

EURODELTA III exercise: An evaluation of air quality models' capacity to reproduce the carbonaceous aerosol



Mihaela Mircea^{a,*}, Bertrand Bessagnet^{b,3}, Massimo D'Isidoro^a, Guido Pirovano^c, Sebnem Aksoyoglu^d, Giancarlo Ciarelli^{d,1}, Svetlana Tsyro^e, Astrid Manders^f, Johannes Bieser^g, Rainer Stern^h, Marta García Vivancoⁱ, Cornelius Cuvelier^j, Wenche Aas^k, André S.H. Prévôt^d, Armin Aulinger^g, Gino Briganti^a, Giuseppe Calori^l, Andrea Cappelletti^a, Augustin Colette^b, Florian Couvidat^b, Hilde Fagerli^e, Sandro Finardi^l, Richard Kranenburg^f, Laurence Rouil^b, Camillo Silibello^l, Gerald Spindler^m, Laurent Poulain^m, Hartmut Herrmann^m, Jose L. Jimenezⁿ, Douglas A. Dayⁿ, Petri Tiitta^o, Samara Carbone^{p,2}

^a ENEA, Italian National Agency for New Technologies, Energy and Sustainable Economic Development, Via Martiri di Monte Sole 4, 40129, Bologna, Italy

^b INERIS, National Institute for Industrial Environment and Risks, Parc Technologique ALATA, 60550, Verneuil-en-Halatte, France

^c RSE S.p.A., via Rubattino 54, 20134, Milan, Italy

^d PSI, LAC, Paul Scherrer Institute, 5232, Villigen, PSI, Switzerland

^e Climate Modelling and Air Pollution Division, Research and Development Department, Norwegian Meteorological Institute (MET Norway), P.O. Box 43, Blindern, N-0313, Oslo, Norway

^f TNO, Dept. Climate, Air and Sustainability, P.O. Box 80015, 3508, TA, Utrecht, the Netherlands

^g HZG, Helmholtz-Zentrum Geesthacht, Institute for Coastal Research, Max-Planck-Straße 1, 21502, Geesthacht, Germany

^h Freie Universität Berlin, Institut für Meteorologie Troposphärische Umweltforschung, Carl-Heinrich-Becker Weg 6–10, 12165, Berlin, Germany

ⁱ CIEMAT, Atmospheric Pollution Unit, Avda. Complutense, 22, 28040, Madrid, Spain

^j ex-European Commission, Joint Research Centre (JRC), 21020, Ispra, Va, Italy

^k Norwegian Institute for Air Research (NILU), Box 100, 2027, Kjeller, Norway

^l ARIANET Srl, Via Gilino n.9, 20128, Milan, Italy

^m Leibniz-Institut für Troposphärenforschung (TROPOS), Permoserstraße 15, 04318, Leipzig, Germany

ⁿ Dept. of Chemistry and CIRES, University of Colorado, Boulder, CO, USA

^o Department of Environmental and Biological Sciences, University of Eastern Finland, P.O. Box 1627, FIN-70211, Kuopio, Finland

^p Air Quality Research, Finnish Meteorological Institute, Erik Palménin aukio 1, 00560, Helsinki, Finland

HIGHLIGHTS

- Carbonaceous aerosol model inter-comparison exercise using same input data.
- Multi-model evaluation of primary and secondary organic aerosol at European level.
- Seasonal, daily and hourly validation of modelled concentrations against measurements.
- Comparison of modelled biogenic and anthropogenic secondary aerosol concentrations.

ARTICLE INFO

Keywords:

Elemental carbon

Organic aerosol

Secondary organic aerosol

Model validation

ABSTRACT

The carbonaceous aerosol accounts for an important part of total aerosol mass, affects human health and climate through its effects on physical and chemical properties of the aerosol, yet the understanding of its atmospheric sources and sinks is still incomplete. This study shows the state-of-the-art in modelling carbonaceous aerosol over Europe by comparing simulations performed with seven chemical transport models (CTMs) currently in air

* Corresponding author. ENEA - National Agency for New Technologies, Energy and Sustainable Economic Development, via Martiri di Monte Sole 4, 40129, Bologna, Italy.

E-mail address: mihaela.mircea@enea.it (M. Mircea).

¹ Now at Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, USA.

² Now at Agrarian Sciences Institute, Federal University of Uberlandia, Amazonas Av., Uberlandia, Brazil.

³ Now at Hangzhou Futuris Environmental Technology Co. Ltd, Zhejiang Overseas High-Level Talent Innovation Park, No. 998 WenYi Road, 311121, Hangzhou, Zhejiang, China.

<https://doi.org/10.1016/j.aeoa.2019.100018>

Received 29 March 2018; Received in revised form 24 January 2019; Accepted 26 January 2019

Available online 14 February 2019

2590-1621/ © 2019 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Model inter-comparison

quality assessments in Europe: CAMx, CHIMERE, CMAQ, EMEP/MSC-W, LOTOS-EUROS, MINNI and RCGC. The simulations were carried out in the framework of the EURODELTA III modelling exercise and were evaluated against field measurements from intensive campaigns of European Monitoring and Evaluation Programme (EMEP) and the European Integrated Project on Aerosol Cloud Climate and Air Quality Interactions (EUCAARI). Model simulations were performed over the same domain, using as much as possible the same input data and covering four seasons: summer (1–30 June 2006), winter (8 January – 4 February 2007), autumn (17 September–15 October 2008) and spring (25 February - 26 March 2009). The analyses of models' performances in prediction of elemental carbon (EC) for the four seasons and organic aerosol components (OA) for the last two seasons show that all models generally underestimate the measured concentrations. The maximum underestimation of EC is about 60% and up to about 80% for total organic matter (TOM). The underestimation of TOM outside of highly polluted area is a consequence of an underestimation of secondary organic aerosol (SOA), in particular of its main contributor: biogenic secondary aerosol (BSOA). This result is independent on the SOA modelling approach used and season. The concentrations and daily cycles of total primary organic matter (TPOM) are generally better reproduced by the models since they used the same anthropogenic emissions. However, the combination of emissions and model formulation leads to overestimate TPOM concentrations in 2009 for most of the models. All models capture relatively well the SOA daily cycles at rural stations mainly due to the spatial resolution used in the simulations. For the investigated carbonaceous aerosol compounds, the differences between the concentrations simulated by different models are lower than the differences between the concentrations simulated with a model for different seasons.

1. Introduction

The carbonaceous aerosol accounts for an important part of atmospheric aerosol which causes negative human health effects (WHO, 2015) and influences regional and global climate (IPCC, 2013). The understanding of its physical and chemical properties as well as its spatial distribution and temporal evolution is still incomplete mainly due to poor knowledge of the emissions of primary carbonaceous particles and gas precursors of secondary carbonaceous aerosols and due to the large number of species involved in the formation and transformation of the carbonaceous particles.

Carbonaceous aerosols consist of various mixtures of elemental (EC) and organic carbon (OC) produced as a result of the interactions between meteorological conditions and emissions that control the atmospheric chemistry formation and transformation processes. Over Europe, the contribution of carbonaceous aerosols to the total atmospheric aerosol mass is accounting for a fraction between 10% and 50% of the particulate matter with particle diameters smaller than 10 μm (PM10) (EEA, 2013). Cavalli et al. (2016), based on comparable data from an action at European scale, confirmed that the carbonaceous aerosols contained in PM10 range from 15% at a Mediterranean site to 43% at the most polluted continental site included in the study. It also showed that the contribution to particulate matter with a diameter smaller than 2.5 μm (PM2.5) is slightly greater at all sites, ranging from 21 to 56%.

EC, or black carbon (BC), the term most often used in the climate-science community (more details on their definition and relationship can be found in Petzold et al. (2013), has a globally averaged radiative forcing comparable to methane (CH_4), the second most important contributor to global warming after the carbon dioxide (CO_2) (IPCC et al., 2013), and it is clearly associated to cardiopulmonary morbidity and mortality (WHO, 2012). The incomplete combustion of fossil fuels, biomass and biofuels releases directly in atmosphere high amounts of EC leading to high pollution in areas with intense road traffic, residential heating (burning of biomass such as wood and fossil fuel coal) and open biomass burning (forest fires and agricultural waste burning).

OC represents the carbon contribution to the organic material (OM) contained in aerosol particles or to the organic aerosol (OA – here called TOM) fraction. According to the site, the contributions of OM to PM10 and PM2.5 mass concentrations vary: Putaud et al. (2010) estimated them in a range from 15 to 25% while Yttri et al. (2007) found contributions up to 40%, with significant differences between urban and rural sites.

Part of OC aerosol fraction has the same sources as EC being emitted into the atmosphere as primary organic aerosol particles (POA – here

called TPOM) while the rest, the so-called secondary organic aerosols (SOA), are formed as a result of organic-aerosol chemistry of volatile organic compounds (VOCs) emitted by anthropogenic (AVOC) and biogenic (BVOC) sources. In Europe, SOA is the main contributor to OC ($52 \pm 17\%$) irrespective of aerosol size fraction and site location (Belis et al., 2013).

SOA is one of the least understood constituent of carbonaceous aerosols; most of the modelling studies show large underestimation of its concentrations in almost all environments (Prank et al., 2016; Simpson et al., 2007; Hodzic et al., 2010; Aksoyoglu et al., 2014; Tsigaridis et al., 2014; Fountoukis et al., 2014). Part of the uncertainties associated to SOA predictions are related to the modelling approaches currently in use: two-product model (SORGAM) (Odum et al., 1996; Schell et al., 2001) and volatility basis set approach (VBS) (Donahue et al., 2006, 2011). SORGAM, the first computational model of SOA formation based on physical and chemical principles, assumes specific compounds and surrogate species for gas precursor classes and describes their degradation in low-volatility products which subsequently partition between gas and particle phases. The VBS approach aims to describe the formation of the all OA compounds, therefore, aggregates them in bins according to their volatility (Donahue et al., 2006) or their volatility and oxidation state (Donahue et al., 2011). In addition to this, VBS uses evolving yields (not fixed as in two-product approach) and evolving volatility due to chemistry processes such as aging, oligomerization, etc. It also considers more precursor classes, adding intermediate volatility organic compound (IVOCs) to semi-volatile organic compounds (SVOCs) and the partitioning of POA aerosol which is treated as a non-volatile compound in two-product approach. The level of complexity included in the VBS approach such as the number of volatility bins/species (Shrivastava et al., 2011) and the parameters used to describe the pathways of organic aerosol (OA) formation and evolution (Zhao et al., 2016) has an important impact on SOA predictions.

Intermediate-volatility organic compound emissions, currently not included in the emissions inventories, can account up to 30 times of POA emissions and explain, on average, up to 30% of the annual SOA in urban areas (Zhao et al., 2016; Ots et al., 2016). In Europe, in areas dominated by isoprene emitted by vegetation, isoprene chemistry has a strong contribution to SOA concentrations as shown by Bessagnet et al. (2008).

A recent study over Europe carried out by Prank et al. (2016) showed that both OC and EC mass concentrations in PM2.5 are underestimated in the range from 40 to 80% and from 20 to 60%, respectively. These results are derived from the comparison of four chemical transport models: CMAQ, EMEP MSC-W, LOTOS-EUROS and

SILAM. Previous studies with EMEP MSC-W model also showed EC underestimations (Genberg et al., 2013; Aas et al., 2012; Tsyro et al., 2007), the fractional bias of EMEP model was less than 20% for most of the examined sites (Genberg et al., 2013).

This study was carried out within the EURODELTA III (ED-III) model intercomparison exercise (Bessagnet et al., 2016) as an activity of EMEP (The European Monitoring and Evaluation Programme) Task Force on Measurement and Modelling (TFMM), under the UNECE (United Nations Economic Commission for Europe) Convention on Long-range Transboundary Air Pollution (LRTAP), and shows an evaluation of the capabilities of state-of-the-art chemical transport models (CTM) largely used in air quality assessments, to simulate carbonaceous aerosols in Europe in different seasons. The past multi-model intercomparison studies have considered mainly particulate matter, total PM mass or some PM components (Vautard et al., 2007; Thunis et al., 2007; Stern et al., 2008; Colette et al., 2011; Prank et al., 2016), addressing only marginally the carbonaceous aerosols issues mainly due to limited availability of EC and OC measurements. In addition, they were focused on specific environmental conditions (e.g. Stern et al., 2008) or limited areas (Vautard et al., 2007; Thunis et al., 2007) or used different horizontal spatial resolutions (Prank et al., 2016). The current study is focused on evaluation of modelled EC and OA in PM_{2.5} based on the daily measurements for EC and hourly measurements for OA.

