

# Air quality trends in Europe over the past decade: a first multi-model assessment

A. Colette<sup>1</sup>, C. Granier<sup>2,3,4,5</sup>, Ø. Hodnebrog<sup>6</sup>, H. Jakobs<sup>7</sup>, A. Maurizi<sup>8</sup>, A. Nyiri<sup>9</sup>, B. Bessagnet<sup>1</sup>, A. D'Angiola<sup>2</sup>, M. D'Isidoro<sup>8,\*</sup>, M. Gauss<sup>9</sup>, F. Meleux<sup>1</sup>, M. Memmesheimer<sup>7</sup>, A. Mieville<sup>10</sup>, L. Rouil<sup>1</sup>, F. Russo<sup>8</sup>, S. Solberg<sup>11</sup>, F. Stordal<sup>6</sup>, and F. Tampieri<sup>8</sup>

<sup>1</sup>Institut National de l'Environnement Industriel et des Risques (INERIS), Verneuil-en-Halatte, France

<sup>2</sup>Université Paris 6 Pierre et Marie Curie, CNRS-INSU, LATMOS-IPSL, Paris, France

<sup>3</sup>NOAA Earth System Research Laboratory, Boulder, CO, USA

<sup>4</sup>Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA

<sup>5</sup>Max Planck Institute for Meteorology, Hamburg, Germany

<sup>6</sup>Universitetet i Oslo (UiO), Oslo, Norway

<sup>7</sup>Rhenish Institute for Environmental Research at the University of Cologne (FRIUUK), Köln, Germany

<sup>8</sup>Institute of Atmospheric Sciences and Climate, Consiglio Nazionale Delle Ricerche, Bologna, Italy

<sup>9</sup>Meteorologisk institutt (met.no), Oslo, Norway

<sup>10</sup>Laboratoire d'Aerologie, Toulouse, France

<sup>11</sup>Norsk Institutt for Luftforskning (NILU), Oslo, Norway

\* now at: ENEA, Bologna, Italy

Received: 16 June 2011 – Published in Atmos. Chem. Phys. Discuss.: 4 July 2011

Revised: 19 October 2011 – Accepted: 30 October 2011 – Published: 22 November 2011

**Abstract.** We discuss the capability of current state-of-the-art chemistry and transport models to reproduce air quality trends and interannual variability. Documenting these strengths and weaknesses on the basis of historical simulations is essential before the models are used to investigate future air quality projections. To achieve this, a coordinated modelling exercise was performed in the framework of the CityZEN European Project. It involved six regional and global chemistry-transport models (BOLCHEM, CHIMERE, EMEP, EURAD, OSLOCTM2 and MOZART) simulating air quality over the past decade in the Western European anthropogenic emissions hotspots.

Comparisons between models and observations allow assessing the skills of the models to capture the trends in basic atmospheric constituents (NO<sub>2</sub>, O<sub>3</sub>, and PM<sub>10</sub>). We find that the trends of primary constituents are well reproduced (except in some countries – owing to their sensitivity to the emission inventory) although capturing the more moderate trends of secondary species such as O<sub>3</sub> is more challenging. Apart from the long term trend, the modelled monthly vari-

ability is consistent with the observations but the year-to-year variability is generally underestimated.

A comparison of simulations where anthropogenic emissions are kept constant is also investigated. We find that the magnitude of the emission-driven trend exceeds the natural variability for primary compounds. We can thus conclude that emission management strategies have had a significant impact over the past 10 yr, hence supporting further emission reductions.

## 1 Introduction

Air quality (AQ) management is an essential aspect of environmental policy. Since the major pollution smog events that occurred in the United Kingdom in the 1950s, the awareness of policy makers, economical stakeholders and the general public kept increasing at a steady pace over the last decades. The issue soon became the focus of international negotiations as it appeared that polluting activities in a given country could have a significant impact on the air quality of its neighbours – making internationally coordinated management strategies more relevant at the regional scale. In addition, the need for coordinated political actions was further



Correspondence to: A. Colette  
(augustin.colette@ineris.fr)

justified as it became obvious that the economic cost of innovative technologies and stringent management policies to control pollutant emissions in the competitive and interrelated economic context should be shared and optimised at the European level. Scientific collaboration and multilateral policy negotiation led thus to the 1979 Convention on Long-range Transboundary Air Pollution (LRTAP) and its Gothenburg Protocol accepted in 1999 (UNECE, 1999) as well as the EU National Emissions Ceiling (NEC) Directive (EC, 2001). In 2005, the European Commission published its Thematic Strategy on Air Pollution under the 6th Environmental Action Programme: The Clean Air For Europe (CAFE, 2005) programme has established a long-term policy strategy targeting the adverse effects of air pollution on human health and environment. It determined a set of objectives to be reached within the ongoing revision of the Gothenburg protocol and the NEC directive. Therefore, it is now timely to assess the actual efficiency of the adopted control measures on air quality trends.

European Air Quality management caused the development of operational air pollution monitoring networks throughout the whole of Europe. Such regulatory AQ monitoring networks started in the 1990s, and the observed records are now long enough to assess trends. These initiatives have been accompanied by a number of scientific programmes aimed at improving our understanding of processes playing a role in air quality. Complex numerical models designed to capture air quality variability have been built, and these models reflect our understanding and ability to simulate atmospheric physical and chemical processes. Hence, we now have the suitable tools and observational data for a detailed assessment of our capability to reproduce current atmospheric pollution trends and assess the efficiency of existing control strategies.

