. Moreover, in RMs, this source incorporates all the resuspended PM components, including dust deposited on the road and abrasion emissions (pavement, tyre and brake wear) while the CTM source (75 RTW) represents only the latter. The cause of the poor $RMSE_u$ values in ship and power plants is likely associated with the strong bias already evidenced between the two families of models in the previous tests (Figs. 10 and 11 left). In the biomass burning source, a problem of amplitude, in addition to a

moderate bias has been identified. It is well known that this source undergoes considerable seasonal variations and episodes take place during the cold season. The result of the $RMSE_u$ test points out that the two families of models reproduce differently both the extent and timing of such variability.

3.4. Influence of the site on the model SA performance

A summary of the SA performance tests for all the sites is given in Fig. 12. In this plot, CTMs are compared with their ensemble reference for each site. At the Lens site, the performance of RMs as well as those of CTMs compared with RMs is also reported (circles).

In general, the variation in the model performance between sites is less than 10% and 15% for the z-scores and $RMSE_u$, respectively. Such homogeneity in the SA model performance suggests that the geographical patterns influencing the allocation of sources depend more on the variations in the input data (emission inventory and meteorological fields) than on the differences between models. The results obtained with the OPT set of sources are always better than those of the MDT set suggesting that the models are more comparable when accomplishing a more detailed apportionment. The overall performance of the RMs at the site of Lens is comparable with the one of the CTMs with the OPT set of sources. As expected, differences are more evident when comparing the two families of models (RMs and CTMs, blue empty circles) than when comparing models within each family, which is reflected by the lower number of successful candidates in the $RMSE_u$ tests (Fig. 12 bottom).

4. Key findings of the intercomparison and discussion

The RMs applications showed a convergence in the use of the EPA-PMF5 tool, in part due to its free availability and user-friendliness and in part due to the good performances of EPA-PMF tools in past exercises, which led to more comparable results with respect to previous intercomparisons (Belis et al., 2015b).

It appears that the industry source in RMs needs a better definition because it encompasses a wide range of different local emissions with different chemical fingerprints that, placed under a common category, led to considerable variability between results. More specific allocation within this category (e.g. in subcategories), by taking more advantage of source profiles repositories for instance, would lead to a better

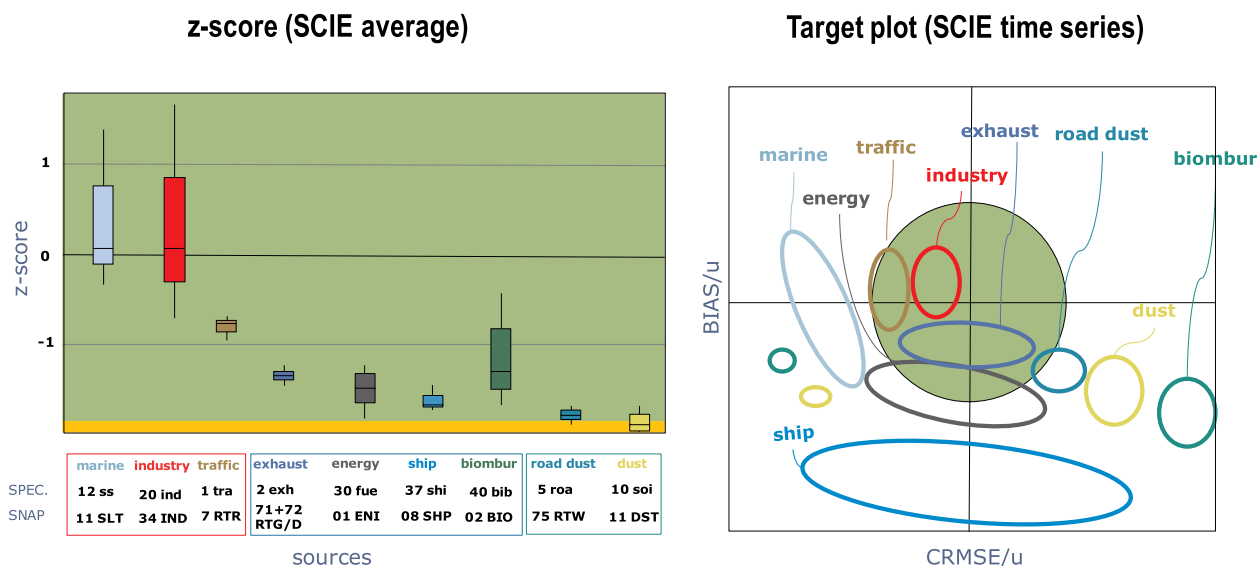


Fig. 11. z-scores (left) and target plots (right) for the comparison between RMs and CTMs at Lens. In the target plot, the sources are represented schematically using the boundaries of their scores. Green background: acceptability area. The sources are encoded according to both the SPECIEUROPE (Table 2) and SNAP (Table 3) nomenclatures. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