The model results were evaluated against EC measurements available from four EMEP intensive campaigns (Bessagnet et al., 2016; Vivanco et al., 2017) and against OA measurements available from two intensive measurement field campaigns carried out in a joint framework of EMEP (Tørseth et al., 2012) and EUCAARI (the European Integrated Project on Aerosol Cloud Climate and Air Quality Interactions) project. This is a first model inter-comparison which performs models' validation against hourly OA data obtained with aerosol mass spectrometer (AMS) over whole Europe (Crippa et al., 2014) allowing, thus, an in-depth investigation of models capacities to reproduce the diurnal variations. The models conducted simulations over the same period of time and over the same domain, using the same horizontal spatial resolution and common anthropogenic emission inventory, boundary and meteorological conditions. Therefore, the differences between simulated carbonaceous aerosol concentrations may be attributed to the differences in models' formulation (transport, gas and aerosol chemistry, aerosol dynamics, deposition), to the land-use databases included in the models and due to the biogenic emission precursors independently estimated by every modelling team. The simulations were performed by several teams using seven chemical transport models (CTMs): CAMx (CAMX), CHIMERE (CHIM), CMAQ (CMAQ), EMEP/ MSC-W (EMEP), LOTOS-EUROS (LOTOS-EUROS), FARM⁴ (MINNI) and RCGC (RCG), largely used in European and national air pollution studies. The models, the simulations and the measurements used for models validation are described in Section 2 while Section 3 shows the results of the inter-comparison. The analyses address the measured and modelled concentrations of EC and of three OA components: total organic matter (TOM) which is mainly the sum of total primary organic matter (TPOM) and of secondary organic aerosol (SOA).

2. Description of models, simulations and measurements

2.1. Models

The models, input data and simulations setup used in ED III exercise are described in detail in Bessagnet et al. (2016). Apart from CMAQ, all the models were run on the same domain with a horizontal spatial resolution of $0.25^\circ \times 0.25^\circ$. CMAQ model used a Lambert conformal conic projection and its results were interpolated to the prescribed ED-III grid.

⁴ MINNI is the Italian integrated modelling system used to support air quality policies at national and regional levels.

The models were run in their default configurations with regard to the description of atmospheric physical and chemical processes and vertical grid layers.

The models' characteristics relevant for carbonaceous aerosol predictions are summarized in Table 1. It can be noted that three CTMs (CHIMERE, MINNI, CAMx) used the same biogenic emission model MEGAN v2.0.4 (The Model of Emissions of Gases and Aerosols from Nature) (Guenther et al., 2006, 2012), but they used different land use databases (GLOBCOVER, Corine Land Cover 2006, USGS). MEGAN estimates mainly emissions of gaseous organic compounds from urban, rural and agricultural ecosystems through simple mechanistic algorithms that account for the major known processes controlling biogenic emissions. The model is run offline in MINNI, CHIMERE and CAMX coupled to land cover data and using the meteorological conditions provided by ECMWF IFS (Integrated Forecast System). EMEP, LOTOS-EUROS and RCGC used parameterizations derived from Simpson et al. (1999) for the temporal variations according temperature and light, with maps of tree species from Koeble and Seufert (2001). LOTOS-EUROS and RCGC used the same land use database: Corine Land Cover 2000 and the same number of classes to describe land cover while EMEP used CCE/SEI for Europe, elsewhere GLC2000 (<http://forobs.jrc.ec.europa.eu/products/glc2000/glc2000.php>). CMAQ used the BEIS 3.14 (Biogenic Emission Inventory System: Pierce et al. (2002) or Vukovich and Pierce, 2002; <https://www.elementascience.org/articles/10.12952/journal.elementa.000111/> Schwede et al., 2005) module developed by the US EPA. BEIS biogenic model was incorporated into the Sparse Matrix Operational Kernel Emissions (SMOKE) system and was adapted for the European domain (Bieser et al., 2011).

Apart from CMAQ and RCGC, all models used the ECMWF IFS meteorology but in different ways. CHIMERE added an urban mixing parameterization while RCGC used 3D-data for wind, temperature, humidity and density to produce a diagnostic meteorological analysis. Precipitation and cloud data, and boundary layer heights were retrieved from the IFS data set. Boundary layer parameters as friction velocity and Monin-Obukhov-length were calculated applying standard boundary layer theory. EMEP, LOTOS-EUROS and MINNI also re-processed some ECMWF IFS data. For example, MINNI recomputed with its own scheme the PBL mixing height and used the estimations delivered by ECMWF IFS only for a sensitivity test (Bessagnet et al., 2016). Overall, the meteorological fields actually used by models were not identical both horizontally and vertically.

The CMAQ model used meteorological conditions simulated by the non-hydrostatic operational weather prediction model COSMO (Consortium for Small scale MOdeling) model in CLimate Mode (COSMO-CLM) version 4.8 clm 11 (Geyer, 2014). SOA formation was simulated with a two products model SORGAM (Schell et al., 2001) by three models RCGC, CMAQ and MINNI and VBS was employed by EMEP, LOTOS-EUROS and CAMx. In CHIMERE, SOA formation was represented according to a single-step oxidation of the relevant anthropogenic and biogenic precursor (Bessagnet et al., 2009).

The aerosol microphysics was described in CHIMERE by 8 bins while all the others models use modal approach with 2 modes (EMEP, LOTOS-EUROS, RCGC, CAMx) and with 3 modes (CMAQ, MINNI). The aerosol dynamics due to nucleation, condensation and coagulation processes was considered only by CHIMERE, CMAQ and MINNI.

2.2. Simulations setup

The anthropogenic emissions were provided by INERIS gridded over the simulated domain. The emissions were generated by merging different databases: TNO-MACC $0.125^\circ \times 0.0625^\circ$ emissions for 2007, EMEP $0.5^\circ \times 0.5^\circ$ emission inventory for 2009 and emission data from the GAINS database (<http://gains.iiasa.ac.at/gains>) (Bessagnet et al., 2016). Emissions were harmonized following the methodology described in Terrenoire et al. (2015). The temporal distributions of

Table 1
Description of models and input data relevant for carbonaceous aerosol predictions.

	CAMx	CHIMERE	CMAQ	EMEP	LOTOS-EUROS	MINNI	RCGC
version	v5.40	CHIMERE2013	v4.7.1	rv4.1.3	v1.8	FARM v3.1.12	v2.1
NATURAL EMISSIONS							
Biogenic VOC	MEGAN v2.1	MEGAN v2.04	BEIS 3.14 (Guenther et al., 2000)	Based upon maps of 115 species from Koeble and Seufert (2001), and hourly temperature and light (Simpson et al., 2012)	Based upon maps of 115 species from Koeble and Seufert (2001), and hourly temperature and light (Beltman et al., 2013)	MEGAN v2.04	Based upon maps of 115 species from Koeble and Seufert (2001), and hourly temperature and light using emissions factors of Simpson et al. (1999)
LANDUSE							
Landuse database	USGS	GLOBCOVER (24 classes)	Corine Land Cover 2006 (22 classes) + 120 vegetation classes (Smiatek, 1998).	CCE/SEI for Europe, elsewhere GLC2000	Corine Land Cover 2000 (13 classes)	Corine Land Cover 2006 (22 classes)	Corine Land Cover 2000 (13 classes)
METEOROLOGY							
Driver	ECMWF IFS	ECMWF IFS + urban mixing	COSMO CLM	ECMWF IFS	ECMWF IFS	ECMWF IFS	ECMWF IFS + observations
PROCESSES							
Secondary Inorganic equilibrium model	ISORROPIA v 1.7 (Nenes et al., 1998)	ISORROPIA (Nenes et al., 1999)	ISORROPIA v1.7 (Nenes et al., 1998)	MARS (Binkowski and Shankar, 1995)	ISORROPIA v.2 (Fountoukis and Nenes, 2007)	ISORROPIA v1.7 (Nenes et al., 1998)	ISORROPIA
SOA model	CAMx-VBS (beta version) (Koo et al., 2014)	After Bessagnet et al. (2009) scheme medium complexity	SORGAM (Schell et al., 2001)	VBS-NPAS (Bergström et al., 2012; Simpson et al., 2012)	VBS (Bergström et al., 2012)	SORGAM (Schell et al., 2001)	SORGAM (Schell et al., 2001)
Aerosol model	Bulk approach (2 modes)	8 bins (40 nm–10 μm)	AEROS (3 modes) Carlton et al. (2010)	Bulk approach (2 modes)	Bulk approach (2 modes)	AEROS (3 modes) (Binkowski, 1999, Binkowski and Roselle, 2003); Coagulation/condensation/nucleation	Bulk approach (2 modes)
Aerosol physics	No dynamics	Coagulation/condensation/nucleation	Coagulation/condensation/nucleation	No dynamics	No dynamics	No dynamics	No dynamics
Wet deposition of particles	IC& BC SC (Seinfeld and Pandis, 1988)	Same as fine aerosols fraction	as appropriate for fine aerosol fraction (Byun and Schere, 2006)	Scavenging ratios (Simpson et al., 2012)	BC SC Scott (1979)	as appropriate for fine aerosol fraction (EMEP, 2003)	Sub-cloud scavenging coefficients
Gas phase chemistry	CB05 (Yarwood et al., 2005)	MELCHIOR2 (Lattuati, 1997)	CB05 (Yarwood et al., 2005)	EmChem09 (Simpson et al., 2012)	TNO-CBM-IV (Sauter et al., 2014)	SAPRC99 (Carter, 2000)	CBM-IV

Note 1: The emitted carbonaceous species provided to modelling teams:

EC₂₅: Elemental carbon: < 2.5 μm; EC_{Co}: Elemental carbon: 2.5 μm < - < 10 μm.

POMS_{2_25}: Primary organic matter from SNAP2: < 2.5 μm; POMS_{2_Co}: Primary organic matter from SNAP2 2.5 μm: < - < 10 μm.

OPOM₂₅: Other primary organic material (non SNAP2 POM): < 2.5 μm; OPOM_{Co}: Other primary organic material (non SNAP2 POM): 2.5 μm < - < 10 μm.

Note 2: The species considered inert and are included in the outputs as TPOM: TPOM₂₅ = POMS_{2_25} + OPOM₂₅; TPOM_{Co} = OPOM₂₅ + OPOM_{Co}.

emissions were computed by each modelling team using the same seasonal, weekly and hourly profiles. The anthropogenic emissions were vertically distributed by CTMS on to their own vertical levels starting from prescribed vertical profiles based on [Bieser et al. \(2011\)](#) for industrial sectors and [Mailler et al. \(2013\)](#) for residential heating.

Models used their own split for NO_x (nitrogen oxides), SO_x (sulphur oxides) and NMVOC (non methane volatile organic compounds) emissions depending on the gas phase chemical mechanism used.

The country emissions were re-gridded over the whole Europe with the same methodology for SNAP 2 emissions from non-industrial

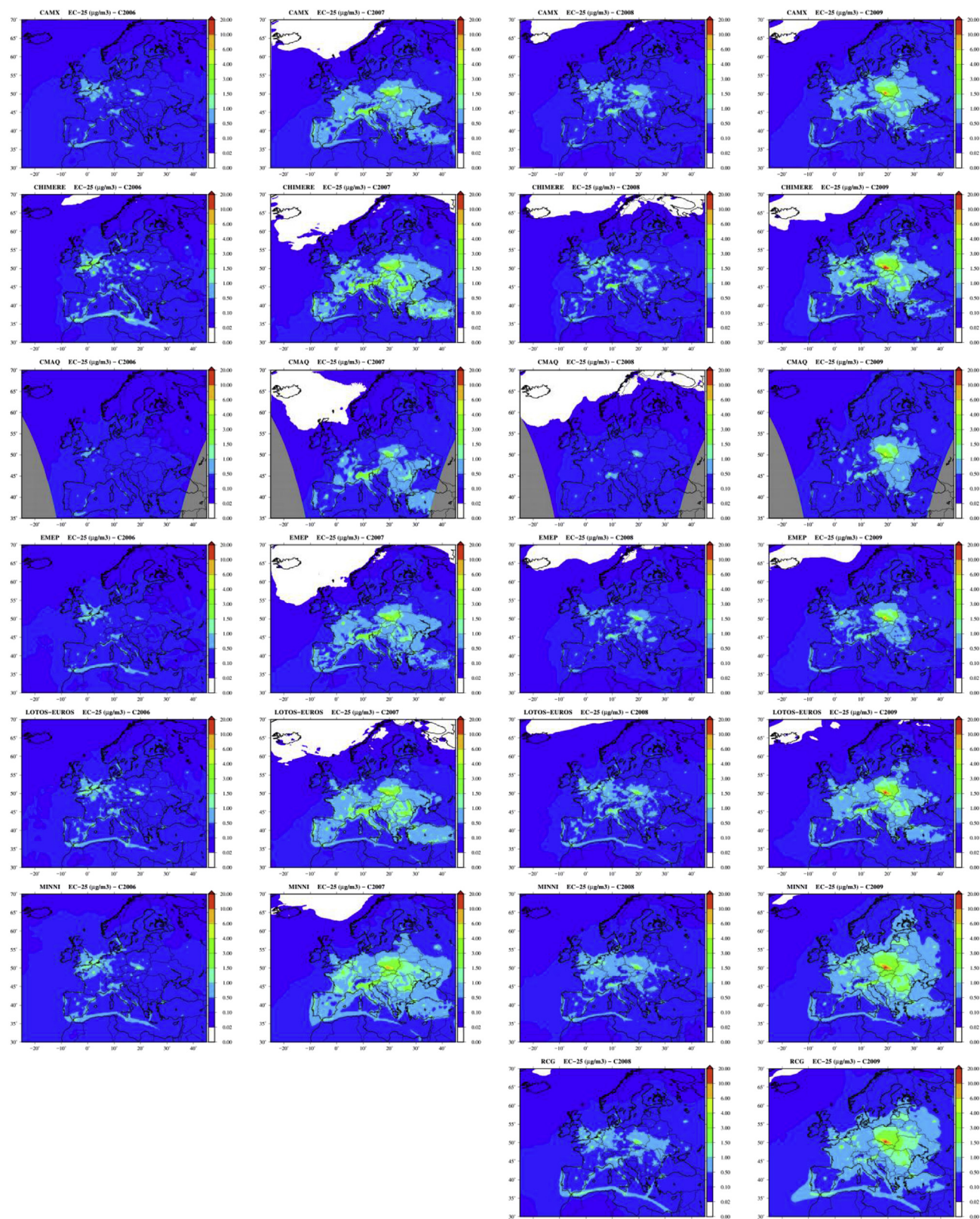


Fig. 1. Averaged EC concentration ($\mu\text{g m}^{-3}$) contained in PM_{2.5} aerosol fraction (EC-25) predicted by the models for 2006, 2007, 2008 and 2009 EMEP campaigns.

combustion, mostly residential emissions from wood burning and from domestic use of coal in two Polish regions. The domestic combustion emissions of PM10 and PM2.5 in two Polish regions that use coal were multiplied by a factor of 4–8 (Bessagnet et al., 2014). In wintertime, residential emissions of particulate matter are dominant in most of European countries, with the highest levels of emissions in Romania, Poland and France and the lowest levels of emissions in Germany, Sweden, Spain. The speciation of PM2.5 and PM coarse was made with a PM speciation profile provided by IIASA (Personal Communication from IIASA) to estimate the fraction of non-carbonaceous species, EC and OC per activity sectors and country. The SVOCs emissions from wood burning recently discussed by Denier van der Gon et al. (2015) were not included. Due to all harmonization efforts described above, no significant difference is expected between the anthropogenic emissions used by the models.