Furthermore, the changing economic and industrial context requires periodical revisions of regulations. Currently, the compatibility of climate and air quality policies is questioned and it is unclear whether current mitigation strategies will be as efficient as expected a few decades ahead. Similar impact assessment studies were performed in the context of previous negotiations (Gothenburg Protocol and EC Directive). But uncertainties in emission projections and modelling were high and the actual impact of adopted policies was not correctly foreseen. In the present phase of revision of the emission control legislation, it is thus essential to ensure that current chemistry transport models used to assess the impact of future projections can capture air quality trends and variability over the past decade.

The goal of the present paper is thus to investigate air quality trends and verify if the processes involved are suitably reproduced in existing chemistry and transport models in order to assess their strengths and weaknesses in dealing with policy-related issues such as the impact of future emission projections. To address this question, a coordinated modelling exercise was conducted in the context of

the CityZen Project (megaCITY – Zoom for the Environment, <http://www.cityzen-project.eu/>) funded by the Seventh European Framework Programme for research. The scope was to attempt to reproduce air quality trends in air pollution hotspots with an ensemble of models in order to investigate the performance of existing tools. Six chemistry-transport models were involved: BOLCHEM, CHIMERE, OSLOCTM2, EMEP, EURAD and MOZART, reflecting a variety of approaches: regional or global coverage, online or offline chemistry and transport coupling. Only anthropogenic emissions (based on national totals officially reported within the CLRTAP) were prescribed uniformly for all models while the choice of remaining forcing data (meteorology, biogenic emissions, boundary conditions, etc...) was left open. That way, we ensured the ensemble of simulations would constitute an envelope of trajectories that adequately represents our understanding of the processes involved. The geographical focus is centred on the Western Europe air pollution hotspots constituted by the densely populated cluster of large cities in Benelux, Southern United Kingdom, Western Germany and Northern France. This area was chosen because it is both an area of high emissions and high population exposure. In addition, it offers some degree of homogeneity in terms of economical activities and air pollution regulation trends. The 1998–2007 decade was chosen because of (1) the availability of monitoring data and (2) the robustness of emissions inventories during that period.

This paper is organized as follows: observed air quality trends in the Western European pollution hotspots are investigated in Sect. 2, the modelling setup is presented in Sect. 3 and a short model evaluation is discussed in Sect. 4. The discussion of the capability of the models involved to capture observed trends is detailed in Sect. 5 and the interannual variability is addressed in Sect. 6. Section 7 is devoted to the investigation of the respective roles of anthropogenic emission reduction and meteorological variability on the observed evolution of air pollution.

## 2 Observed air quality trends

### 2.1 Scope and available databases

Before proceeding to the assessment of model performance in terms of air quality trend modelling, we present the observational data that will be used as a reference for the model validation. We limited our scope to the comparison to in-situ surface monitoring stations and we left aside total vertical columns derived by satellite (Kononov et al., 2010) or tropospheric profiles (Thouret et al., 1998; Logan et al., 1999).

Also, we focus only on ozone (O<sub>3</sub>), nitrogen dioxide (NO<sub>2</sub>) and particulate matter with a diameter smaller than 10 μm (PM<sub>10</sub>). Since these basic compounds have been regulated for several years, they are widely monitored, so that we can compile a significant dataset of stations offering a good

coverage (including in urban areas) over the past 10 yr. Unfortunately the same does not hold true for  $\text{PM}_{2.5}$ , whereas this metric would have been better suited to investigate trends in human health exposure.

Building a reliable dataset to assess long term trends is a notoriously difficult task. Two main approaches are found in the literature. The first one consists of using a subset of well documented records (Vautard et al., 2006; Løvblad et al., 2004) follow this strategy by focusing on stations of the EMEP network – i.e. records that are specifically designed for trend assessments. But such stations are all located in rural background areas (because they are designed to monitor transboundary fluxes of air pollution) making it impossible to study urban agglomerations (Derwent et al., 2003; Harrison et al., 2008; Ordonez et al., 2005) include urban sites but limit their geographical scope to a given area – making it possible to check the consistency of individual records. Our aim to document trends over a large hotspot of emissions could thus only be fulfilled by using an alternative approach that consists of relying on a much larger set of stations (at the cost of including sites not designed specifically for trend assessment studies). Here we follow an approach similar to (EEA, 2009) or (Konovalov et al., 2010), considering that the hypothetical degradation of the dataset is compensated by its statistical significance (dubious records having less weights on the statistical indicators inferred).

The focus of the present work being a study of anthropogenic emissions hotspots, regulatory air quality monitoring stations constitute the main source of data. These data were obtained through the public database of the European Environmental Agency AIRBASE (<http://air-climate.eionet.europa.eu/databases/AIRBASE/>, version 3 downloaded in spring 2010).

We also included a few measurements of Sulphate ( $\text{SO}_{4\text{p}}$ ), total Nitrate ( $\text{NO}_{3\text{t}} = \text{NO}_{3\text{p}} + \text{HNO}_{3\text{g}}$ ) and total Ammonia ( $\text{NH}_{4\text{t}} = \text{NH}_{4\text{p}} + \text{NH}_{3\text{g}}$ ) (subscripts are defined as follow: “p” for particulate, “g” for gaseous, “t” for total) collected at remote background sites of the EMEP network (Co-operative programme for monitoring and evaluation of the long range transmission of air pollutants in Europe) reported by the parties of the CLRTAP and available through the EBAS repository (<http://ebas.nilu.no/>). However, we could not gather enough records for a robust assessment of 10-yr trends for these compounds in the emission hotspots. Hence these data will be used exclusively in the model evaluation to discuss the uncertainty of total particulate matter modelling.