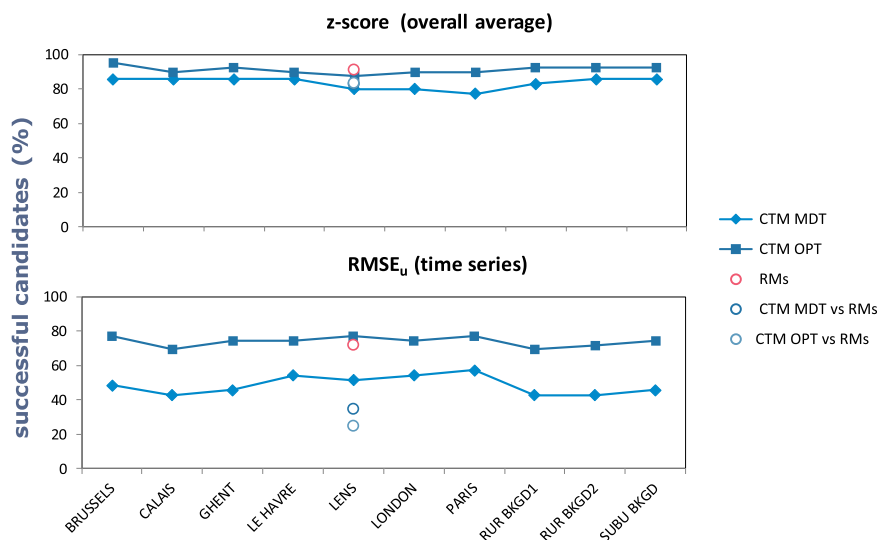


Fig. 12. Synthesis of the intercomparison CTM performance tests (expressed as percentage of successful candidates) for all the studied sites (solid markers). The comparison between RMs and between CTMs and RMs, available only for the city of Lens, are also displayed (empty circles).

identification and consequently to a more accurate quantification of the contributions. In this regard, the characterisation of new markers or chemical fingerprint remains necessary to better take into account the specificity of a wider range of industrial sources.

In this exercise, the availability of an extended set of organic markers, supported the allocation of primary biogenic organic aerosols, a category for which more scientific evidence is necessary in order to be included in the list of the sources that can be subtracted from the total PM for the calculation of the number of exceedances in the framework of the Air Quality Legislation (Directive 2008/50/EC; European Commission -, 2011).

The high number of RM groups involved in the study made it possible to investigate the influence of their experience on the SA performance. RM practitioners declaring to have conducted more than 10 studies (24%) always showed good performances while the less experienced ones presented more variable scores.

Thanks to the set-up adopted in the CTM performance tests it was possible to quantify with real-world data the differences between the CTM TS and BF/ERI approaches in the allocation of SIA under non-linear situations, confirming previous estimations made with theoretical examples (Clappier et al., 2017). The most emblematic source from this point of view is agriculture, whose ammonia emissions contribute to the formation of the most important components of the SIA, ammonium nitrate and ammonium sulphate. Considerable differences in the total mass attributed to this source and relative CTM chemical profiles were observed between BF/ERI and TS approaches. Similarly, differences were also identified in the performance tests ($RMSE_u$) in other SIA precursor-emitting sources such as road traffic, power plants and to a lesser extent biomass burning.

The abovementioned differences between TS and BF/ERI were evident in the CTM performance tests when analysing the SCIE daily time averages while they were not detected in test on the overall SCIE averages. A possible explanation is that the non-linearity that gives rise to discrepancies between models depends on the considered time window (i.e. daily for z-scores; seasonal, annual for $RMSE_u$).

The specific emission inventory, including details on the type of fuel, and the comparability table for the sources prepared for this exercise, made it possible to quantitatively assess the differences between CTMs and RMs SA results. However, it should be recalled that the two types of models use different source definitions and, therefore, only a subset of sources was selected for comparison.

The comparison of CTM source contributions with the measured mass and the RM SCIEs highlights a generalised difficulty of CTMs in apportioning all the PM_{10} mass. This problem has been associated with the underestimation of this pollutant concentration in the base case model results, especially during high pollution episodes, due to the poor organic PM fraction reconstruction and to problems in reproducing the dust PM fraction.

Since traffic (including both exhaust and resuspension components) and industry passed both performance tests, they were the most comparable sources between CTMs and RMs. Although exhaust (only in OPT) also passed these tests, the bias in the z-score was not negligible. The marine source obtained good z-scores (little bias) but showed differences in the amplitude of the time series between the two families of models in the target plot. The good comparability between RMs and CTMs for industry may look odd when considering the high variability of the RM SCIEs for this source. However, such comparability is clear in both (a): the z-score tests, where the RM uncertainty is not considered, and (b) the $RMSE_u$ tests, where the uncertainties of the RM SCIE are accounted for and lead to scores very close to the centre of the target plot. A possible explanation is that despite the considerable dispersion of the RM SCIEs for this source, their average is not biased. On the contrary, the most critical sources in the comparison between CTMs and RMs are dust and road dust, likely due to the different way in which the former was reconstructed by the different models and a possible underestimation of the latter in the emission inventory because only the PM deriving from abrasion is included while the re-suspended dust is not considered. These sources are only partially represented in the emission inventories because their quantification is not formally required by the official reporting schemes in the EU. For example, road dust resuspension is considered a re-emission and not a primary PM emission source (Denier Van Der Gon et al., 2018). Significant underestimation, albeit within the tolerance of the test, is observed for shipping, power plants, biomass burning and to a lesser extent, traffic exhaust.

The differences between TS and BF/ERI were not evident when comparing CTMs and RMs. The reason is that the objective of the performance tests is to compute the distance between the reference and every candidate and not those between candidates. Consequently, the acceptability criteria are set on the basis of the reference SCIE (RMs in this case). As shown in Fig. 10, most of the RM-CTM differences are much higher than those between TS and BF/ERI and, therefore, the latter are less visible in the CTMs-RMs test. It follows that the

performance tests among CTMs (section 3.2.3) are the most appropriate to assess the differences between CTM approaches.