Biogenic VOC (BVOC) emissions from the vegetation were calculated by each team using their own models as described in Table 1.

Wildfire emissions were provided by the GFASv1.0 database (Kaiser et al., 2012) only for the 2006 campaign. The Global Fire Assimilation System (GFASv1.0) calculates biomass burning emissions by assimilating Fire Radiative Power (FRP) observations from the MODIS instruments on-board of the Terra and Aqua satellites. It corrects for gaps in the observations, which are mostly due to cloud cover, and filters spurious FRP observations of volcanoes, gas flares and other industrial activity. For all models the wildfire emissions were assigned in the whole planetary boundary layer.

The simulations used boundary conditions data from the MACC reanalysis (Inness et al., 2013; Benedetti et al., 2009, <http://www.gmes-atmosphere.eu/services/gac/reanalysis/>). The reanalysis production stream provides analyses and 1-day forecasts of global fields of O₃, CO, NO₂, SO₂, HCHO, CO₂, CH₄ and aerosols such as elemental carbon, organic carbon, dust and sulphate. Other reactive gases are available from the coupled chemistry transport model (Bessagnet et al., 2016).

2.3. Measurements

The four EMEP intensive measurement campaigns simulated in

EURODELTA III exercise were carried out for the following periods: 1 June - 30 June 2006, 8 January - 4 February 2007, 17 September - 15 October 2008, 25 February - 26 March 2009. All the data reported from these intensive periods are available from the EMEP data base (<http://ebas.nilu.no>). The EC measurements available from these campaigns, regular data submitted to EMEP, were used in the present study and they may be somewhat different from those used by Aas et al. (2012).

The 2006 (summer) and 2007 (winter) campaigns are described in details in Aas et al. (2012). The 2008 (fall) and 2009 (spring) campaigns were carried out jointly with EUCAARI (European Integrated Project on Aerosol Cloud Climate and Air Quality Interactions) project (Kulmala et al., 2009; Crippa et al., 2014; Pandolfi et al., 2014; Yttri et al., 2018). These campaigns were used in this study because they provided significantly more OA measurements from aerosol mass spectrometer (AMS DeCarlo et al., 2006) in comparison to 2006 and 2007 campaigns (Aas et al., 2012). The data were previously used in aerosol modelling evaluations by Knotte et al. (2011), Mensah et al. (2012), Tsigaridis et al. (2014), Paglione et al. (2014), Kiendler-Scharr et al. (2016); Ciarelli et al. (2016, 2017). The OA measurements were performed at different sites (see Table S1), which are classified as urban (Barcelona - BCN and Helsinki - HEL), rural (Cabauw - CBW, Payerne - PAY, Montseny - MSY, Melpitz - MPZ, Chilbolton - CHL, Harwell - HAR, K - Puszt - KPO, Puijo - PUI and Vavihill - VAV), remote (Mace Head - MHD and Hyttiälä - HYY) and high altitude (Puy de Dome - PDD). The sum of hydrocarbon-like (HOA), biomass burning (BBOA) and cooking (COA) organic aerosol concentrations was compared to simulated TPOM (total primary organic matter) concentrations and the sum of oxygenated components (OOA with both oxygenated semi-volatile - SV-OOA and low-volatility - LV-OOA components) concentrations was compared to simulated SOA. The sum of all these organic components (TPOM and SOA) was evaluated against simulated TOM (total organic matter) concentrations. More details on OA compounds and their naming convention can be found in Murphy et al. (2014).

The lack of EC and OA data in Europe made difficult to address their spatial and temporal variation in more detail at the regional scale for the investigated periods. Also, due to rather poor data coverage, the statistics are partly biased and seasonal variability is only indicative.

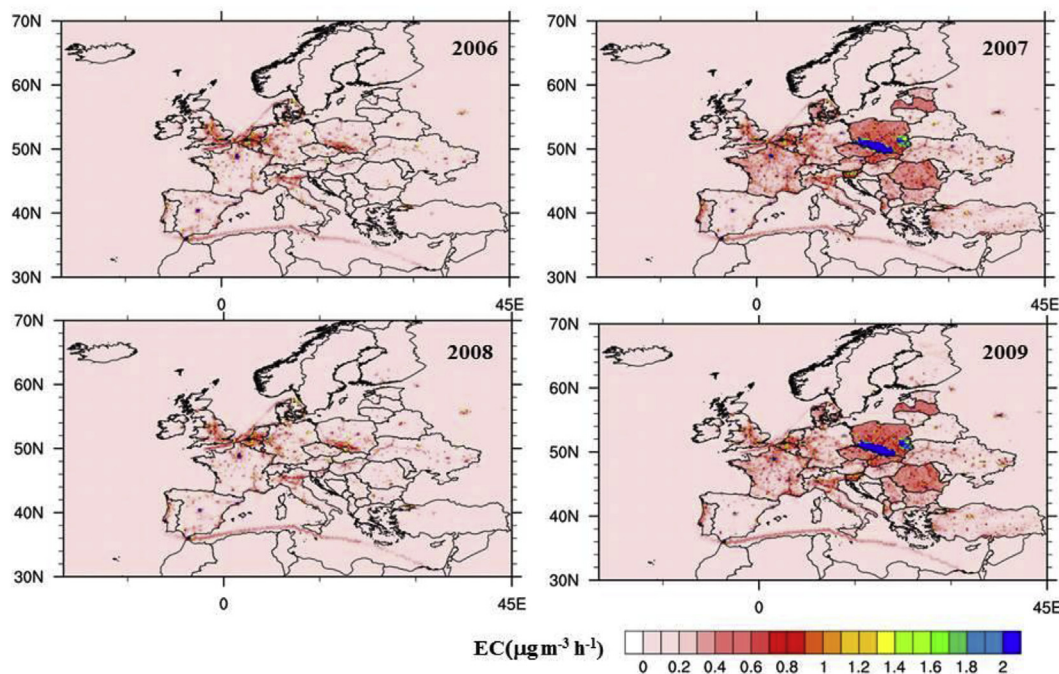


Fig. 2. Averaged EC emission fluxes ($\mu\text{g m}^{-3} \text{h}^{-1}$) for 2006, 2007, 2008 and 2009 EMEP campaigns.

3. Results

Apart from RCG, all models performed simulations of the four campaign periods mentioned in section 2.3. This allowed models' inter-comparison and their evaluation against measurements in different seasons. The analyses were carried out separately for EC and OA since they have different sources and undergo different physical and chemical processes in the atmosphere but also due to lack of simultaneous measurements at stations. Both compounds are considered by CMAQ and MINNI in the fine fraction of aerosol (PM_{2.5}) while CAMX, CHIM, EMEP, LOTOS-EUROS and RCG include its presence also in the coarse fraction contribution thus to PM₁₀ mass. The measurements of EC concentrations contained in the PM_{2.5} matrix (EC-25) are more available and, therefore, they are shown here. The EC contained in PM₁₀ (EC-10), including both the fine and coarse aerosol fractions, is shown in the Supplementary Material. The analysis of EC concentrations includes all seven CTMs while the analysis of OA compounds excludes LOTOS-EUROS as their implementation of VBS was still in the exploratory phase.

Only for EC, the hourly model predictions were averaged to the temporal resolution of the observations at stations. For both EC and OA compounds, several statistical indicators such as mean bias (MB), mean normalised bias (MNB), mean fractional error (MFE) and mean fractional bias (MFB) were used (Table S2). The last two scores do not assume that observations are the absolute truth and give indications about the level of accuracy considered to be close to the best a model can be expected to achieve (Boylan and Russell, 2006). Other two indicators, Rmin (the ratio between the averaged measured concentration and the minimum of averaged simulated concentration) and Rmax (the

ratio between the averaged measured concentration and the maximum of averaged simulated concentration) (Table S2) allow identifying a reproducibility range of the observations in this inter-comparison exercise.

The following discussion of the results is taking into account the differences in actual meteorology used by each modelling system that were extensively analysed and discussed in Bessagnet et al. (2016) and its Supplement.

3.1. Elemental carbon (EC)

The modelled average EC-25 concentrations for summer 2006, winter 2007, autumn 2008 and spring 2009 campaign periods are shown in Fig. 1. For a given campaign, all models predict similar spatial distributions which are closely related to the spatial distribution of emissions shown in Fig. 2. In wintertime, the residential emissions of particulate matter and, thus, of EC from wood-burning or coal, are particularly high in Romania, Poland and France (Bessagnet et al., 2016). In south of Poland, the high concentrations of EC predicted by all models during the cold campaigns (January–February 2007 and February–March 2009) are due the domestic use of coal. High EC concentrations are also predicted in the Po valley (northern Italy) as a result of the synergistic action of the relative high residential biomass burning emissions and stable atmospheric conditions that prevents pollutant transport and dilution. This feature was also observed in a recent study by Glasius et al. (2018) who confirmed that wood combustion is a major source to OC and EC in Northern Europe during winter.

During the warmer periods such as June 2006 and September-

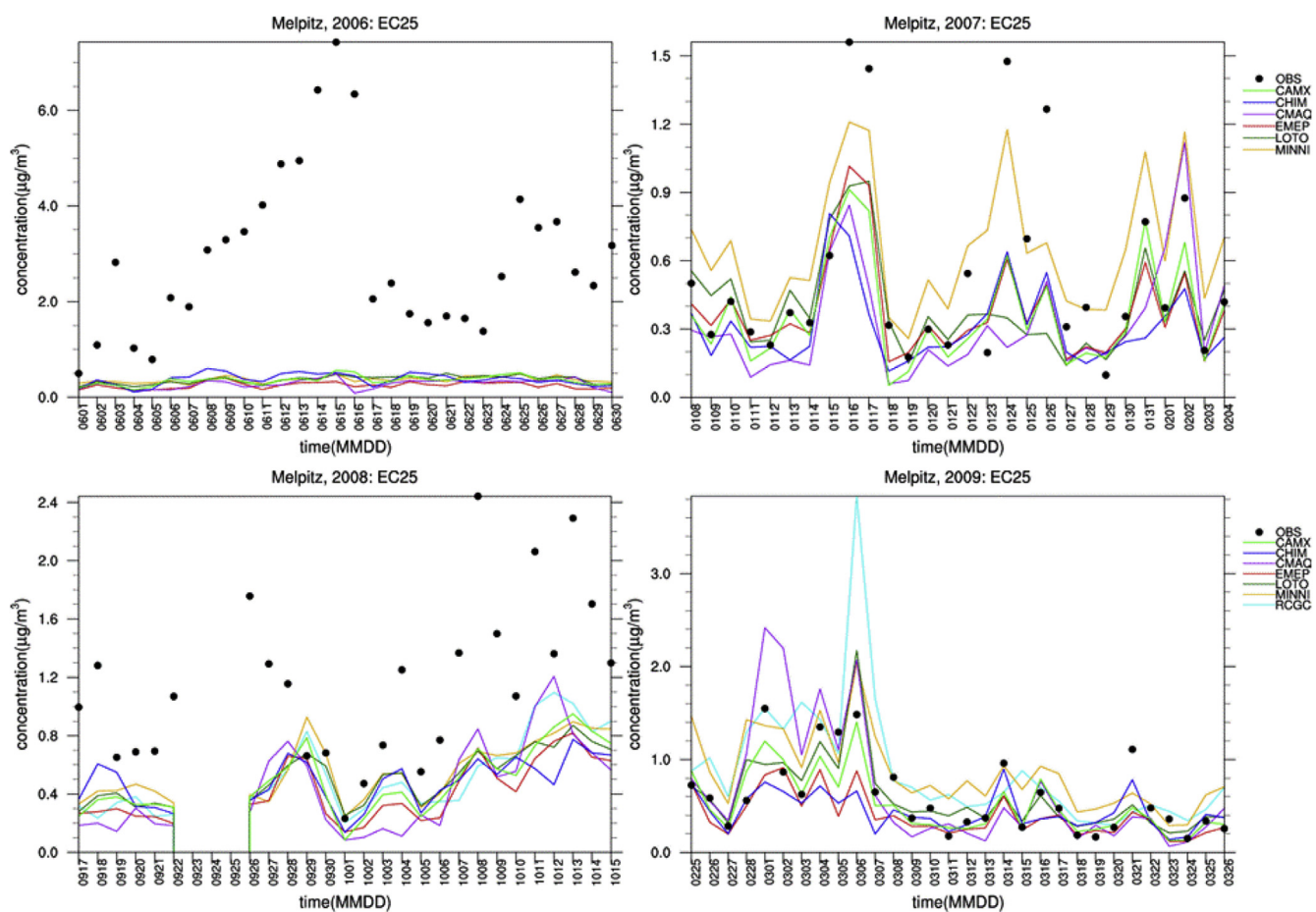


Fig. 3. Daily EC concentrations ($\mu\text{g m}^{-3}$) contained in PM_{2.5} aerosol fraction (EC-25) observed (black dots) and predicted by the models (lines) for 2006, 2007, 2008 and 2009 EMEP campaigns at Melpitz station.