## 2.2 Data filtering

The temporal consistency of the record is a major concern in trend assessment studies. This issue is especially relevant when using surface AQ monitoring stations considering that the networks are often designed for population exposure and regulatory purposes rather than trend assessment. As such, the experimental setup can be modified following a change

in the legislation. Monitoring networks have improved significantly since 1998, but unfortunately the present trend assessment has to be based on a fraction of the network that offers a satisfactory coverage of the past decade.

The consistency of the subset used here was ensured using the following three criteria derived from the guidelines of the European Environmental Agency (EEA, 2009):

- the annual coverage should be larger than 75 %;
- at least 8 of the 10 yr between 1998 and 2007 should be recorded;
- a visual screening of each individual record was performed to discard time series with obvious peculiar behaviour. Developing an automated screening algorithm was beyond the scope of the present study. However the subjective character of visual inspection is balanced by a superior capability of detecting a wide spectrum of awkward features. The visual inspection should thus not be considered as a limitation of the present approach as long as the number of discarded records is as small as possible.

The number of selected stations for each constituent and for both the European region (geographic box extending from  $12^\circ \text{W}$  to  $30^\circ \text{E}$  and  $35^\circ \text{N}$  to  $65^\circ \text{N}$ ) and the Benelux region ( $1^\circ \text{W}$ – $8^\circ \text{E}$ ,  $48^\circ \text{N}$ – $54^\circ \text{N}$ ) is given on Table 1. It is noted that the quantitative thresholds on the annual coverage (first two bullet points above) constitute a much more stringent criterion than the subsequent visual inspection.

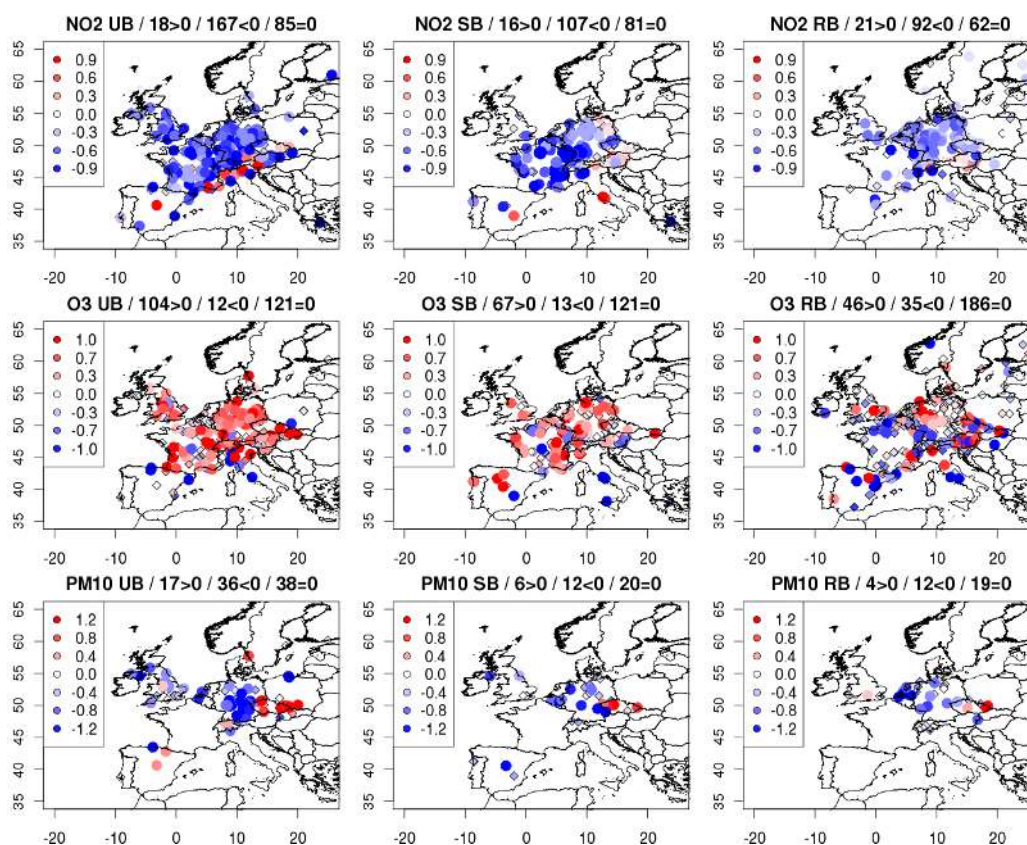
## 2.3 Observed trends

The trends observed at each of the selected stations are displayed in Fig. 1 for  $\text{NO}_2$ ,  $\text{O}_3$  and  $\text{PM}_{10}$ . These trends are computed using time series of monthly averages of daily mean values at each individual location. Each record is de-seasonalised by removing the average seasonal cycle from the raw monthly record and the slope is then computed using a standard linear least square method. Given the fact that the record is only 10 yr long (in the best case), it was considered un-necessary to implement a more elaborate de-seasonalisation procedure. The limited length of the record also led us to focus on linear trends although there are ongoing initiatives to identify change points, piecewise linear or non-linear trends in air quality monitoring (Konovalov et al., 2010; Carslaw et al., 2011). To account for auto-correlation and seasonality, the significance of the trend is assessed with a Mann-Kendall test at the 95 % confidence level (Kendall, 1976; Hipel and McLeod, 2005).