According to the variation in the CTM share of sources it was possible to identify four groups of sites: megacities, medium-sized cities, coastal sites and background sites. Although differences among these groups were perceptible, the medium-sized cities and background sites presented a relatively homogeneous share of sources. Such similarity suggests that the contribution of sources attributed by CTM within the cities and between them do not differ significantly. This could be due either to the high density of urban areas in the studied domain or to the spatial resolution of the simulations that limited the analysis of local hot spots. The abovementioned homogeneity is also reflected in the CTM SA model performance suggesting that the geographical patterns influencing the allocation of sources depend more on the variations in the input data (emission inventory and meteorological fields) than on the differences between models.

The sensitivity analysis with CAMX model suggests that adopting spatial resolutions lower than the one used in this study could lead to the concentrations of PM₁₀ and associated primary pollutants (e.g. traffic) being underestimated by 20–30%, particularly in urban areas. This would affect the ability of the models to apportion the mass to this kind of sources properly. More work is necessary to better understand the implications of this results.

Despite the CHIMERE base case was comparable to the other models, the methodological choice adopted in this particular application (not apportioning nitrate and ammonia to their sources) impacted on the comparability with the other models' results. As a consequence, the scores of sources for which these chemical compounds represent a high share of the mass (power plants, shipping, agriculture) were identified as outliers in the performance tests.

5. Conclusions

The high quality of the input data for both RMs and CTMs and the considerable number of SA results (49) provided the basis to build up an unprecedented database, with key information to support an extensive analysis of RMs and CTMs methodologies used in Europe for SA applications related to PM₁₀ SA.

The adopted methodology proved to be efficient to compare and characterise the SA performances of models of the same family (RM or CTM), different approaches within the CTM family, and CTMs vs RMs. The differences highlighted in the tests provided evidence to understanding the implications that the use of different types of models have on the SA output contributing to a better interpretation of the SA results. Since this is the first application of this methodology to CTMs and RMs-CTMs, further work is needed to create a record of results, including different models and input datasets, to tune up the tests' acceptability criteria.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.aeoa.2019.100053>.

Appendix

A1 Evaluation methodology

The methodology adopted in this study encompasses three types of tests: (a) complementary tests to provide information about the SA result as a whole, (b) similarity tests that aim at establishing whether candidate sources belong to a source category on the basis of their chemical composition and time series, and (c) the performance tests to assess whether the mass of a pollutant attributed to a source category is coherent with the reference value on the basis of pre-established quality criteria (Table A1).

Despite a considerable dispersion observed in some sources (e.g. industry), RMs presented comparable results likely due to the use of the same model (EPA-PMF5) in the majority of cases. More effort is needed to better define sources by developing the existing repositories of source profiles to be used as reference in SA analyses using either RMs or CTMs. Differences were measurable between the TS and BF/ERI CTM approaches, in particular for secondary inorganic aerosol deriving from precursors emitted from different sources (energy, traffic, agriculture). In general, the RMs presented higher SCIEs than CTMs likely due to the underestimation of the total PM₁₀ by the latter, especially during high pollution episodes, which has been partly associated with problems in reconstructing the PM₁₀ organic fraction. In the case of dust and road dust sources, the CTM-RM comparison highlighted areas for improvement in the emission inventories.

The relevant differences in the comparability between RM and CTM and between TS and BF/ERI approaches SCIEs confirmed that the evaluation of the CTM results should not be limited to the total concentration but also extended to the single emission categories. In this regard, more work is needed to establish how this can be achieved in the common modelling practice.

The quantitative evaluation of the RM and CTM SA model applications using a common methodology achieved in this study is a first step towards the joint application of the two families of models to take advantage of the strengths of both of them.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors warmly thank J. Vercauteren (VMM) for providing the CHEMKAR dataset. The CARA program was funded by the French Ministry of environment. IMT Lille Douai participates in the CaPPA project funded by the ANR through the PIA under contract ANR-11-LABX-0005-01, the "Hauts de France" Regional Council and the European Regional Development Fund (ERDF). The C2TN/IST author gratefully acknowledges the FCT support through the UID/Multi/04349/2013 project. J.L. Jaffrezo would like to thank F. Donnaz, F. Masson, and S. Ngo for the chemical analyses of the Lens samples performed at IGE (ECOC, ions, sugars). These were possible on the Air-O-Sol analytical platform supported by Labex OSUG@2020 (ANR10 LABX56). A. Angyal was supported by National Research, Development and Innovation Office – NKFIH, contract number PD 125086. H. Jorquera acknowledges support from Grant CONICYT/FONDAP/15110020. P. Thunis commented on an early version of the manuscript.

Table A1
Evaluation methodology for source apportionment model results (Belis et al., 2015a)

Types of test	Test description
1. Complementary tests: provide ancillary information about the model output performance	1a. Mass apportionment: compares the sum of the mass of the candidates of one result with the total PM mass. 1b. Number of candidates: compares the number of candidates with the average number in all the results.
2. Preliminary tests: evaluate the attribution of candidates to a given source category	2a. Chemical profiles: compare the candidate chemical profiles with those of reference profiles for a given category. 2b. Time series: compare the candidate time trends with those of reference time series for a given category. 2c. Contribution-to-species: compare the source allocation for every single PM component expressed in percentage in the candidate and the reference.
3. Performance tests: evaluate the candidates SCIEs by comparison with a reference value	3a. z-score (ISO 13528): test the bias of the overall SCIE with respect to a pre-defined uncertainty for proficiency test (σ_p). 3b. Uncertainty weighted root mean square error (RMSE _u): test the bias, amplitude and phase of the SCIE time trends.