October 2008, the modelled EC concentrations are below $0.5 \mu\text{g m}^{-3}$ over a large area of Europe. Higher EC concentrations are still predicted in the Mediterranean area along the international maritime routes and the English Channel as well as in the metropolitan areas of London, Paris, Belgium, in south of Poland and Po valley. These areas are characterised by high emissions (Fig. 2) related to transport activities. In spite of using the same anthropogenic emission inventory, the models simulate the international maritime routes in different ways showing also a seasonal dependency. For example, the main shipping routes between the eastern and western parts of the Mediterranean basin are clearly visible in Fig. 1 for all models except CMAQ which shows them only in the summer 2006. This behaviour may be mainly due to the models' differences in meteorological parametrizations employed over water (as it was shown in Bessagnet et al. (2016) for CO concentrations variability between models) since they used the same EC anthropogenic emissions, apportioned between fine and coarse aerosol fractions. However, the overall impact of the models' differences in transport, dispersions and dry/wet deposition schemes as well as of the differences in meteorological input data seems limited since, in most areas over the land, the models predict similar EC concentrations near surface. This indicates a high dependence of air concentrations on emission patterns for a primary pollutant, chemically inert such as EC.

The daily measured EC-25 concentrations for all four campaigns were available only at Melpitz (Fig. 3) and Ispra (Fig. S1) stations. At Melpitz, all models capture the daily variations as well as the absolute concentrations reasonably well during the two cold campaigns, 2007 and 2009. However, the modelled EC concentrations are much lower than the measured ones in autumn 2008 and in June 2006, as already reported for the EMEP model in Aas et al. (2012). Two hypotheses may explain the underestimation of concentrations during the warm period (2006 and 2008): the underestimation of anthropogenic emissions around the site and the inability of models to capture the aerosol transport during peak episodes probably due to the adopted horizontal resolution. This station, located in the eastern part of Germany, is influenced by long-range transport of anthropogenic emissions during easterly air mass inflow from countries within the European Union (EU), e.g. Poland, the Czech Republic and Slovakia, etc. According to Spindler et al. (2013), the winter and summer air mass inflow is about 50% of the time from West and about 15% from East. The eastern air masses are more polluted than the western ones since the emission reduction measures started later in those countries, resulting in a still higher level of emissions. In addition to this, the very high levels of EC measured in June 2006 with respect to other seasons may be due to technical problems campaign related. However, the models' predictions are more similar among them during the warm season (2006, 2008) than during the cold season (2007, 2009). The latter shows differences

up to a factor of approximately 3 during peak episodes. On 6 March 2009, RCGC model predicts much higher EC concentrations with respect to the other models, highly overestimating the measurements. This peak event is reproduced by all models but the simulated concentrations have different values. Since the models used the same emissions and most of the models used the same meteorological data which implies similar pollutant transport, this variability may be explained by the differences in the dispersion parametrizations of each modelling system and/or meteorological conditions actually used by models and the thickness of the layer near the surface.

The different model performances in reproducing wind speeds at 10 m and PBL heights (Bessagnet et al., 2016 and supplementary material) do not have the expected impact on CTMs predictions of EC. The wind overestimations due to the use of ECMWF data do not lead the CTMs using it to predict lower concentrations with respect to CMAQ model which uses other meteorological data and has the lowest absolute bias corresponding to the best performance (Figs. 1 and 2). Besides, MINNI which generally has the largest underestimation of PBL heights does not predict the highest EC concentrations with respect to the other CTMs which use EMWF data, apart for January 2007 at Melpitz (Fig. 2). This result may be explained not only by the lower PBL height used by MINNI but also by a more intensive long-range transport of pollutants as the results of the highest overestimation of wind in this period.

At Ispra, models could reproduce the daily variations of EC-25 concentrations and there is a relatively good agreement between model predictions and measurements especially at low concentrations (Fig. S1). As for Melpitz, the differences between the models' predictions are larger during the peak episodes (Fig. S1). However, the agreement between models predictions and measurements is relatively good and it is interesting to note the models' ability to reproduce the low concentrations observed from 22 to 26 January 2007, from 3 to 5 October 2008 and 18–20 March 2009. Contrary to Melpitz, the EC values observed in summer at Ispra are lower than those observed during the other seasons and they are captured relatively well by the models. In Ispra, CMAQ model shows a very distinct pattern with respect to other models, particularly during cold seasons. Such differences can probably be ascribed to differences in the dispersion condition simulated by CMAQ meteorological driver, with respect to ECMWF. Of course such differences are more influencing during cold season and particularly in the Po Valley that is generally characterised by strong atmospheric stability conditions. In June 2006, at Montelibretti station (Fig. 4), the models still underestimated the measurements, though much less than at Melpitz, capturing well their upward tendency. At the same station, during 2007 winter campaign, all models show the opposite behaviour overestimating the measured EC: the highest concentrations being predicted by CHIMERE and LOTOS-EUROS. It is interesting to note that

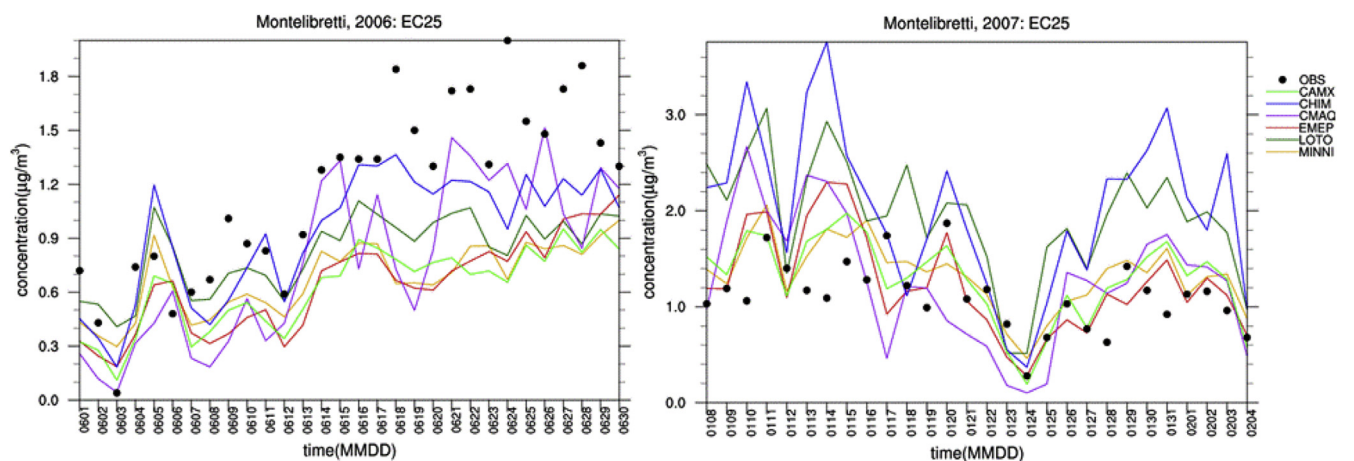


Fig. 4. Daily EC concentrations ($\mu\text{g m}^{-3}$) contained in PM_{2.5} aerosol fraction (EC-25) observed (black dots) and predicted by the models (lines) for 2006 and 2007 EMEP campaigns at Montelibretti station.

all models predict the drop of measured concentration on 12 and 24 January 2007. As for Melpitz and Ispra, the differences between the models' predictions are higher when the modelled EC concentrations are higher. Generally, the models show more and higher peaks than observed that may be explained by the high amount of pollutants transported from Rome metropolitan area characterised by high emissions which could be overestimated. At all sites, the EMEP model, having a relatively thick lowest layer of 92 m, shows the tendency to give the lowest estimation of EC concentrations. Overall, the models' abilities to reproduce the EC concentrations show seasonal and day-to-day variations: the former is mainly controlled by the emission inventory and the latter by both emission inventory and meteorology (transport, dispersion, wet removal, etc.). Another common feature shared by the models is that they simulate successfully the low EC levels but systematically underestimate the peak levels. This may be partly due to the general overestimation of wind speed by all models, except CMAQ (Bessagnet et al., 2016 and supplementary material).

An inclusive view of the models' performances at EMEP stations is shown in composite Taylor diagrams for each campaign (Fig. 5). It can

be noted that the correlations are above 0.6 and standard deviations of models are generally lower than of the observed ones at most of the stations. CMAQ tends to overestimate the variability of observed concentrations, particularly at Ispra site, as already pointed out by Bessagnet et al. (2016) and the models' performances are more dispersed during the cold (2007, 2009) than during the warm (2006, 2008) seasons. This could be related to the model difficulties and diversity (e.g. the use of different minimum K_z diffusion parameter) in the description of the vertical dispersion process in winter when turbulence and boundary layer height are lower with respect to summer, and temperature inversions are frequent and persistent.

Fig. 6 shows the observed and predicted EC concentrations averaged over all stations for the four campaigns and each model. For each campaign, R_{min} and R_{max} indicate the minimum and the maximum of the averaged measured EC reproduced by the present pool of models. Variability of data is indicated with 1σ . In warm seasons (2006, 2008), all models underestimated the measured EC concentrations by about 50–60% while they reproduced those measured in 2009 (R_{max} ratio is close to 1). The winter period (2007) exhibits the largest variability

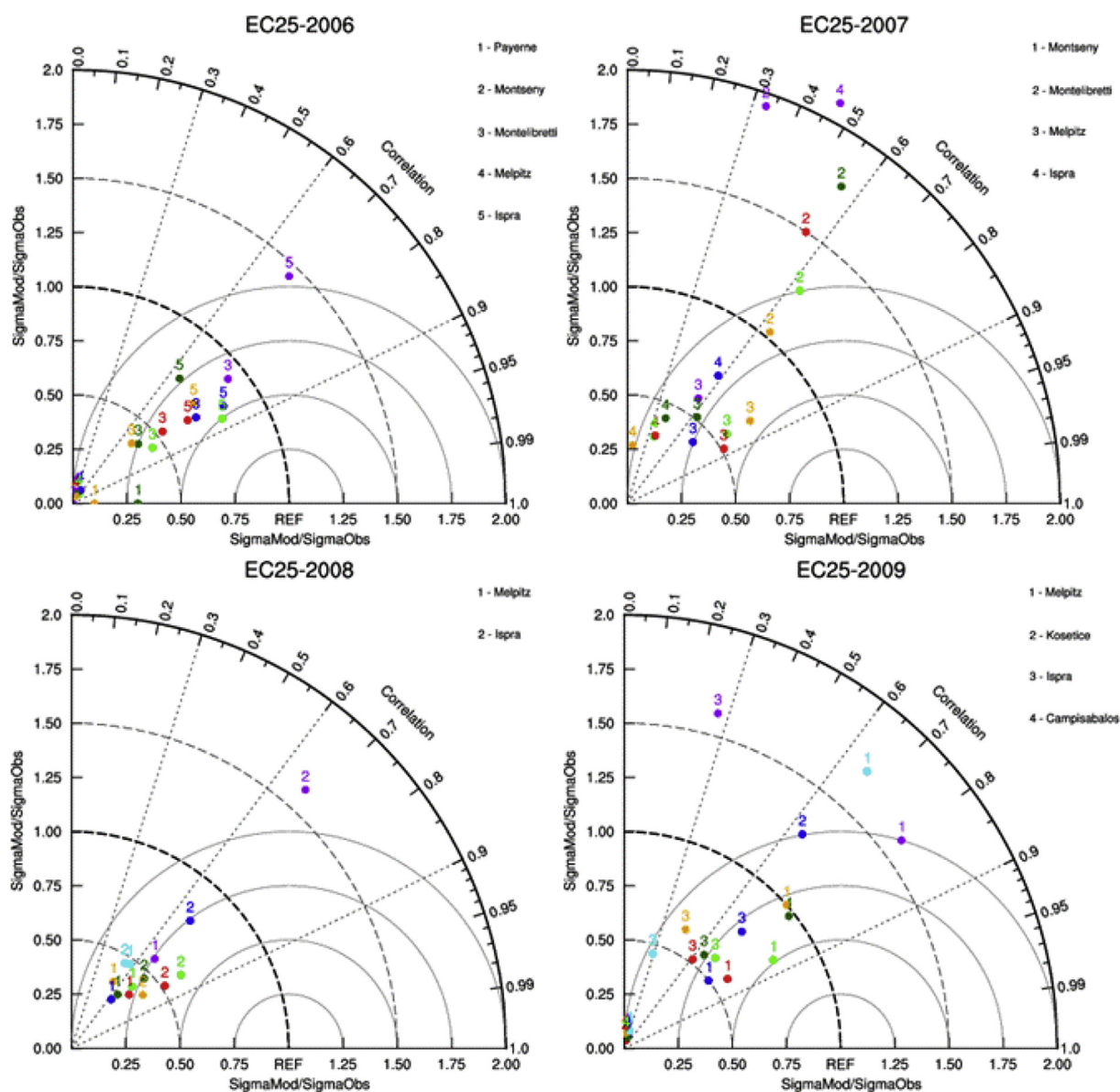


Fig. 5. Taylor diagram for EC concentrations contained in PM_{2.5} aerosol fraction (EC-25) for 2006, 2007, 2008 and 2009 EMEP campaigns at EMEP stations (see legend). CAMX –light green, CHIMERE-blue, CMAQ-violet, EMEP-red, LOTOS-EUROS-green, MINNI-orange, RCGC-light blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