The decrease of  $\text{NO}_2$  concentration is quite robust throughout Europe, except in South-Eastern France and Northern Italy plus a couple of isolated stations. It appears on these maps that the average trend is more pronounced at

**Table 1.** Number of available in-situ surface records obtained from the AIRBASE repository ( $\text{O}_3$ ,  $\text{NO}_2$ , and  $\text{PM}_{10}$ ) or the EMEP network ( $\text{SO}_2$ ,  $\text{SO}_{4\text{p}}$ ,  $\text{NO}_{3\text{t}}$ , and  $\text{NH}_{4\text{t}}$ ) before and after applying the quality check criteria, and for both the whole European domain and the Benelux hotspot.

	Europe			Benelux		
	Total available	8 yr with 75 % annual coverage	Passed visual screening	Total available	8 yr with 75 % annual coverage	Passed visual screening
$\text{O}_3$	1855	717	705	339	166	162
$\text{NO}_2$	1997	669	649	354	160	158
$\text{PM}_{10}$	1533	166	164	252	30	30
$\text{SO}_2$	56	38	37	5	0	0
$\text{SO}_{4\text{p}}$	54	41	37	3	3	3
$\text{NO}_{3\text{t}}$	41	24	22	0	0	0
$\text{NH}_{4\text{t}}$	37	25	21	0	0	0



**Fig. 1.** Trends of daily means of  $\text{NO}_2$ ,  $\text{O}_3$ , and  $\text{PM}_{10}$  ( $\mu\text{g m}^{-3} \text{yr}^{-1}$ ) observed at urban background (UB), suburban background (SB) and rural background (RB) AIRBASE stations. Stations where a statistically significant trend is observed are shown with a large dot a small diamond is used otherwise. The title of each panel also provides the number of stations with a positive, negative or null (not significant) slope.

urban stations: the median trend for all UB (urban background), SB (suburban background) and RB (rural background) stations are  $-0.37$ ,  $-0.27$  and  $-0.14 \mu\text{g m}^{-3} \text{yr}^{-1}$ , respectively. We find an absolute majority of European sta-

tions with a significant negative trend: 62 %, 52 % and 53 % (UB, SB and RB), in line with existing studies with similar temporal and geographical focus (Konovalov et al., 2010; Løvblad et al., 2004; Monks et al., 2009).

These decreasing trends for nitrogen dioxide are reflected in the evolution of  $O_3$  where a slight increase is observed especially at urban sites in and around the Benelux region because anthropogenic emissions are high enough so that a decrease of  $NO_x$  has primarily an impact on the reduction of night-time titration. Hence we find that the average daily mean  $O_3$  trend at UB, SB and RB sites is 0.37, 0.27 and  $0.05 \mu\text{g m}^{-3} \text{yr}^{-1}$ , respectively.

The proportion of sites where the  $O_3$  trend is positive is 30.8% when considering daily means but this number drops to 18.5% when considering  $O_3$  daily peaks, reflecting qualitatively the findings of (Vautard et al., 2006) as they found an opposite trend for background and peak ozone. We find however a smaller difference than in their study because they focused on a different time period (1990–2002) and on remote EMEP stations.

To summarize, the relatively strong decrease of nitrogen oxides over the past decade was unfortunately not accompanied by a sufficient decrease of the other precursors of ozone, notably VOCs, thus leading to moderate observed increases in background ozone in urban areas. Also, in places such as the Benelux region, this decrease was not strong enough to change the photochemical regimes leading to a predominant role of reduced titration of  $O_3$  as a result of  $NO_x$  decrease.

The number of  $PM_{10}$  monitoring stations that pass the filtering described in Sect. 2.2 is by far lower than for  $O_3$  or  $NO_2$ . For instance in France  $PM_{10}$  reporting in AIRBASE started in 2001 and the trend is affected by a change in the metrological correction applied to the measurements in 2007 (Favez et al., 2007) so that no station could be included in the present study. In Germany, UK, and Benelux,  $PM_{10}$  concentrations are systematically decreasing, thanks to the air quality regulation enforced during the past decade. However, the trend of total  $PM_{10}$  levels-off in Northern Germany and the UK as noted by Harrison et al. (2008). In parts of Spain and the Czech Republic, a positive trend is found. This behaviour was discussed by Braniš (2008) who reported a decrease of  $PM_{10}$  during the 1990s due to the economic downturn followed by an increase as a consequence of the increased car traffic in Eastern European countries (of which the Czech Republic is almost the unique representative to pass the filtering of Sect. 2.2).

## 2.4 Sensitivity of the estimated trend to the filtering

In Sect. 2.2, we required somewhat arbitrarily that selected records should cover at least 8 yr in the 1998–2007 decade. One could however question to what extent the findings reported in this paper are sensitive to that threshold.

The number of stations that would have been selected if we had used a threshold of 5, 6, 7, 8, 9, or 10 yr are provided on Table 2 as well as the corresponding median trend. Here we focus only on stations in the larger Benelux region to enhance the homogeneity of the subset. The most stringent criteria of a minimum coverage of 10 yr would have led to a much

smaller subset, hence changing the significance of the set. This is illustrated by looking at the median trend obtained for a 10 yr threshold: it differs from the median obtained with the other thresholds. However it does not mean that it is more representative since the number of stations is lower.

The choice of a 8-yr threshold is further justified by the comparison of the distribution of trends for various thresholds. Figure 2 displays quantile-quantile plots of the distribution of  $PM_{10}$  ( $O_3$  and  $NO_2$  not represented for concision purposes) trends in the Benelux area, taking as reference the distribution of de-seasonalised trends covering at least 8 yr over the 1998–2007 decade. It appears that central quantiles are quite insensitive to that threshold, but the tails of the distribution can be dramatically different when using only stations that cover 5, 6, or 10 yr.