The reference SCIE values for the performance tests were obtained from the average of the candidates excluding the outliers identified with the similarity tests. The z-score indicator (ISO 13528, 2005) was applied to the average SCIEs over the entire dataset reported by participants. The uncertainty for proficiency test (σ_p) was set to 50% of the reference SCIE by analogy with the model quality objectives for PM₁₀ annual mean laid down in Directive (2008)/50/EC. The acceptability interval for the z-scores was -1.96 and 3.99 obtained using a synthetic dataset where the SCIEs of the sources were known. A kernel distribution (R package ks v. 1.9.2) was fitted to more than 200 unbiased z-scores after removing outliers and extracting the 0.005 and 0.995 percentiles to define areas with the same probability density as those used in the abovementioned standard (Belis et al., 2015a). Moreover, the SCIE time series for every candidate were evaluated using the root mean square error (RMSE_u) normalised by the uncertainty of the reference (u) at every time step (Jolliff et al., 2009; Thunis et al., 2012). The uncertainty of the reference was estimated as the standard deviation of the SCIE reported by participants. The RMSE_u values ≤ 1 are considered indicators of good performance.

A2 Glossary of source apportionment intercomparison terminology used in this study

Candidate: PM₁₀ source reported in one source apportionment result.

Contribution-to-species: mass of each of the PM₁₀ chemical components attributed to each candidate expressed as percentage of the chemical species total mass.

Participant: is a single or group of practitioners who deliver a result in an intercomparison exercise.

Reference: is the average of the SCIE of the candidates reported by participants for the same source category.

Result: is the output of a source apportionment model including a list of candidates with the estimation of their contribution or impact to the total PM₁₀.

Root mean square error uncertainty normalised (RMSE_u): is the RMSE weighted by the uncertainty of the reference and is used to assess the performance of a candidate SCIE time series.

Source profile: is the relative abundance of every PM₁₀ component of a source expressed as a proportion of the source total mass. It could be measured directly at the source (measured source profile) or be the result of a SA analysis (derived source profile).

Source category: is a set of sources or processes pooled under a single class due to their chemical or temporal affinities.

Source contribution or impact estimate (SCIE): is the amount of mass attributed to one candidate in one source apportionment result either for the single sample (time step) or for the overall average value.

z-score: is the difference between the candidate and the reference SCIE weighted by the uncertainty for proficiency testing, which is set to 50% of the reference. It is used to evaluate the performance of the candidate SCIE overall average.

A3 Most common abbreviations used in the text (except intercomparison methodology)

BF	brute force approach (CTM)
CTM	chemical transport model
EEA	European Environment Agency
ERI	emission reduction impact approach (CTM)
MDT	mandatory set of sources for CTM
MPE	model performance evaluation
PMF	positive matrix factorisation
RM	receptor model
OPA	other primary aerosol
OPT	optional set of source for CTM
POA	primary organic aerosol
SIA	secondary inorganic aerosol
SNAP	Selected Nomenclature for sources of Air Pollution
SOA	secondary organic aerosol
TS	tagged species approach (CTM)

Table A2

List of species contained in the input dataset used for receptor models.

Chemical group	N	Species list
OC, EC	2	Organic carbon (OC), elemental carbon (EC)
Ions	8	Cl ⁻ , NO ₃ ⁻ , SO ₄ ²⁻ , Na ⁺ , NH ₄ ⁺ , K ⁺ , Mg ²⁺ , Ca ²⁺
Trace elements	25	As, Ba, Cd, Ce, Co, Cs, Cu, La, Mn, Mo, Ni, Pb, Rb, Sb, Sr, V, Zn, Al, Ca, Fe, K, Mg, Na, S, Ti
Anhydrosugars	3	Levoglucosan, Mannosan, Galactosan
PAHs	15	Phenanthrene (Phe), Anthracene (An), Fluoranthene (FlA), Pyrene (Pyr), Retene (Ret), Benzo[a]anthracene (BaA), Chrysene (Chr), Benzo[e]pyrene (BeP), Benzo[b]fluoranthene (BbF), Benzo[k]fluoranthene (BkF), Benzo[a]pyrene (BaP), Benzo[ghi]perylene (BghiP), Dibenzo[a,h]anthracene (DBahA), Indeno[1,2,3-cd]pyrene (IP), Coronene (COR)
Hopanes	10	Trisnorneohopane, 17a(H)-Trisnorhopane, 17a(H)-21b(H)-Norhopane, 17a(H)-21b(H)-Hopane, 17a(H)-21b(H)-22S-Homohopane, 17a(H)-21b(H)-22R-Homohopane, 17a(H)-21b(H)-22S-Bishomohopane, 17a(H)-21b(H)-22R-Bishomohopane, 17a(H)-21b(H)-22S-Trishomohopane, 17a(H)-21b(H)-22R-Trishomohopane
n-alkanes	29	C12, C13, C14, C15, C16, C17, C18, C19, C20, C21, C22, C23, C24, C25, C26, C27, C28, C29, C30, C31, C32, C33, C34, C35, C36, C37, C38, C39, C40
POA	3	Arabitol, Sorbitol, Mannitol
Others	3	Pristane, Phytane, Glucose