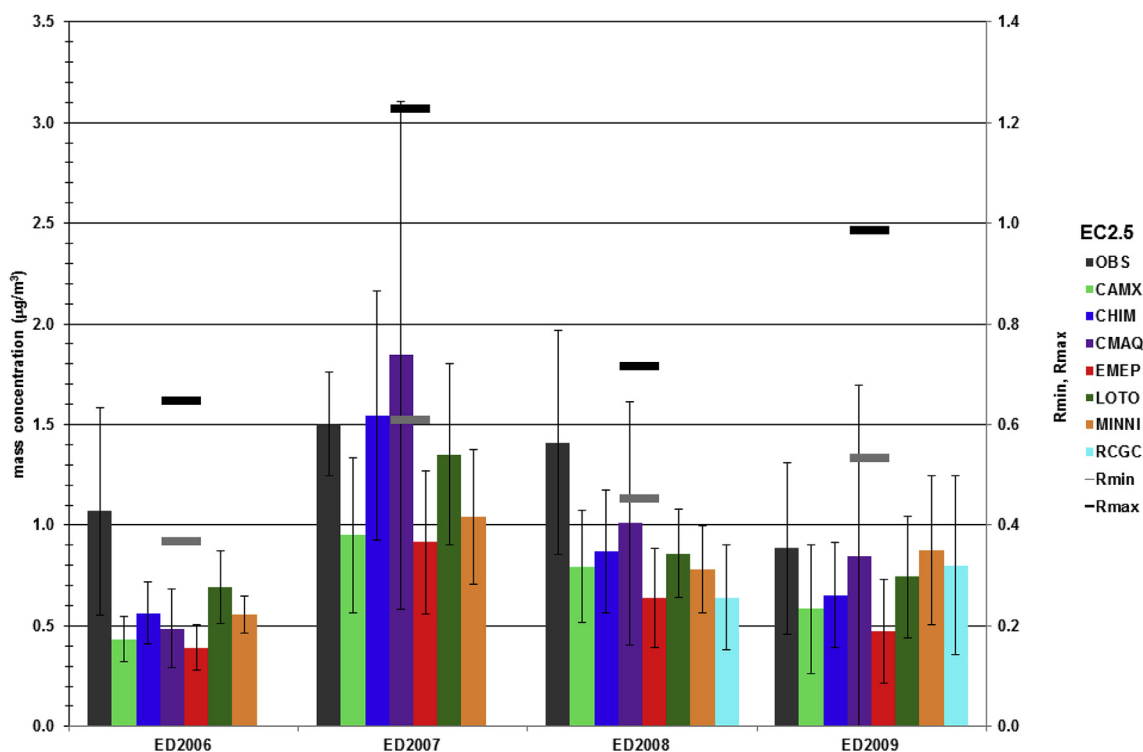


Fig. 6. Averaged observed and predicted EC concentrations ($\mu\text{g m}^{-3}$) contained in PM_{2.5} aerosol fraction (EC-25) for 2006, 2007, 2008 and 2009 EMEP campaigns. The whiskers show the \pm standard deviation. Grey and black symbols show Rmin and Rmax, respectively (right axis).

between model results (also shown by the distance between Rmax and Rmin). All the models underestimate the measurements except CHIMERE which captures the measured levels quite well and CMAQ which shows a substantial overestimation. CMAQ also has the highest standard deviation.

All models except CMAQ and MINNI include EC in both fine and coarse aerosol fractions. As expected, the average EC-25 concentrations (Fig. 1) are slightly lower than EC-10 (Fig. S2) but their spatial distributions are similar. The EC-10 model performances (Fig. S3) are also similar to those obtained for EC-25 (Fig. 6). Due to lack of measurements of both size fractions at the same stations, it is impossible to conduct a direct comparison of EC-25 and EC-10 data.

The evaluation of models against observations for such a limited number of stations and relatively short periods might not be regarded as satisfactory. Nevertheless, given the fact that the model runs were performed with harmonized setup and input information, it provides unique information on the current state of EC modelling.

3.2. Organic aerosol (OA)

The modelled distributions of OA compounds, TOM (TOM-25), TPOM (TPOM-25), SOA (SOA-25) averaged concentrations contained in PM_{2.5} aerosol fraction for summer 2006, winter 2007, autumn 2008 and spring 2009 campaign periods are shown in Figs. 7–9, respectively.

From TOM patterns may be noted the low concentrations simulated by CMAQ over large areas of Europe, followed by EMEP and CHIMERE in 2007, 2008 and 2009 periods (Fig. 7). All models predict the highest TOM concentrations during the cold periods (2007, 2009), showing high concentrations in South of Poland, South of Romania and Po valley. CAMX, CHIMERE and MINNI, in particular, show extended areas with concentrations above $6 \mu\text{g m}^{-3}$ corresponding to intense anthropogenic emissions (see previous section for EC and Fig. 2). Moreover, during the cold periods, SOA contribution (Fig. 9) to TOM is below $1 \mu\text{g m}^{-3}$ for all models, except CAMx. TPOM is the main contributor to TOM in these areas (Fig. 8) as evidenced by the overlap of its

highest values, comparable, concentrations to those of TOM. Over the whole Europe, the anthropogenic SOA (ASOA) (Fig. S4) contribution to TOM concentrations (Fig. 7) is lower than $0.5 \mu\text{g m}^{-3}$ for all models except CAMx, which predicts concentrations above $1.0 \mu\text{g m}^{-3}$. In spite of the diversity of SOA modelling approaches used, all models predict the highest ASOA concentrations in the Po valley making evident the key role of the emissions of ASOA precursors. The EMEP model simulates very high concentrations not only in the Po valley but also over the Adriatic Sea and the Mediterranean basin in September-October 2008 and, particularly, in June 2006. This may indicate enhanced production of secondary aerosol from the ASOA precursors related to maritime transport. Other mechanisms responsible for the transport of pollutants from land to the sea are excluded since similar patterns are not evident for the EC concentrations mainly driven by these physical processes. In case of CAMx, high ASOA concentrations over the sea may be determined by the combination of ASOA production in coastal areas with a less efficiently removal over the sea with respect to the land. The high concentrations of ASOA and consequently of SOA, predicted by CAMx and EMEP models may be explained by the fact that their VBS approaches consider IVOC species of anthropogenic emissions. CMAQ has been shown to underestimate anthropogenic SOA by about an order-of-magnitude in US studies (Hayes et al., 2015; Baker et al., 2015), which is consistent with this evaluation. This underestimation has been corrected in very recent versions (Woody et al., 2016; Murphy et al., 2017), which were not used in this paper. The contribution of biogenic SOA (BSOA) (Fig. S5) to SOA (Fig. 9) and, consequently to TOM (Fig. 7), is higher in June 2006 and September-October 2008. The highest BSOA concentrations over extended areas are predicted by CHIMERE and EMEP in June 2006, followed by MINNI but only over the Scandinavian Peninsula. CMAQ, MINNI and RCGC simulated the highest BSOA concentrations over the Scandinavian Peninsula and this may be due to the combination of the vegetation land cover data and biogenic emission model. In fact, in this area, RCGC model predicts the highest concentrations of α -pinene, a BSOA precursor (Fig. S6).

TPOM (Fig. 8), a primary OA compound by definition, has high

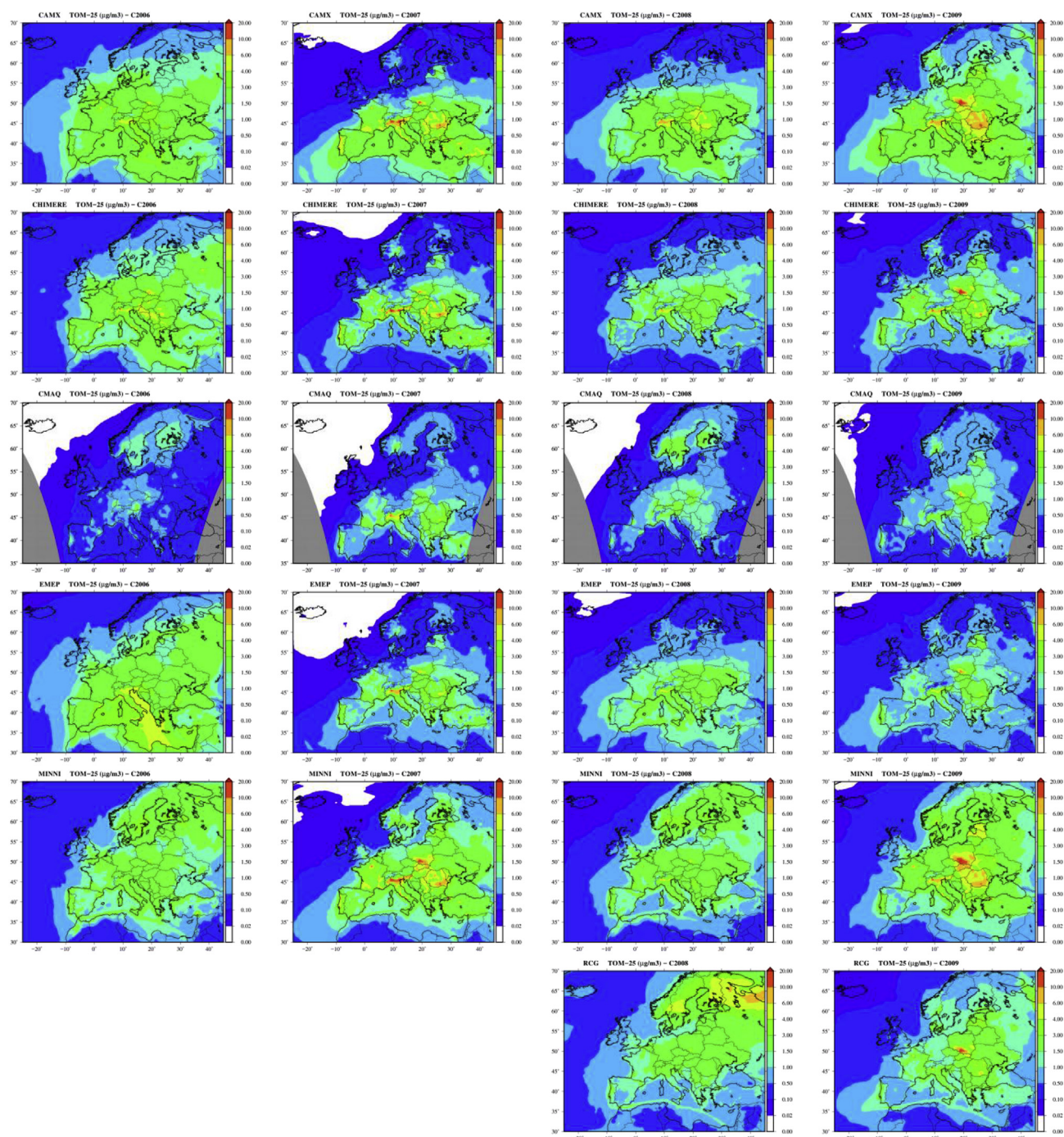


Fig. 7. Averaged TOM concentrations ($\mu\text{g m}^{-3}$) contained in PM_{2.5} aerosol fraction (TOM-25) predicted by the models for 2006, 2007, 2008 and 2009 EMEP campaigns.

concentrations over extended areas in cold seasons: 2007 and 2009, and similar spatial distribution with EC (Fig. 1) since they undergo similar transformations in the atmosphere (transport, dispersion, removal), are of anthropogenic origin and are spatially distributed with the same proxy. It is interesting to note the resemblance between MINNI and CAMx patterns and between CHIMERE, CMAQ and EMEP patterns. Since, apart CMAQ, all models used as input ECMWF-IFS meteorology and the same anthropogenic emissions, it seems straightforward that the models' formulations are responsible for the differences between the two groups of models. But given that a similar behaviour is not observed in EC patterns (Fig. 1), the control of spatial distributions of primary carbonaceous aerosol, TPOM and EC, by both spatial

distribution of emission levels and model formulation can be inferred.

SOA, the secondary OA (sum of ASOA and BSOA), has the opposite behaviour with respect to TPOM, having high concentrations over extended areas during the warm seasons: June 2006 and September-October 2008 (Fig. 9). Its main contributor in these periods is BSOA (Fig. S5) which concentrations are controlled by BVOC precursors emitted by vegetation. As for TOM, CMAQ simulates the lowest SOA concentration, over extended areas of Europe in all periods.