A closer look at the median of the distribution of trends (Table 2) shows that strong differences can be found when using different criteria, even if the sign of the trend is quite robust. The table also features the result of the Wilcoxon test (Hollander and Wolfe, 1999) that measures the similarity of two distributions by comparing the rank of individuals (offering a similar yet more quantitative information than the quantile-quantile plot). The p-value of that test is given, it provides the probability that the distribution is similar to the distribution obtained when using all stations covering at least 8 yr. It can be seen that for all three pollutants, using 7, 8 or 9 yr as a threshold yields similar distributions. But using 5, 6, or 10 yr as a threshold would give quite different estimates.

## 3 Modelling setup

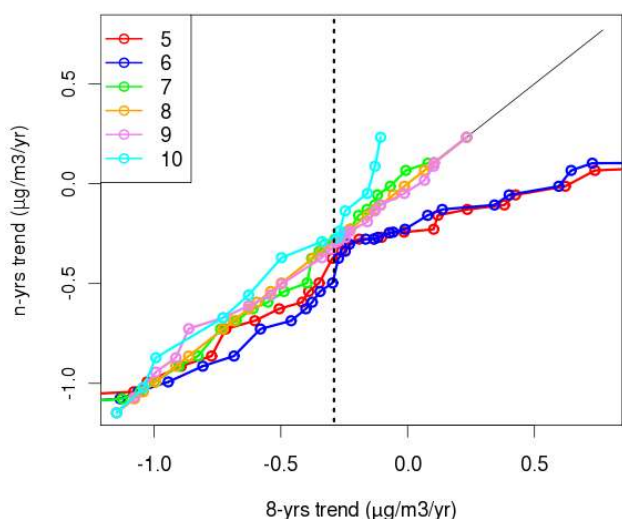
In order to produce an ensemble of models that best represents our ability to capture air quality trends, it was decided to keep the modelling setup as flexible as possible, the only restriction being to use the same emission inventory for anthropogenic emissions. As such, the present experiment is not a model inter-comparison initiative, but rather an attempt to assess the uncertainties in air pollution trend modelling.

### 3.1 Inventory of anthropogenic emissions

We use the EMEP emission inventory (Vestreng et al., 2005) which is based on official emission data reported by individual countries under the LRTAP convention. This inventory is the most widely used and the only available for the whole decade 1998–2007. When launching the experiment (August 2009) only the 1998–2007 period was available from the website <http://www.emep.int>. Beyond the European domain (for global CTMs), these emissions are merged into the so-called MACCity inventory (Granier et al., 2011). Several published studies documented the shortcomings of anthropogenic emission inventories in general (Granier et al., 2011; Monks et al., 2009) or for the EMEP inventory in particular (Vestreng et al., 2009; Jonson et al., 2006; Kononov

**Table 2.** Sensitivity of the trend computed in the Benelux region to the threshold used in the quality checking procedure. For each minimum number of years covered and each pollutant, we provide the number of available stations, the p-value of the Wilcoxon test of similarity of the distributions compared to the reference (with a 8-yr threshold), and the median of the distribution of the trends at all stations.

	NO <sub>2</sub>		Median trend	O <sub>3</sub>		Median trend	PM <sub>10</sub>	Median trend	
	# sta.	p-val.	( $\mu\text{g m}^{-3} \text{ yr}^{-1}$ )	# sta.	p-val.	( $\mu\text{g m}^{-3} \text{ yr}^{-1}$ )	# sta.	p-val.	( $\mu\text{g m}^{-3} \text{ yr}^{-1}$ )
5	194	0.69	-0.343	203	0.91	0.188	91	0.08	-0.228
6	180	0.60	-0.322	186	0.92	0.204	65	0.06	-0.222
7	167	0.87	-0.344	174	0.99	0.204	39	0.91	-0.303
8	158	1	-0.347	162	1	0.204	30	1	-0.290
9	110	0.54	-0.366	132	0.72	0.178	25	0.93	-0.277
10	49	0.33	-0.315	51	0.49	0.203	13	0.46	-0.338



**Fig. 2.** Comparison of the distributions of PM<sub>10</sub> trends ( $\mu\text{g m}^{-3} \text{ yr}^{-1}$ ) in Benelux, depending on the threshold of minimum number of years used in the data quality check procedure (8 yr being used as a reference on the x-axis). The vertical dashed line shows the median of the reference distribution.

et al., 2006). The main advantage of this inventory is that it is built from officially reported national emissions. Nevertheless because of national regulatory issues, some countries might choose not to report emission for given activity sectors, which leads to the problem of data completeness and gap-filling process. In a context of changing regulatory context, this issue is especially relevant for trend assessment. In addition to this issue which is specifically relevant for the EMEP inventories, other shortcomings are well known for any anthropogenic emissions inventory. The injection height or temporal (seasonal or daily) profiles are also an issue. Last but not least, only a few species are reported and hypothesis have to be made regarding the chemical speciation (the ratio between NO and NO<sub>2</sub> amongst the total NO<sub>x</sub>, the speciation of primary particulate matter or volatile organic compounds).

In addition to the above shortcomings, a couple of hypothesis had to be made to these emissions to improve their inter-annual consistency. Particulate matter (PM) emitted in North African countries were not reported before 2007, hence they were reset to zero for that year. There are no reported PM emissions over the sea areas in 1999, hence for that year, and over the sea exclusively, we used PM emissions reported for 2000. There are no PM emissions reported in 1998, hence we used PM emissions of 1999. These assumptions will certainly have an impact on the results discussed below, especially over sea surfaces where PM emission are constant during the first three years of the decade.