Table A3

Description of the SNAP macro-sectors used in this study

#	SNAP nomenclature ¹	MACC II nomenclature ²	Description
1	Combustion in energy and transformation industries	Energy industries	Combustion processes for large-scale energy production and its transformation including emissions of boilers, gas turbines and stationary engines.
2	Non-industrial combustion plants	Residential & commercial combustion	Combustion processes aimed at producing heat (heating) for non-industrial activities: commercial and institutional installations, including residential (heating and domestic combustion processes such as fireplaces, stoves, etc.). The macro-sector is subdivided according to the used fuel: 2.1 includes the emissions deriving from combustion of fossil fuels and 2.2 those from the combustion of biomass burning.
3	Combustion in manufacturing industry ³	Industry (combustion)	Combustion processes strictly related to industrial activity. Includes all the processes that require locally produced energy through combustion: boilers, furnaces, first melting of metals, production of gypsum, asphalt, cement, etc.
4	Production processes ³	Industry (processes)	Industrial emissions originated by all processes for the production of a given good or material different than combustion. Includes all the processes in the iron and steel, mechanics, organic and inorganic chemistry, wood, food production industries, among others.
5	Extraction and distribution of fossil fuels and geothermal energy	Extraction and distribution of fossil fuels	Land and off-shore emissions deriving from production, distribution, storage of solid, liquid and gaseous fuel. It also includes emissions from geothermal energy extraction processes.
#	SNAP nomenclature ¹	MACC II nomenclature ²	Description
6	Solvents and other products use	Product use	Emissions from activities that involve the use of products containing solvents, excepting their production (e.g. painting and degreasing operations including the domestic use of such products).
7	Road transport	Road transport	Emissions due to cars, light and heavy vehicles, motorcycles and other means of transport on the road, including both emissions due to exhaust and wear from brakes, wheels and the road. The macro-sector is subdivided in five subcategories: 7.1 exhaust gasoline, 7.2 exhaust diesel, 7.3 exhaust natural gas/LPG, 7.4 non-exhaust evaporation, and 7.5 non-exhaust wear of brakes etc.
8	Other mobile sources and machinery	Non-road transport and other mobile sources	Emissions from rail transportation, inland navigation, military vehicles, maritime traffic, air traffic and mobile non-road internal combustion sources, such as agricultural and forestry vehicles (chainsaws, pruning equipment, etc.), those linked to gardening activities (lawn mowers, etc.) and industrial vehicles (bulldozers, caterpillars, etc.). In this study only the international shipping emissions are considered.
9	Waste treatment and disposal	Waste treatment	Emissions from waste incineration, spreading, and landfill including waste management related activities such as the treatment of waste water, composting, the production of biogas, the spreading of sludge, etc.
10	Agriculture	Agriculture	Emissions deriving from all agricultural practices with the exception of the thermal heating groups (included in the macro-sector 3) and the motor vehicles (included in the macro-sector 8). Includes emissions from crops with and without fertilizers and /or pesticides, pesticides, herbicides, incineration of residues carried out on site, emissions due to breeding activities (enteric fermentation, production of organic compounds) and nursery production.
11	Nature and other sources and sinks		It includes all the non-anthropogenic activities that generate emissions (activity of plants, shrubs and grass, lightning, spontaneous emissions of gas, emissions from the soil, volcanoes, natural combustion, etc.) and those activities managed by man that are connected to them (managed forests, planting, repopulation, arson combustion of forests). In this study are considered only the following subcategories: 11.1 natural dust, 11.2 sea salt and 11.3 biogenic secondary organic aerosol.

¹EEA, 2000.² adapted from Kuenen et al. (2014).³ sectors 3 and 4 are merged.

References

- Amato, F., Alastuey, A., Karanasiou, A., Lucarelli, F., Nava, S., Calzolari, G., Severi, M., Becagli, S., Gianelle, V.L., Colombi, C., Alves, C., Custódio, D., Nunes, T., Cerqueira, M., Pio, C., Eleftheriadis, K., Diapouli, E., Reche, C., Minguillón, M.C., Manousakas, M.-I., Maggos, T., Vratolis, S., Harrison, R.M., Querol, X., 2016. AIRUSE-LIFE+: a harmonized PM speciation and source apportionment in five southern European cities. *Atmos. Chem. Phys.* 16, 3289–3309. <https://doi.org/10.5194/acp-16-3289-2016>.
- Argyropoulos, G., Samara, C., 2011. Development and application of a robotic chemical mass balance model for source apportionment of atmospheric particulate matter. *Environ. Model. Softw* 26, 469–481.
- Baker, K.R., Kelly, J.T., Fox, T., 2013. Estimating Secondary Pollutant Impacts from Single Sources Air and Waste Management Association - Guideline on Air Quality Models 2013: the Path Forward, vol. 2, pp. 810–827.
- Baudic, A., Gros, V., Sauvage, S., Locoge, N., Sanchez, O., Sarda-Estève, R., Kalogridis, C., Petit, J.E., Bonnaire, N., Baisnée, D., Favez, O., Albinet, A., Sciare, J., Bonsang, B., 2016. Seasonal variability and source apportionment of volatile organic compounds (VOCs) in the Paris megacity (France). *Atmos. Chem. Phys.* 16, 11961–11989.