The spatial distribution of the anthropogenic secondary aerosol (ASOA) formed, according to the SOA modelling approach used, by the photochemical oxidation of traditional anthropogenic VOC precursors such aromatics, alkanes, alkenes, etc. and/or the photochemical

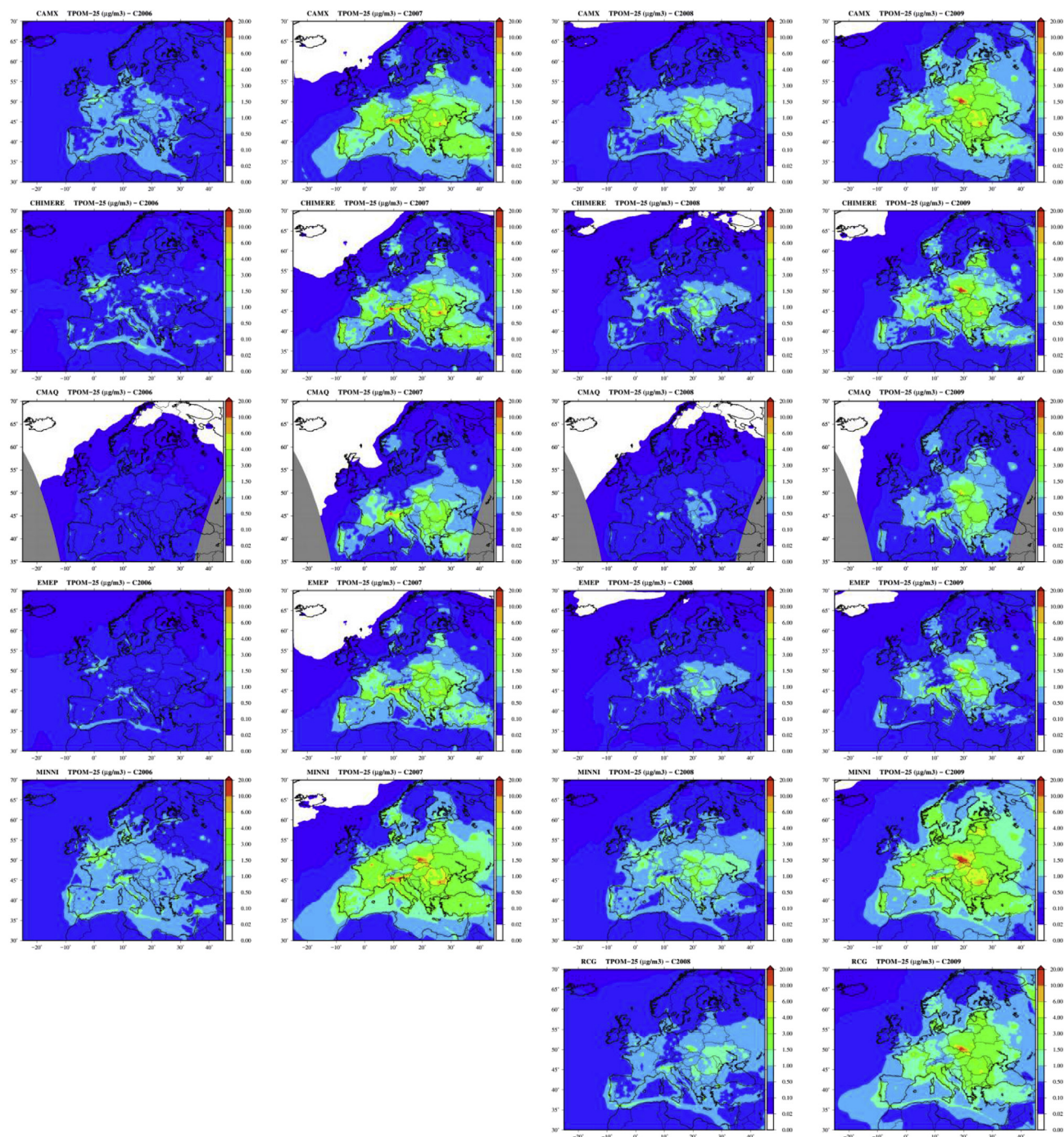


Fig. 8. Averaged TPOM concentrations ($\mu\text{g m}^{-3}$) contained in PM_{2.5} aerosol fraction (TPOM-25) predicted by the models for 2006, 2007, 2008 and 2009 EMEP campaigns.

oxidation of anthropogenic S/IVOC precursors, are shown in Fig. S4. As already mentioned, CAMX has the highest ASOA concentrations during all campaigns and CMAQ predict the lowest, close to zero over extended areas which is consistent with the large underestimations observed over the US and mentioned above. Since the AVOC anthropogenic emissions used by models as input were the same, the differences between the model simulations may be due to SOA modelling approach used (VBS, SORGAM, etc) as well as due to the way of including emissions (speciation, volatility, etc). The further discussion addresses only the former factor since the lack of information on the latter in the framework of the present inter-comparison exercise. The most striking features noticed in ASOA pattern (Fig. S4) are:

- i) EMEP uses a VBS approach as CAMx but it does not show high concentrations over extended areas except for 2006 and, to a lower extent, for 2008 campaigns;
- ii) CMAQ simulates the lowest concentrations over the whole domain, much lower than those predicted by MINNI and RCG in spite of the fact that all three models use SORGAM approach;
- iii) the periods have negligible impact on ASOA patterns, (in particular for CHIMERE and MINNI models).

Therefore, for given AVOC emissions, from i) and ii) may be inferred that the same SOA modelling approach (VBS or SORGAM) can give opposite results and the use of the VBS approach is not a guaranty for

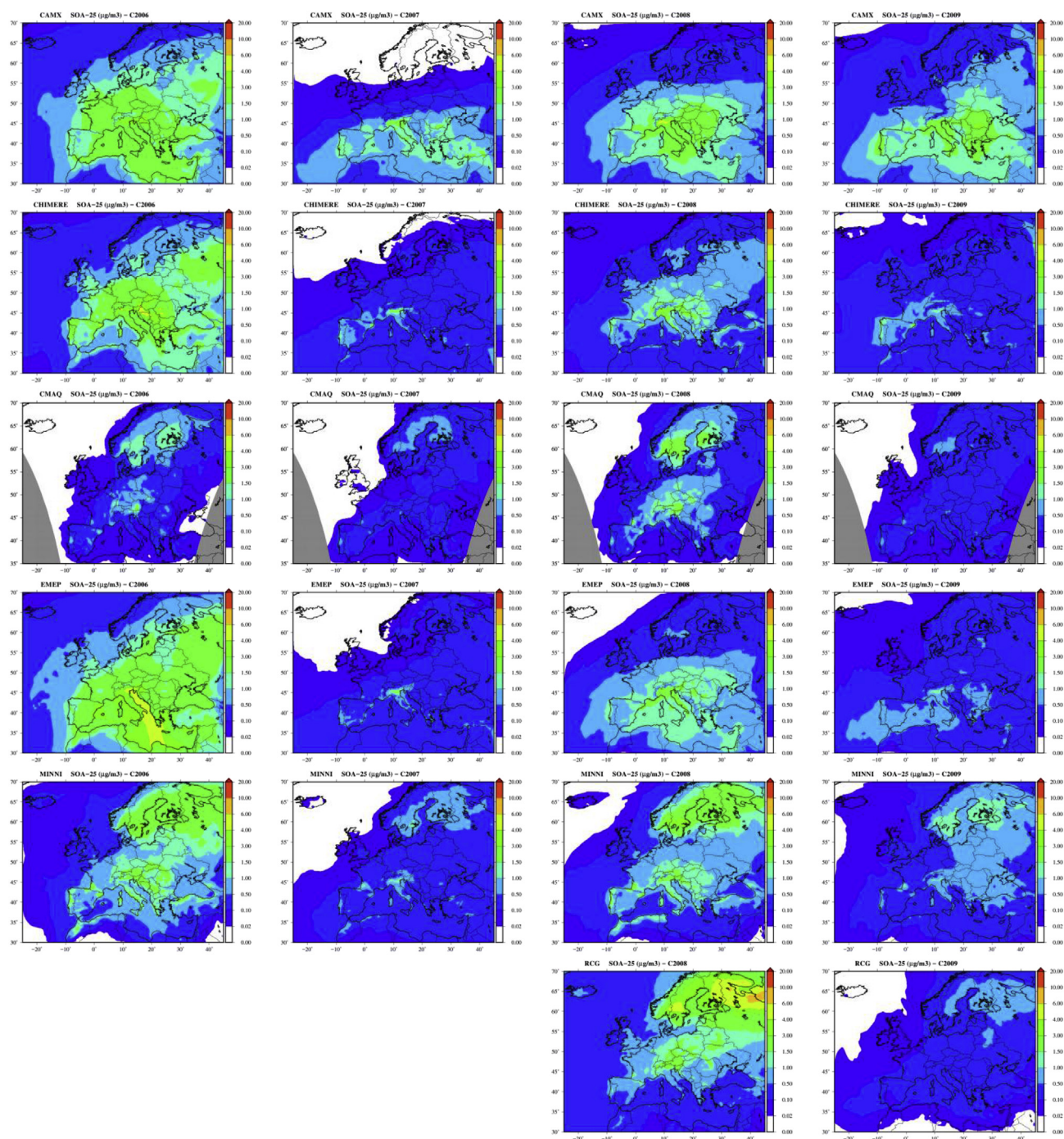


Fig. 9. Averaged SOA concentrations ($\mu\text{g m}^{-3}$) contained in PM_{2.5} aerosol fraction (SOA-25) predicted by the models for 2006, 2007, 2008 and 2009 EMEP campaigns.

the production of higher ASOA concentrations than SORGAM approach and from iii) that the ASOA patterns vary more between the models than with season/period. It can be also noticed that SOA mechanisms of CAMx and EMEP are more sensitive to temperature with respect to the other models, predicting higher ASOA concentrations during 2006 and 2008 periods. The role of temperature on ASOA in models implementing VBS could be related also to the previous point, meaning that during warmer season TPOM can evaporate, then oxidise and finally contribute to ASOA.

CHIMERE, MINNI and to a large extent, EMEP, show similar results for spatial distribution of ASOA (Fig. S4) in spite of using different chemical mechanism as a consequence of reproducing in a similar way

the 2 m temperature and 10 m wind speed: negative bias for temperature and positive bias for wind (Bessagnet et al., 2016 and supplementary material). However, the differences observed for CAMx and CMAQ with respect to the predictions of the other models are more likely due to their ASOA mechanisms than due to their performances in reproducing the meteorological conditions. This hypothesis is supported by the fact that the impact of average differences in temperature (ca. 8 °C in supplementary material of Bessagnet et al. (2016)) between 2008 autumn and 2009 spring is negligible when comparing the ASOA concentrations while the absolute temperature bias for CAMx and CMAQ are below 1° and 1.5 °C, respectively. The wind speed bias and the PBL bias of the two models are quite similar for these periods.

As for ASOA, BSOA concentrations are also controlled by SOA modelling approach and biogenic emissions. The latter was different among modelling systems being produced by BVOC models in relation to the land cover database used and the meteorological conditions. Therefore, the spatial distributions of modelled BSOA concentrations show more spatial variability (Fig. S5). Fig. S6 shows the α -pinene (APINEN) concentrations simulated by EMEP, CAMx and RCG. It can be noted that RCG predicts high α -pinene concentrations in June 2006, much higher than observed at Birkenes, a rural station, located in the south of Norway (58.38 lat, 8.25 lon, 219.0 m), where median concentrations of monoterpenes from May to August in 2012 were around $1 \mu\text{g m}^{-3}$ (Langebner et al., 2014). Probably for this reason, a clear relationship between α -pinene (APINEN) (Fig. S6), a precursor monoterpene compound, and BSOA (Fig. S5) can be seen only for RCG: both species have high concentrations in Scandinavian Peninsula. EMEP and CHIMERE do not show a similar behaviour; therefore, it can be concluded that the BVOC emissions and SOA modelling approaches for other precursors than APINEN control the BSOA concentrations. A more in-depth analysis of impact SOA precursors requires more data that are not available in the present EURODELTA exercise. A recent study (Aksoyoglu et al., 2017) conducted with CAMx for June 2006 has shown that terpenes, both mono- and sesqui-terpens are the main BSOA precursors using VBS approach.

It can be also noted that the lowest BSOA concentrations are simulated by CAMx which simulates the highest ASOA concentrations (Fig. S4). This points out too low BVOC emissions or/and unfavourable setup of VBS approach for BSOA formation and aerosol microphysics treatment. The BSOA differences between CMAQ and MINNI in all seasons except January 2007 (winter season is characterised by the lowest BVOC emissions) may be due to the fact that BVOC emissions from BEIS (CMAQ) are lower than those from MEGAN (MINNI) as shown for isoprene in Pouliot and Pierce (<https://www.epa.gov/air-emissions-modeling/beis-version-history>). On the other hand, the similarity of MINNI and RCGC results are probably due to the insignificant differences between the CORINE 2006 and 2000 and similar parametrizations for BVOC species in MEGAN and KB approach for the species used as SOA precursors in SORGAM. Moreover, the SOA fields of MINNI differ from those simulated by the other two models using MEGAN (CAMx and CHIMERE), but they are similar to those modelled by RCGC based on KB parametrisation for BVOC emissions. Thus, comparing the SOA fields produced by CMAQ, MINNI and RCG that use SORGAM model, CORINE land use (2006, 2006 and 2000 respectively) and different BVOC emissions (BEIS, MEGAN and KB respectively) results that SOA levels are highly dependent on the combination of BVOC model and land-use data as it involves various assumptions to couple the available land cover classes to BVOC emission model requirements. However, the spatial distributions of SOA and BSOA concentrations simulated by RCG, CMAQ and MINNI are similar having maxima over Northern Europe. The models equipped with VBS (CAMx and EMEP) also exhibit similar distributions of SOA and BSOA concentrations even

if they use different BVOC models (MEGAN and KB) and land cover (USGS, CCE/SEI for Europe, elsewhere GLC2000). They predict extended areas with high SOA concentrations located over Europe but not over Scandinavian Peninsula as CMAQ, MINNI, RCG.