For each grid point of the inventory we fitted a linear least square regression for the total emissions of PM, NO<sub>x</sub>, and non-methane volatile organic compounds, as well as the NMVOC/NO<sub>x</sub> ratio and we plotted the map of the slope on Fig. 3. NO<sub>x</sub> emissions decreased throughout Europe, except on the ship tracks because of a significant increase of the traffic (Eyring et al., 2010; Endresen et al., 2007). NMVOC decreased also, except in Poland. And the trend in the ratio NMVOC/NO<sub>x</sub> shows some interesting patterns with regards to the modelled trends of O<sub>3</sub> that will be discussed later. Note that the trend of primary PM emission is much more variable geographically.

### 3.2 Chemistry transport models

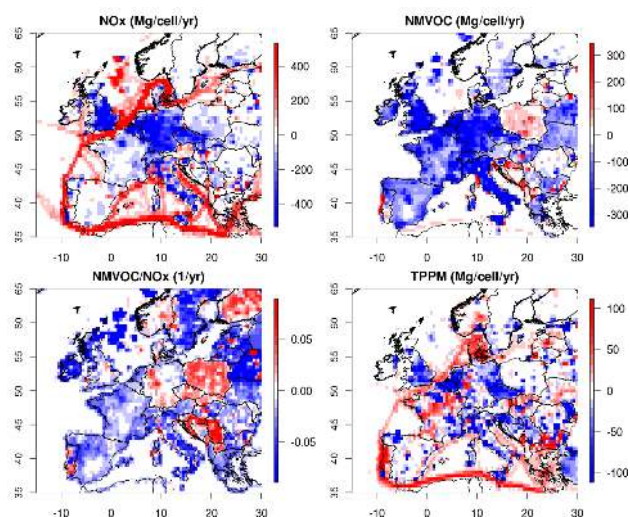
The main technical characteristics of the four regional and two global chemistry-transport models used in the present study are summarized in Table 3.

#### 3.2.1 BOLCHEM

The BOLCHEM model is developed by the Institute of Atmospheric Sciences and Climate of the Italian National Council of Research. It is an online coupled atmospheric dynamics and composition model. The meteorological part is BOLAM (<http://www.isac.cnr.it/~dinamica/bolam>) while the composition part deals with gas and aerosol chemistry and

**Table 3.** Technical characteristics of the chemistry-transport models used in the present study.

		BOLCHEM	CHIMERE	EMEP	EURAD	OSLOCTM2	MOZART 4
Chemical Mechanism	Gas-phase	SAPRC90 (Carter, 1990)	Melchior 2 reduced: 44 species, 120 reactions (Latuati, 1997).	EMEP (Simpson et al., 1993), with subsequent updates (Simpson et al., 2011)	RADM2, RACM-MIM (Stockwell et al., 1997; Geiger et al., 2003; Karl et al., 2006)	QSSA solver (Hestvedt et al., 1978) (Berntsen and Isaksen, 1997) for ozone/NO <sub>x</sub> /hydrocarbon (51 species, 104 reactions)	MOZART chemical scheme (Emmons et al., 2010)
	Aerosols	Aero3 (Binkowski and Roselle, 2003)	ISORROPIA (equilibrium), 14 aerosols species, 8 size bins (Nenes et al., 1998)	EQSAM (Metzger et al., 2002)	MADE/SORGAM (Ackermann et al., 1998; Schell et al., 2001)	M7 aerosol model (Vignati et al., 2004) for BC/OC, sea salt, dust and sulfate. Nitrate according to Myhre et al. (2006)	MOZART aerosol scheme (Emmons et al., 2010), includes BC/OC, Sulphate, Nitrate, Secondary Organic aerosols
Geometry	Modelling domain (resolution)	Europe (0.5 deg)	Europe (0.5 deg)	Europe (50 km)	Europe (50 km)	Global (T42, ~2.81 deg)	Global (T63, ~1.8 deg)
	Number of vertical levels (min, max)	16 (40 m a.g.l.–500 hPa)	8 (30 m a.g.l.–500 hPa)	20 (90 m a.g.l.–100 hPa)	23 (20 m a.g.l.–100 hPa)	60 (8 m a.g.l.–0.1 hPa)	28
Meteorology	Mesoscale model	Hydrostatic limited area model BOLAM coupled online with atmospheric composition modules (Buzzi et al., 1994)	WRF v3.2.1. (Skamarock et al., 2008). Simulations compliant with the CORDEX requirements (Giorgi et al., 2009)	PARLAM-PS (for 1998–2006), HIRLAM (2007) (Undén et al., 2003)	MM5 (Grell et al., 1994)	N/A	N/A
	Large scale Forcing	ERA-interim.	ERA-Interim	IFS (ECMWF) run in-house with data assimilation	NCEP/GFS	IFS (ECMWF) run in-house with data assimilation	NCEP/NCAR analyses
Boundary Conditions		LMDz-INCA monthly climatology (1997–2001) (Hauglustaine et al., 2004)	LMDz-INCA monthly climatology (1997–2001) (Hauglustaine et al., 2004)	O <sub>3</sub> : observation-based climatology modulated by potential vorticity, latitude and inter-annually (Logan, 1999)	O <sub>3</sub> : observation-based climatology modulated inter-annually (Logan, 1999)	N/A	N/A
Biogenic emissions		Isoprene, monoterpenes and other VOCs computed according to Symeonidis et al. (2008)	MEGAN v. 2.04 (Guenther et al., 2006)	Isoprene computed according to Simpson (1995)	Isoprene and Monoterpene emissions according to Lamb et al. (1993)	POET (Granier et al., 2005; Olivier et al., 2003)	MEGAN v2.04 (Guenther et al., 2006)
Natural emissions		none	none	Volcano emissions for Italy according to EMEP emission data base	none	Volcanic emissions of SO <sub>2</sub> taken from Spiro et al. (1992) with vertical distribution from Graf et al. (1997).	POET (Granier et al., 2005)
Biomass burning		none	GFED monthly (A. Heil, personal communication, 2010)	N/A	none	RETRO (monthly, 1998–2000) (Schultz et al., 2008) and GFEDv2 (8-days, 2001–2007) (Van Der Werf et al., 2006)	GFED monthly (A. Heil, personal communication, 2010)