- Belis, C.A., Karagulian, F., Larsen, B.R., Hopke, P.K., 2013. Critical review and meta-analysis of ambient particulate matter source apportionment using receptor models in Europe. *Atmos. Environ.* 69, 94–108.
- Belis, C.A., Pernigotti, D., Karagulian, F., Pirovano, G., Larsen, B.R., Gerboles, M., Hopke, P.K., 2015. A new methodology to assess the performance and uncertainty of source apportionment models in intercomparison exercises. *Atmos. Environ.* 119, 35–44.
- Belis, C.A., Karagulian, F., Amato, F., Almeida, M., Artaxo, P., Beddows, D.C.S., Bernardoni, V., Bove, M.C., Carbone, S., Cesari, D., Contini, D., Cuccia, E., Diapouli, E., Eleftheriadis, K., Favez, O., El Haddad, I., Harrison, R.M., Hellebust, S., Hovorka, J., Jang, E., Jorquera, H., Kammermeier, T., Karl, M., Lucarelli, F., Mooibroek, D., Nava, S., Nøjgaard, J.K., Paatero, P., Pandolfi, M., Perrone, M.G., Petit, J.E., Pietrodangelo, A., Pokorná, P., Prati, P., Prevot, A.S.H., Quass, U., Querol, X., Saraga, D., Sciare, J., Sfetsos, A., Valli, G., Vecchi, R., Vestenius, M., Yubero, E., Hopke, P.K., 2015. A new methodology to assess the performance and uncertainty of source apportionment models II: the results of two European intercomparison exercises. *Atmos. Environ.* 123 (Part A), 240–250.
- Belis, C., Blond, N., Bouland, C., Carnevale, C., Clappier, A., Douros, J., Fragkou, E., Guariso, G., Miranda, A.I., Nahorski, Z., Pisoni, E., Ponche, J.-L., Thunis, P., Viaene, P., Volta, M., 2017. Strengths and weaknesses of the current EU situation. In: Guariso, G., Volta, M. (Eds.), *Air Quality Integrated Assessment, SpringerBriefs in Applied Sciences and Technology*. https://doi.org/10.1007/978-3-319-33349-6_4.
- Bove, M.C., Brotto, P., Cassola, F., Cuccia, E., Massabò, D., Mazzino, A., Piazzalunga, A., Prati, P., 2014. An integrated PM_{2.5} source apportionment study: positive Matrix Factorisation vs. the chemical transport model CAMx. *Atmos. Environ.* 94, 274–286.
- Bove, M.C., Massabò, D., Prati, P., 2018. PMF5.0 vs. CMB8.2: an inter-comparison study based on the new European SPECIEUROPE database. *Atmos. Res.* 201, 181–188. <https://doi.org/10.1016/j.atmosres.2017.10.021>.
- Brandt, J., Silver, J.D., Christensen, J.H., Andersen, M.S., Bønløkke, J.H., Sigsgaard, T., Geels, C., Gross, A., Hansen, A.B., Hansen, K.M., Hedegaard, G.B., Kaas, E., Frohn, L.M., 2013. Contribution from the ten major emission sectors in Europe and Denmark to the health-cost externalities of air pollution using the EVA model system – an integrated modelling approach. *Atmos. Chem. Phys.* 13, 7725–7746.
- Burr, M.J., Zhang, Y., 2011. Source apportionment of fine particulate matter over the Eastern U.S. Part I: source sensitivity simulations using CMAQ with the Brute Force method. *Atmos. Pollut. Res.* 2, 300–317.
- Byun, D.W., 1999. Dynamically consistent formulations in meteorological and air quality models for multiscale Atmospheric studies. Part I: governing equations in a generalized coordinate system. *J. Atmos. Sci.* 56, 3789–3807.
- Byun, D., Schere, K.L., 2006. Review of the governing equations, computational algorithms, and other components of the models-3 community multiscale Air quality (CMAQ) modeling system. *Appl. Mech. Rev.* 59, 51–77.
- Clappier, A., Belis, C.A., Pernigotti, D., Thunis, P., 2017. Source apportionment and sensitivity analysis: two methodologies with two different purposes. *Geosci. Model Dev.* 12.
- Denier Van Der Gon, Hugo, Jan, Hulskotte, Jozwicka, Magdalena, Kranenburg, Richard, Kuenen, Jeroen, Visschedijk, Antoon, 2018. European emission inventories and projections for road transport non-exhaust emissions. In: Amato, F. (Ed.), *Non-Exhaust Emissions: an Urban Air Quality Problem for Public Health*. Academic Press, San Diego, pp. 102–120.
- Yarwood, G., Morris, R.E., Wilson, G.M., 2004. Particulate matter source apportionment Technology (PSAT) in the CAMx photochemical grid model. In: *Proceedings of the 27th NATO/CCMS International Technical Meeting on Air Pollution Modeling and Application*. Springer Verlag.
- 2008/50/EC, D. 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, Edited, European Parliament.
- Dunker, A.M., Koo, B., Yarwood, G., 2015. Source apportionment of the anthropogenic increment to ozone, formaldehyde, and nitrogen dioxide by the path-integral method in a 3D model. *Environ. Sci. Technol.* 49, 6751–6759.
- European Environment Agency - EEA, 1996. EMEP/CORINAIR Atmospheric emission inventory guidebook - first edition. Technical report No 1/1996. <https://www.eea.europa.eu/publications/emep-corinair-atmospheric-emission-inventory/>.
- European Environment Agency - EEA, 2000. EMEP/CORINAIR Atmospheric emission inventory guidebook - second edition 1999. Technical report No 30/2000. <https://www.eea.europa.eu/publications/EMEP/CORINAIR>.
- European Environment Agency - EEA, 2015. Airbase dataset. <https://www.eea.europa.eu/data-and-maps/data/airbase-the-european-air-quality-database-2>.
- European Environment Agency - EEA, 2018. Air Quality in Europe- 2018 Report. European Environment Agency, Copenhagen, ISBN 978-92-9213-989-6. <https://doi.org/10.2800/777411>.
- ENVIRON, 2014. User's guide comprehensive air quality model with extensions. Version 6.1. http://www.camx.com/files/camxusersguide_v6-10.pdf. Fragkou et al., 2012.
- European Commission - EC, 2011. Commission Staff Working Paper "Establishing Guidelines for Demonstration and Subtraction of Exceedances Attributable to Natural Sources under the Directive 2008/50/EC on Ambient Air Quality and Cleaner Air for Europe" SEC(2011) 208 Final.
- Foley, K.M., Roselle, S.J., Appel, K.W., Bhawe, P.V., Pleim, J.E., Otte, T.L., Mathur, R., Sarwar, G., Young, J.O., Gilliam, R.C., Nolte, C.G., Kelly, J.T., Gilliland, A.B., Bash, J.O., 2010. Incremental testing of the Community Multiscale Air Quality (CMAQ) modeling system version 4.7. *Geosci. Model Dev.* 22.
- Fragkou, E., Douros, I., Moussiopoulos, N., Belis, C.A., 2012. Current trends in the use of models for source apportionment of air pollutants in Europe. *Int. J. Environ. Pollut.* 50, 363–375.