Both VBS and SORGAM approaches as well as CHIMERE produce the highest SOA and BSOA concentrations during warm periods (June 2006, September–October 2008). In addition to that, the VBS based models (CAMx, EMEP) and CHIMERE seem more sensitive to meteorology, temperature and radiation, simulating higher concentrations of BSOA over large areas in summer 2006 with respect to autumn 2008. This behaviour is less evident for SORGAM based models, CMAQ and MINNI which shows higher concentrations in Scandinavian Peninsula more related to BVOC precursor emissions. The dependence of BSOA on meteorological conditions cannot be inferred from the present experiment due to its complexity: strong dependence on land cover type due to BSOA dependence on BVOC precursor's emissions (indirect effect), and on temperature during formation process (direct effect).

Fig. 10 shows, on average, how the models reproduce the concentrations of TOM, TPOM and SOA retrieved from AMS measurements during 2008 and 2009 EUCAARI/EMEP campaigns at the stations listed in Table S1 (Crippa et al., 2014). It can be observed that the models overestimate TPOM, particularly in 2009, and, in both periods the underestimation of TOM is due to the underestimation of SOA concentrations. The models reproduce TOM measurements up to 68% (RCG) in 2008 and to 90% (MINNI) in 2009 (Rmin, Rmax axis) while SOA measurements up to 48% (RCG) in 2008 and 30% (CAMx) in 2009. The models' biases vary from $-2.47 \mu\text{g m}^{-3}$ (CMAQ in 2008) to $-0.29 \mu\text{g m}^{-3}$ (MINNI in 2009) for TOM and from $-2.64 \mu\text{g m}^{-3}$ (CMAQ in 2008) to $-1.39 \mu\text{g m}^{-3}$ (CAMx in 2009) for SOA. In 2009, the overestimation of TPOM offsets the underestimation of SOA leading to lower biases for TOM. EMEP and CMAQ reproduce relatively well TPOM for this period while CHIMERE, CAMx and RCG overestimate it within a factor 2. MINNI model highly overestimates TPOM in both campaigns and consequently reproduce relatively well TOM. Since the models used the same TPOM emissions, spatial distributed, the differences in models' ability to reproduce TPOM concentrations may be due to model physics, in particular due to different representation of the removal processes, as well as to the influence of the different vertical dilution strength simulated by the different models (e.g. PBL height in MINNI). However, the large areas with high TPOM concentrations predicted by MINNI, CHIMERE, CAMx and RCG (Fig. 8) and the overestimation of TPOM at stations also suggest that TPOM emissions may be too high.

The models capabilities to reproduce TOM, TPOM and SOA measurements at individual stations are shown in Fig. 11. All the statistical indicators (Table S2) calculated for OA compounds are based on hourly measurements performed with aerosol mass spectrometer (AMS) and described in detail in Crippa et al. (2014). Most of the correlations (CORR) are between 0.3 and 0.9 at stations such as Chilbolton (CHL), Harwell (HAR), Hyytiäläe (HYY), K-Pustza (KPO), Mace Head (MHD),

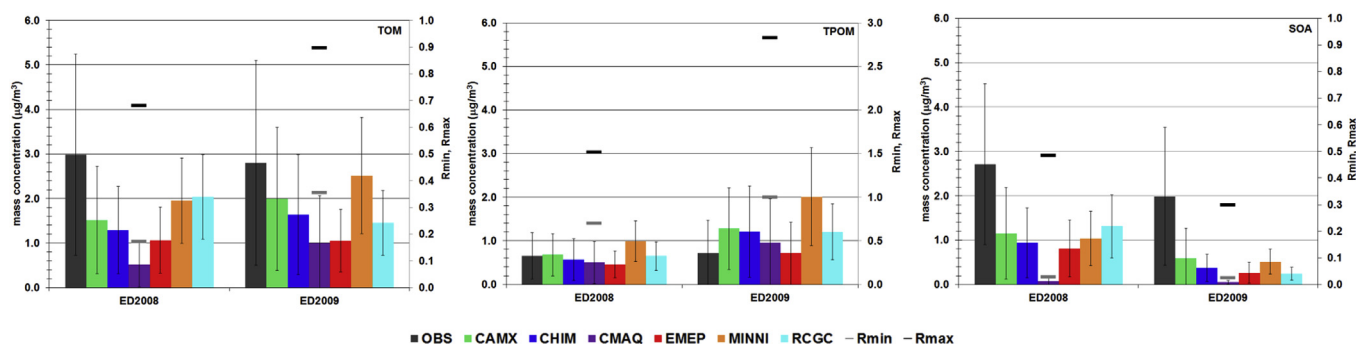


Fig. 10. Averaged observed and predicted TOM, TPOM and SOA concentrations ($\mu\text{g m}^{-3}$) for 2008 and 2009 EMEP/EUCAARI campaigns. The whiskers show \pm standard deviation. Grey and black symbols show Rmin and Rmax, respectively (right axis).

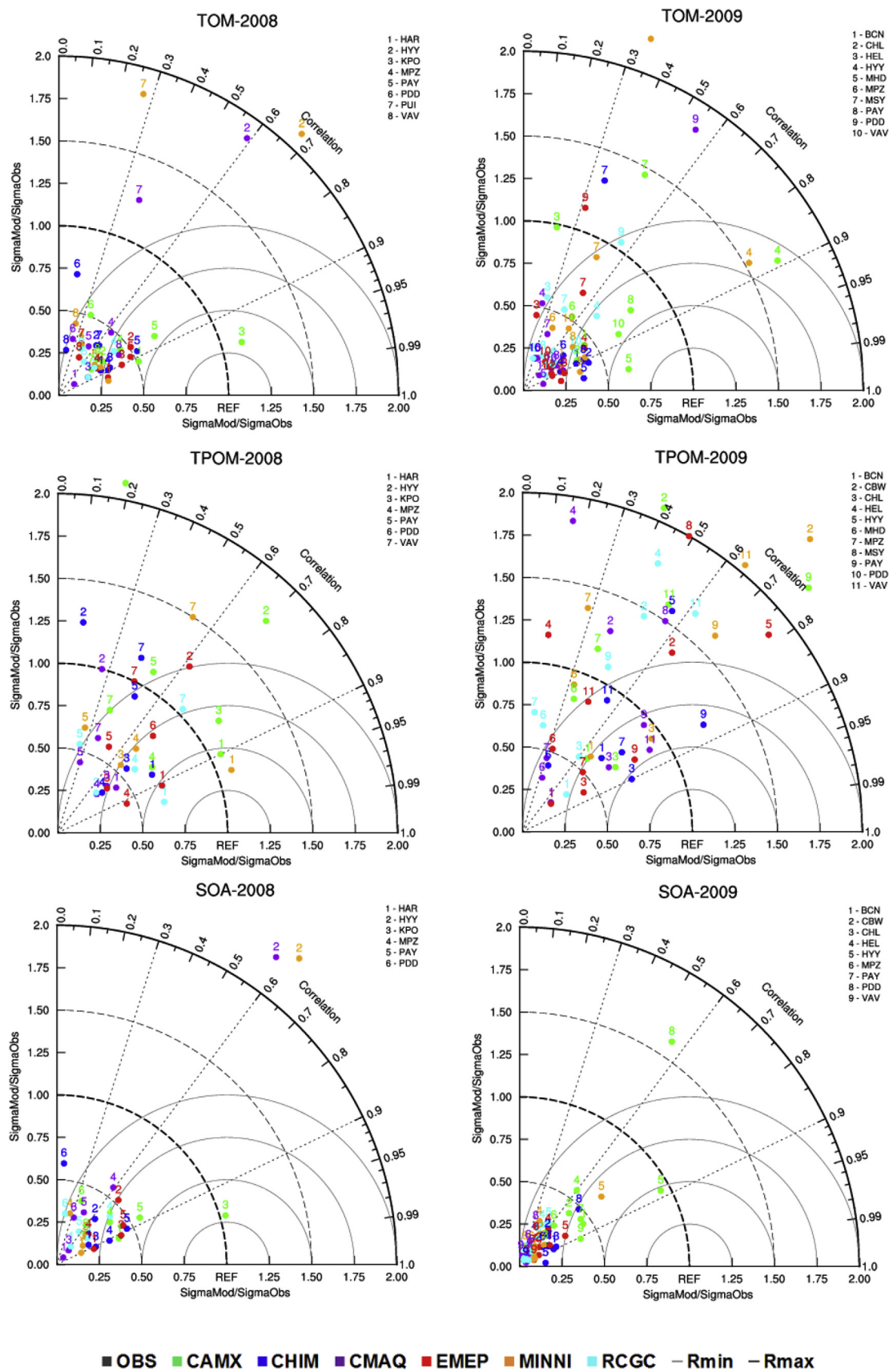


Fig. 11. Taylor diagram for predicted and measured TPOM, TOM and SOA concentrations ($\mu\text{g m}^{-3}$) for 2008 and 2009 EUCAARI/EMEP campaigns. CAMX—light green, CHIMERE—blue, CMAQ—violet, EMEP—red, MINNI—orange, RCGC—light blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Melpitz (MPZ), Payerne (PAY) as can be seen also in Fig. S7. The difficulties of the models to simulate the OA measurements at the high altitude station Puy de Dome (PDD) are partly due to the coarse spatial resolution used in the present intercomparison exercise that makes difficult the horizontal and vertical interpolation, in addition to meteorology, emissions, model formulation that preclude an adequate description of vertical transport in atmospheric models. In spite of this limitation, Barcelona (BCN) and Helsinki (HEL) stations are relative well reproduced by all models (Fig. S7). Fig. 11 also shows that the spread in models' performances is higher for TPOM than for TOM and SOA but this is not confirmed for the other statistical indicators defined in Table S2 and shown in Figs. S7a,b,c. It can be also noted that in 2009, the results of CMAQ for SOA differs from those of the other models as also seen in Fig. 9. Overall, for TOM and SOA, the models have lower variability than the measurements: SigmaMod/SigmaObs values are lower than 1 at almost all stations while for TPOM can be also higher than 1.

As expected, for 2008 period, the mean fractional bias (MFB) values for SOA are similar to MFB values for TOM confirming that the models underestimation of TOM may be attributed to the SOA underestimation (Figs. S7a,b,c). SOA also exhibits the highest diversity in model performances: mean fractional error (MFE) range from 0 to 200, the upper bound of this statistical index. Overall, the analysis of CORR, MFE and MFB at stations (Figs. S7a,b,c) shows that:

- i) for a given compound, the models have similar values at a given station: there are stations where the models' performances are relatively good and stations where the models fail to reproduce OA concentrations;
- ii) for a given station, all models show that a good CORR may correspond to high MFE and MFB values proving the models capacity to capture the main processes contributing to the pollution but not including enough sources to reach the pollution levels;
- iii) for a given station, the models' performances vary with campaign period (season): the diversity in model results is higher in 2009 with respect to 2008 for TPOM and SOA (Figs. S7b and c);
- iv) for all models, MFB and MFE values do not show any sensitivity to the magnitude of measured concentration for any OA compound: the models have similar performances in predicting both low and high concentrations, at least for the range of concentrations investigated;

The curves in Figs. S7a,b,c show the statistical indicators of each model sorted in ascending order to better evidence the differences between the models, irrespective to the measurement station. As expected, the differences between the models' performances are higher for SOA, followed by TOM and TPOM. It can be also observed that CMAQ has the highest MFB and MFE values for SOA as already shown in Fig. 10, but its CORR is similar with the other models.

Figure S8a,b,c,d shows the daily cycles measured and simulated of TOM, TPOM and SOA at stations with data for both seasons: Melpitz (MPZ), Payerne (PAY), Hyytiala (HYY) and Puy de Dome (PDD). The first two are classified as rural, the third as remote and the fourth is a high altitude station. At MPZ (Fig. 8a), the models capture relatively well the hourly evolution of all three OA and underestimate the measured concentrations at any time of the day or night, regardless of season, except for TPOM in 2009. TPOM is overestimated in 2009 at all the stations by most of models, while, in 2008, is underestimated only at the rural stations MPZ and PAY. In both seasons, at MPZ and PAY, TPOM has a well-defined daily cycle, with maxima in the morning and in the evening. This is not observed at HYY and PDD, both stations being located far from anthropogenic emissions sources. The behaviour of the hourly concentration at MPZ indicates a direct relationship with the anthropogenic emission daily cycle for traffic and domestic combustions while other transport and/or chemical reaction pathways seems relevant at the other stations and are not adequately included in

models' input or formulation.