**Fig. 3.** Map of EMEP expert emissions trends (linear least square fit of annual totals) over 1998–2007 for  $\text{NO}_x$ , NMVOC,  $\text{NMVOC}/\text{NO}_x$  and total primary PM (TPPM). Units are  $\text{Mg}/\text{yr}$  except for  $\text{NMVOC}/\text{NO}_x$  ( $\text{yr}^{-1}$ ).

physics. More details can be found in Mircea et al. (2008) and at <http://bolchem.isac.cnr.it>.

### 3.2.2 CHIMERE

The CHIMERE model is developed, maintained and distributed by Institut Pierre Simon Laplace (CNRS) and INERIS (Bessagnet et al., 2008). It is used for daily operational forecasting in France (Honoré et al., 2007) and beyond (e.g. through the MACC project of the European Global Monitoring for Environment and Security Programme) as well as long-term studies (Vautard et al., 2006; Beekmann and Vautard, 2010). More details can be found on the website: <http://www.lmd.polytechnique.fr/chimere>.

### 3.2.3 EMEP

The EMEP model is a Eulerian Chemical Transport Model developed at the EMEP Centre MSC-W, hosted by the Norwegian Meteorological Institute. It has been publicly available as Open Source code since 2008. The latest version can be obtained from <https://wiki.met.no/emep/page1/unimodopensource2011>. The model has been documented by Simpson et al. (2003) and Simpson et al. (2011). It is used to provide the scientific basis to the LRTAP convention, in particular for establishing source-receptor relationships of air pollution, but also for daily chemical weather forecasting within the MACC project.

### 3.2.4 EURAD

The EURAD model (Jakobs et al., 2002; Memmesheimer et al., 2004, 2007) is used to carry out chemical transport

simulations for the area considered. The model calculates the transport, chemical transformations and deposition of air pollutants in the troposphere from the surface up to about 16 km. It is being implemented operationally for daily forecast in Germany and beyond in the framework of the European project MACC. Meteorological fields are provided by the meteorological model MM5. Gas phase kinetics is computed using the RACM-MIM chemistry mechanism (Geiger et al., 2003). The MADE-SORGAM (Schell et al., 2001) model is used to account for the formation of secondary organic and inorganic particles in the atmosphere. More details can be found on the website: <http://www.eurad.uni-koeln.de>.

### 3.2.5 OSLOCTM2

OSLOCTM2 is a global offline chemistry transport model driven by ECMWF meteorological data (Isaksen et al., 2005; Søvde et al., 2008). In this study the model was run with tropospheric and stratospheric chemistry including both gas-phase chemistry using the Quasi Steady-State Approximation (Hesstvedt et al., 1978; Berntsen and Isaksen, 1997), and aerosols using the M7 (Vignati et al., 2004) and nitrate (Myhre et al., 2006) modules. The period 1997–2007 was simulated, the first year of which was considered as spin-up. In OSLOCTM2 advection is done using the second order moment scheme (Prather, 1986), convection is based on the Tiedtke mass flux parameterization (Tiedtke, 1989), and transport in the boundary layer is treated according to the Holtslag K-profile method (Holtslag et al., 1990). The calculation of dry deposition is based on Wesely (1989).

### 3.2.6 MOZART

MOZART (Model for OZone And Related chemical Tracers) is a chemistry transport model (CTM) developed jointly by the (US) National Center for Atmospheric Research (NCAR), the Geophysical Fluid Dynamics Laboratory (GFDL), and the Max Planck Institute for Meteorology (MPI-Met) to simulate the distribution of gaseous and particulate compounds in the Earth's atmosphere. The MOZART-4 version of the model (Emmons et al., 2010) was used in this study. The MOZART-4 source code and standard input files are available for download from the NCAR Community Data Portal (<http://cdp.ucar.edu>).

## 4 Model evaluation

The present model ensemble was designed to assess the capability of state-of-the-art chemistry transport models to capture the trends of main pollutants. This chapter presents a short model evaluation to understand where the models stand.

The  $\text{O}_3$ ,  $\text{NO}_2$  and  $\text{PM}_{10}$  scores of each model compared to AIRBASE suburban stations are given on Table 4. Only one type of station is discussed for concision purposes. Bias,



**Table 4.** Model performances at AIRBASE suburban stations computed over 10 yr on the basis of daily means.

	NO <sub>2</sub>			O <sub>3</sub>			PM <sub>10</sub>		
	R	Bias ( $\mu\text{g m}^{-3}$ )	RMSE ( $\mu\text{g m}^{-3}$ )	R	Bias ( $\mu\text{g m}^{-3}$ )	RMSE ( $\mu\text{g m}^{-3}$ )	R	Bias ( $\mu\text{g m}^{-3}$ )	RMSE ( $\mu\text{g m}^{-3}$ )
BOLCHEM	0.658	-1.46	12.6	0.782	-4.31	17.6	0.37	-4.48	13.9
CHIMERE	0.633	-13.4	16.6	0.797	18.5	23.5	0.576	-11.3	15.5
EMEP	0.574	-14.1	17.4	0.74	13.3	21	0.542	-14.4	17.9
EURAD	0.644	-6.46	16.0	0.659	15	25.2	0.524	-3.75	15.1
OSLOCTM2	0.544	-14.7	18.0	0.75	17.1	23.5			
MOZART	0.35	-18.1	21.3	0.627	4.76	22			