- Gariazzo, C., Silibello, C., Finardi, S., Radice, P., Piersanti, A., Calori, G., Cecinato, A., Perrino, C., Nussio, F., Cagnoli, M., Pelliccioni, A., Gobbi, G.P., Di Filippo, P., 2007. A gas/aerosol air pollutants study over the urban area of Rome using a comprehensive chemical transport model. *Atmos. Environ.* 41, 7286–7303.
- Golly, B., Brulfert, G., Berlioux, G., Jaffrezo, J.L., Besombes, J.L., 2015. Large chemical characterisation of PM₁₀ emitted from graphite material production: application in source apportionment. *Sci. Total Environ.* 538, 634–643.
- Hendriks, C., Kranenburg, R., Kuenen, J., van Gijlswijk, R., Wichink Kruit, R., Segers, A., Denier van der Gon, H., Schaap, M., 2013. The origin of ambient particulate matter concentrations in The Netherlands. *Atmos. Environ.* 69, 289–303.
- Hong, S.-Y., Noh, Y., Dudhia, J., 2006. A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Weather Rev.* 134, 2318–2341.
- Hopke, P.K., 2016. Review of receptor modeling methods for source apportionment. *J. Air Waste Manag. Assoc.* 66, 237–259.
- Hsu, Y., Divita, F., Dorn, J., 2014. SPECIATE version 4.4: DATABASE DEVELOPMENT DOCUMENTATION. Abt associates inc - U.S. EPA EPA contract No. EP-D-08-100/FINAL REPORT EPA/600/R-13/307 U.S.A. https://www.epa.gov/sites/production/files/2015-10/documents/speciate_version4.4_finalreport.pdf.
- ISO 13528, 2005. Statistical Methods for Use in Proficiency Testing by Interlaboratory Comparisons, Edited. (ISO) International Organization for Standardization.
- Izadmanesh, Y., Ghasemi, J.B., Tauler, R., 2017. Receptor modeling of environmental aerosol data using MLPCA-MCR-ALS. *Chemometr. Intell. Lab. Syst.* 167, 50–62.
- Jolliffe, J.K., Kindel, J.C., Shulman, I., Penta, B., Friedrichs, M.A.M., Helber, R., Arnone, R.A., 2009. Summary diagrams for coupled hydrodynamic-ecosystem model skill assessment. *J. Mar. Syst.* 76, 64–82.
- Karagulian, F., Belis, C.A., Dora, C.F.C., Prüss-Ustün, A.M., Bonjour, S., Adair-Rohani, H., Amann, M., 2015. Contributions to cities' ambient particulate matter (PM): a systematic review of local source contributions at global level. *Atmos. Environ.* 120, 475–483.
- Kranenburg, R., Segers, A.J., Hendriks, C., Schaap, M., 2013. Source apportionment using LOTOS-EUROS: module description and evaluation. *Geosci. Model Dev. (GMD)* 6, 721–733.
- Kuenen, J.J.P., Visschedijk, A.J.H., Jozwicka, M., Denier van der Gon, H.A.C., 2014. TNO-MACC II emission inventory; a multi-year (2003 - 2009) consistent high-resolution European emission inventory for air quality modelling. *Atmos. Chem. Phys.* 14, 10963–10976.
- Lin, J.C., Wen, D., 2015. A method to quantitatively Apportion pollutants at high spatial and temporal resolution: The Stochastic Lagrangian Apportionment method (SLAM). *Environmental Science and Technology* 49 (1), 351–360. <https://doi.org/10.1021/es505603v>.
- Menut, L., Bessagnet, B., Khvorostyanov, D., Beekmann, M., Blond, N., Colette, A., Coll, I., Curci, G., Foret, G., Hodzic, A., Mailler, S., Meleux, F., Monge, J.-L., Pisoni, I., Siour, G., Turquety, S., Valari, M., Vautard, R., Vivanco, M.G., 2013. Chimere 2013: a model for regional atmospheric composition modelling. *Geosci. Model Dev. (GMD)* 6, 981–1028.
- Napelenok, S.L., Vedantham, R., Bhawe, P.V., Pouliot, G.A., Kwok, R.H.F., 2014. Source-receptor reconciliation of fine-particle emissions from residential wood combustion in the southeastern United States. *Atmos. Environ.* 98, 454–460.
- Norris, G., Vedantham, R., Wade, K., Brown, S., Prouty, J., Foley, C., 2008. Agency, U.S. E.P., Development, O.o.R.a. (Ed.), EPA Positive Matrix Factorization (PMF) 3.0 Fundamentals & User Guide. U.S. Environmental Protection Agency, Washington, DC, 20460.
- Norris, G., Duvall, R., Brown, S., Bai, S., 2014. EPA Positive Matrix Factorization (PMF) 5.0 Fundamentals and User Guide. U.S. Environmental Protection Agency, National Exposure Research Laboratory, Washington, DC 20460.
- Paatero, P., 1997. Least squares formulation of robust non-negative factor analysis. *Chemometr. Intell. Lab. Syst.* 37, 23–35.
- Paatero, P., 1999. The multilinear engine—a table-driven, least squares program for solving multilinear problems, including the n-way parallel factor analysis model. *J. Comput. Graph. Stat.* 8, 854–888.
- Paatero, P., Tapper, U., 1993. Analysis of different modes of factor analysis as least squares fit problems. *Chemometr. Intell. Lab. Syst.* 18, 183–194.
- Pernigotti, D., Belis, C.A., 2018. DeltaSA tool for source apportionment benchmarking, description and sensitivity analysis. *Atmos. Environ.* 180, 138–148.
- Pernigotti, D., Belis, C.A., Spanó, L., 2016. SPECIEUROPE: the European data base for PM source profiles. *Atmos. Pollut. Res.* 7, 307–314.
- Pillon, S., Liguori, F., Dalan, F., Maffei, G., 2011. PM source apportionment analysis in the Venetian area. *Int. J. Environ. Pollut.* 47 (1–4), 134–148. <https://doi.org/10.1504/IJEP.2011.047330>, 47330.
- Stein, U., Alpert, P., 1993. Factor separation in numerical simulations. *J. Atmos. Sci.* 50, 2107–2115.
- Strunk, A., Ebel, A., Elbern, H., Friese, E., Goris, N., Nieradzik, L.P., 2010. Four-Dimensional variational assimilation of atmospheric chemical data – application to regional modelling of air quality. In: Lirkov, I., Margenov, S., Wasniewski, J. (Eds.), *Large-Scale Scientific Computing*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 214–222.
- Tauler, R., Viana, M., Querol, X., Alastuey, A., Flight, R.M., Wentzell, P.D., Hopke, P.K., 2009. Comparison of the results obtained by four receptor modelling methods in aerosol source apportionment studies. *Atmos. Environ.* 43, 3989–3997.
- Thunis, P., Pederzoli, A., Pernigotti, D., 2012. Performance criteria to evaluate air quality modeling applications. *Atmos. Environ.* 59, 476–482.
- Thunis, P., Degraeuwe, B., Pisoni, E., Trombetti, M., Peduzzi, E., Belis, C.A., Wilson, J., Clappier, A., Vignati, E., 2018. PM 2.5 source allocation in European cities: a SHERPA modelling study. *Atmos. Environ.* 187, 93–106.
- Thunis, P., Clappier, A., Tarrason, L., Cuvelier, C., Monteiro, A., Pisoni, E., Wesseling, J., Belis, C.A., Pirovano, G., Janssen, S., Guerreiro, C., Peduzzi, E., 2019. Source Apportionment to Support Air Quality Planning: Strengths and Weaknesses of Existing Approaches Environment International (under review).