For SOA, at MPZ, the measurements have a well-defined daily cycle with low concentrations during daytime and high concentrations during nighttime, which is not well reproduced by models in 2009. This behaviour of measurements is not observed at the other rural station PAY, and, also at HYY and PDD where the changes of hourly SOA concentrations have an indefinite evolution both for measured and modelled concentrations. The differences among modelled concentrations and between modelled and measured concentrations for SOA are larger than for TPOM (within about $2 \mu\text{g m}^{-3}$ with respect to about $1 \mu\text{g m}^{-3}$, respectively) suggesting that further investigation of the SOA formation process through combination of models' formulation (BVOC and SOA models) in various environmental conditions is necessary. As shown previously, the daily cycles confirms that the errors in TOM predictions are mainly due to the errors in SOA predictions at all four stations investigated. SOA is the major contribution to TOM at all the sites, independently of their location.

4. Conclusions

In general, the CTMs' capabilities to reproduce carbonaceous aerosols in Europe are similar, for both EC and OA compounds. The simulated concentrations of EC and TPOM, which are primary anthropogenic compounds, show similar spatial distribution, reflecting the spatial distribution of anthropogenic emissions. For these components, it can be observed more variability in modelled concentrations due to meteorological conditions (season) than due to model formulations (given that the models used the same anthropogenic emissions). On average, all models calculate for those components the highest concentrations during the cold periods: 2007 and 2009 due to the increase of domestic combustion (residential heating, in particular). For these campaign periods, the strong dependency of average concentrations on the distribution and magnitude of anthropogenic emissions is clearly observed and explain the fewer differences between the models' simulations as well as the relatively good reproduction of the averaged measured concentrations by all models. However, most models have the tendency to underestimate the average EC concentrations; the maximum underestimation is about 60% as reported also by Prank et al. (2016). CAMx and EMEP models tend to predict lower EC concentrations for all campaign periods while CHIMERE and CMAQ overestimate the average EC in 2007 campaign. In case of EMEP, this behaviour could be explained by a relatively thick lowest layer which implies a higher dilution of emissions and by accounting EC ageing which lead to a greater wet scavenging. The dependency of these results on the measurement dataset (day and/or station) can be clearly observed for CAMQ which predicts also very low concentrations, underestimating the ones measured, in 2007 at Melpitz and Montelibretti stations (Figs. 3 and 4). The results also show that, on daily basis, the models reproduce well the drops in EC concentrations related mainly to meteorological based processes such as precipitation episodes, but not the peaks during the episodes caused by the transport of pollutants from areas with emissions not adequately described in the inventories and/or stable atmospheric condition that favour pollutants' accumulation. Since the models used the same emissions and meteorology, the large differences between the models' accuracy of the predictions of peak pollution episodes may be ascribed to different models' physics. In addition to this, the anthropogenic emissions may play an important role as demonstrated by TPOM which has a high contribution to TOM in the cold season (2009) since most of the models overestimate the measured concentrations (Fig. 10).

These results for primary anthropogenic pollutants point out that a future models' evaluation has to be carried out at stations representative for different environmental conditions (rural vs. urban), on hourly basis, for a winter and a summer period (a month at least). Focusing on limited time periods but covered with extensive monitoring data will allow also to improve the understanding of atmospheric

conditions on SOA model predictions with respect to local emissions. Moreover, a substantial enhancement of the understanding in the differences of models' predictions requires more sensitivity tests with different models' configurations (activating or not aerosol dynamics, dry and wet deposition processes, etc.) in addition of using the same input data (emissions, meteorology and boundary conditions) as in ED-III. The models' predictions of OA compounds confirmed that SOA is the main contributor to TOM in all seasons as shown by measurements. Therefore, TOM uncertainties are comparable and mainly due to SOA uncertainties, except the areas with high anthropogenic pollution. All models, at most stations (except high altitude and urban stations), have relatively good correlations, above 0.5, for hourly predicted concentrations of TOM, TPOM and SOA. SOA, and consequently TOM, concentrations are underestimated by all models. This result confirms that there missing processes and/or emissions that characterize SOA concentrations, underestimated more in winter than summer.

SOA is more underestimated in 2009 than in 2008 and the opposite is true for TOM. The overestimations of TPOM concentrations in 2009 compensate for SOA underestimation reducing, thus, the underestimation of TOM concentrations. The daily cycles of TOM, TPOM and SOA are relatively well simulated at rural stations, MPZ and PAY, and, generally, show high underestimations of measured SOA and TOM concentrations. The lack of a well-defined measured daily cycle of SOA at the other stations prevents models based on algorithms that follow daily temperature and radiation cycles from reproducing it.

The simulations show that the models predict comparable levels of SOA, independently of the SOA approach used (VBS: CAMx, EMEP, SORGAM: MINNI, RCG and CHIMERE). The similar performance showed by VBS and SORGAM models suggests that there are relevant processes outside the SOA algorithm that are missing, particularly concerning missing precursors emissions (e.g IVOC) and processes (e.g heterogeneous SOA chemistry). Apart from CAMx and CMAQ, all models show spatial distributions of SOA concentrations highly similar with those of BSOA, which is its main contributor. The differences of modelled BSOA patterns are due to the fact that the models used different approaches to calculate BVOC emissions (MEGAN, BEIS and KB) and different land cover databases (USGS, CCE/SEI for Europe, elsewhere GLC2000, and CORINE 2000) and addition to different approaches for SOA. The comparison of CMAQ and MINNI results point out to the fact that the BVOC emissions from BEIS (CMAQ) may be lower than those from MEGAN (MINNI) leading to less BSOA formed with SORGAM approach. The VBS approaches of CAMx and EMEP produce different spatial distributions of SOA due to different BVOC approaches and land cover databases. Moreover, the VBS approach seems more sensitive to meteorological conditions than SORGAM approach, since the spatial distribution of BSOA varies more from one season to another. The variation is not only due to SOA approach used but also due to variation of BVOC emissions. However, a more accurate evaluation of the SOA contribution and the role of BVOC models and land cover databases to the differences in SOA concentrations cannot be made with the present simulations.

The high variability in SOA model predictions in relation to the season and the location of station requires further investigations with both air quality models and field campaigns. Simultaneous measurements of SOA components and BVOC (SOA precursors) are necessary to better understand and describe SOA and BVOC processes in models for eliminating the OA underestimations. Future modelling exercises need to perform in deep investigations at European scale in order to elucidate the underlying complexity of OA production through sensitivity tests to both input data and changes in model formulation.

Declaration of interest

None.

Acknowledgements

This work was supported by the Co-operative Programme for Monitoring and Evaluation of the Long-range Transmission of Air Pollutants in Europe (EMEP) under UNECE and EU 6th Framework programmes EUCAARI (Contract No. 34684). We would like to express our thanks to all those who are involved in the EMEP monitoring efforts and have contributed through operating sites, performing chemical analysis and by submissions of data.

INERIS was financed by the French Ministry of the Environment, Energy and Marine Affairs. RSE contribution to this work has been financed by the research fund for the Italian Electrical System under the contract agreement between RSE S.p.A. and the Ministry of Economic Development – General Directorate for Nuclear Energy, Renewable Energy and Energy Efficiency in compliance with the decree of 8 March 2006. PSI contribution in this work was funded by the Swiss Federal Office for the Environment (FOEN).

ENEA work was funded by the Italian Ministry for Environment and Territory and Sea and the computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff (<http://www.cresco.enea.it/english>).

Jose L. Jiminez was supported by US EPA STAR 83587701-0. EPA has not reviewed this manuscript and thus no endorsement should be inferred.

Samara Carbone was supported by TEKES CLEEN MMEA - programme, and the Graduate School in Physics, Chemistry, Biology, and Meteorology of Atmospheric Composition and Climate Change (University of Helsinki).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.aeoa.2019.100018>.

References

- Aas, W., Tsyro, S., Bieber, E., Bergström, R., Ceburnis, D., Ellermann, T., Fagerli, H., Frölich, M., Gehrig, R., Makkonen, U., Nemitz, E., Otjes, R., Perez, N., Perrino, C., Prévôt, A.S.H., Putaud, J.-P., Simpson, D., Spindler, G., Vana, M., Yttri, K.E., 2012. Lessons learnt from the first EMEP intensive measurement periods. *Atmos. Chem. Phys.* 12, 8073–8094. <https://doi.org/10.5194/acp-12-8073-2012>.
- Aksoyoglu, S., Keller, J., Ciarelli, G., Prévôt, A.S.H., Baltensperger, U., 2014. A model study on changes of European and Swiss particulate matter, ozone and nitrogen deposition between 1990 and 2020 due to the revised Gothenburg protocol. *Atmos. Chem. Phys.* 14, 13081–13095. <https://doi.org/10.5194/acp-14-13081-2014>.
- Aksoyoglu, S., Ciarelli, G., El-Haddad, I., Baltensperger, U., Prévôt, A.S.H., 2017. Secondary inorganic aerosols in Europe: sources and the significant influence of biogenic VOC emissions, especially on ammonium nitrate. *Atmos. Chem. Phys.* 17, 7757–7773. <https://doi.org/10.5194/acp-17-7757-2017>.
- Baker, K.R., Carlton, A.G., Kleindienst, T.E., Offenberg, J.H., Beaver, M.R., Gentner, D.R., Goldstein, A.H., Hayes, P.L., Jimenez, J.L., Gilman, J.B., de Gouw, J.A., Woody, M.C., Pye, H.O.T., Kelly, J.T., Lewandowski, M., Jaoui, M., Stevens, P.S., Brune, W.H., Lin, Y.-H., Rubitschun, C.L., Surratt, J.D., 2015. Gas and aerosol carbon in California: comparison of measurements and model predictions in Pasadena and Bakersfield. *Atmos. Chem. Phys.* 15, 5243–5258. <https://doi.org/10.5194/acp-15-5243-2015>.
- 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. <https://doi.org/10.1016/j.atmosenv.2012.11.009>.
- Beltman, J.B., Hendriks, C., Tum, M., Schaap, M., 2013. The impact of large scale biomass production on ozone air pollution in Europe. *Atmos. Environ.* 71, 352–363.
- Benedetti, A., Morcrette, J.-J., Boucher, O., Dethof, A., Engelen, R.J., Fisher, M., Flentje, H., Huneeus, N., Jones, L., Kaiser, J.W., Kinne, S., Mangold, A., Razingger, M., Simmons, A.J., Suttie, M., the GEMS-AER team, 2009. Aerosol analysis and forecast in the European Centre for medium-range weather forecasts integrated forecast system: 2. Data assimilation. *J. Geophys. Res.* 114, D13205. <https://doi.org/10.1029/2008JD011115>.
- Bergström, R., Denier van der Gon, H.A.C., Prévôt, A.S.H., Yttri, K.E., Simpson, D., 2012. Modelling of organic aerosols over Europe (2002–2007) using a volatility basis set (VBS) framework: application of different assumptions regarding the formation of secondary organic aerosol. *Atmos. Chem. Phys.* 12, 8499–8527. <https://doi.org/10.5194/acp-12-8499-2012>.
- Bessagnet, B., Menut, L., Curci, G., Hodzic, A., Guillaume, B., Liousse, C., Moukhtar, S.,

- Copenhagen ISBN 978 92 890 0265 3.
- WHO, 2015. WHO Regional Office for Europe, OECD. *Economic Cost of the Health Impact of Air Pollution in Europe: Clean Air, Health and Wealth*. WHO Regional Office for Europe, Copenhagen, Denmark.
- Woody, M.C., Baker, K.R., Hayes, P.L., Jimenez, J.L., Koo, B., Pye, H.O.T., 2016. Understanding sources of organic aerosol during CalNex-2010 using the CMAQ-VBS. *Atmos. Chem. Phys.* 16, 4081–4100. <https://doi.org/10.5194/acp-16-4081-2016>.
- Yarwood, G., Rao, S., Yocke, M., Whitten, G.Z., 2005. Updates to the carbon Bond chemical mechanism: CB05, final report prepared for US EPA. available at: http://www.camx.com/publ/pdfs/CB05_Final_Report_120805.pdf, Accessed date: 8 October 2016.
- Yttri, K.E., Aas, W., Bjerke, A., Cape, J.N., Cavalli, F., Ceburnis, D., Dye, C., Emblico, L., Facchini, M.C., Forster, C., Hanssen, J.E., Hansson, H.C., Jennings, S.G., Maenhaut, W., Putaud, J.P., Tørseth, K., 2007. Elemental and organic carbon in PM10: a one year measurement campaign within the European Monitoring and Evaluation Programme EMEP. *Atmos. Chem. Phys.* 7, 5711–5725. <https://doi.org/10.5194/acp-7-5711-2007>.
- Yttri, K.E., Simpson, D., Bergström, R., Ceburnis, D., Eckhardt, S., Hueglin, C., Kiss, G., Noeygaard, J.K., Perrino, C., Prevot, A., Putaud, J.-P., Spindler, G., Szidat, S., Vanan, M., Zhang, Y.L., Aas, W., 2018. Source Apportionment of the Carbonaceous Aerosol at European Rural Background Sites. (In preparation).
- Zhao, B., Wang, S., Donahue, N.M., Jathar, S.H., Huang, X., Wu1, W., Hao, J., Robinson, A.L., 2016. Quantifying the effect of organic aerosol aging and intermediate volatility emissions on regional-scale aerosol pollution in China. *Sci. Rep.* 6, 28815. <https://doi.org/10.1038/srep28815>.