- Uranishi, K., Ikemori, F., Nakatsubo, R., Shimadera, H., Kondo, A., Kikutani, Y., Asano, K., Sugata, S., 2017. Identification of biased sectors in emission data using a combination of chemical transport model and receptor model. *Atmospheric Environment* 166, 166–181. <https://doi.org/10.1016/j.atmosenv.2017.06.039>.
- Viana, M., Kuhlbusch, T.A.J., Querol, X., Alastuey, A., Harrison, R.M., Hopke, P.K., Winiwarter, W., Vallius, M., Szidat, S., Prévôt, A.S.H., Hueglin, C., Bloemen, H., Wählin, P., Vecchi, R., Miranda, A.I., Kasper-Giebl, A., Maenhaut, W., Hitznerberger, R., 2008. Source apportionment of particulate matter in Europe: a review of methods and results. *J. Aerosol Sci.* 39, 827–849.
- VMM, 2013. CHEMKAR PM10 – Stedencampagne Chemische karakterisering van fijn stof in Antwerpen, Gent, Brugge en Oostende 2011-2012. Report D/2013/6871/021 Vlaamse Milieumaatschappij.
- Wagstrom, K.M., Pandis, S.N., Yarwood, G., Wilson, G.M., Morris, R.E., 2008. Development and application of a computationally efficient particulate matter apportionment algorithm in a three-dimensional chemical transport model. *Atmos. Environ.* 42, 5650–5659.
- Waked, A., Favez, O., Alleman, L.Y., Piot, C., Petit, J.-E., Delaunay, T., Verlinden, E., Golly, B., Besombes, J.-L., Jaffrezou, J.-L., Leoz-Garziandia, E., 2014. Source apportionment of PM10 in a north-western Europe regional urban background site (Lens, France) using positive matrix factorization and including primary biogenic emissions. *Atmos. Chem. Phys.* 14, 3325–3346.
- Wang, Z.S., Chien, C.-J., Tonnesen, G.S., 2009. Development of a tagged species source apportionment algorithm to characterize three-dimensional transport and transformation of precursors and secondary pollutants. *J. Geophys. Res.* 114.
- Wang, J., Wang, S., Voorhees, A.S., Zhao, B., Jang, C., Jiang, J., Fu, J.S., Ding, D., Zhu, Y., Hao, J., 2015. Assessment of short-term PM2.5-related mortality due to different emission sources in the Yangtze River Delta, China. *Atmos. Environ.* 123, 440–448.
- Watson, J.G., Antony Chen, L.-W., Chow, J.C., Doraiswamy, P., Lowenthal, D.H., 2008. Source apportionment: findings from the U.S. Supersites program. *J. Air Waste Manag. Assoc.* 58, 265–288.