Root Mean Square Error (RMSE) and correlation are all computed from daily mean values. Note that aggregated metrics or daily maxima are often used for model performances assessment but daily values were considered more appropriate for the investigation of trends. Figure 4 displays the mean seasonal cycles (monthly values based on 10 yr of daily means) observed and modelled at AIRBASE stations. Model and data are displayed for all types of stations (UB, SB, RB) even if the models do not capture very well the variability brought about by the typology of the stations. Modelling a whole decade could only be achieved at the cost of using a relatively coarse spatial resolution, making it difficult to reproduce the differences between UB, SB and RB stations.

#### 4.1 Nitrogen dioxide

All models exhibit a negative NO<sub>2</sub> bias when compared to suburban stations. This feature was expected as we used at best a 50 km spatial resolution. The small bias of modelled NO<sub>2</sub> levels with BOLCHEM and EURAD were however unexpected at this resolution. They are probably the result of a different representation of the vertical mixing as suggested by their strong seasonal cycle but we cannot rule out an influence of heterogeneous chemistry in the NO<sub>x</sub> removal (which would be corroborated in the strong difference on total PM<sub>10</sub> discussed below in Sect. 4.4). The other models perform better when compared to RB stations, as expected given the resolution.

Note that the average bias of global models is in-line with RCTMs such as CHIMERE and EMEP even considering their much coarser resolution. This result was not expected and constitutes an interesting finding of the study. However, the comparison might have been less favourable to GCTMs if we had focused on higher-quantile metrics (such as daily maximum values that were unfortunately unavailable in some of the global model outputs).

It is also interesting to point out the moderate importance of the seasonality in emissions. All regional models use the seasonal profile recommended by EMEP while – in these simulations – global models have no seasonality in anthropogenic emissions. The results shown on Fig. 4 show that

the main driver of seasonality is probably not the prescribed cycle of emission but rather other factors such as vertical mixing (main driver of the wintertime maximum) or biogenic emissions (that could be responsible for the summer secondary maxima modelled by OSLOCTM2).

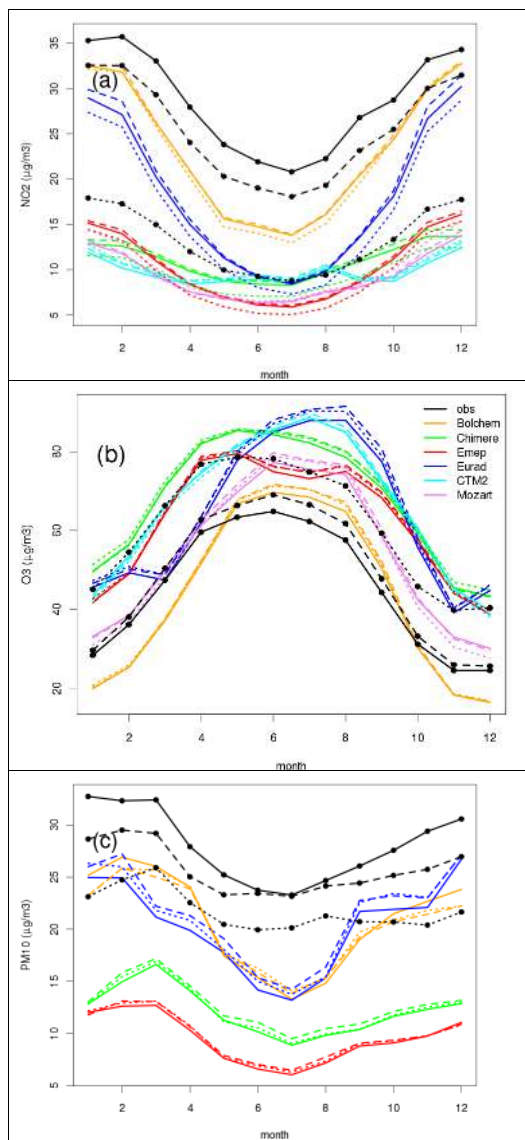
#### 4.2 Ozone

As far as ozone is concerned, the results are in line with previous model inter-comparison initiatives (van Loon et al., 2007; Vautard et al., 2009). BOLCHEM is the only model to have a negative (albeit small) bias at suburban stations, owing to the larger NO<sub>2</sub> concentrations compared to other models. All the other regional CTMs show a positive bias. The best example of this behaviour is CHIMERE that has the largest bias but a very good correlation, hence similar RMSE scores than the other models.

The seasonal cycle of ozone is also very insightful (Fig. 4b). The springtime ozone build-up is quite consistent in all models but the summer time behaviour is very different. The correlation of this average monthly cycle (compared to observations) is 0.97, 0.99, 0.95, 0.85, 0.96 and 0.96 for BOLCHEM, CHIMERE, EMEP, EURAD, OSLOCTM2 and MOZART, respectively. Average O<sub>3</sub> concentrations level off between June and August in CHIMERE, EMEP and BOLCHEM (and in the observations), while they keep increasing according to EURAD, MOZART and OSLOCTM2. This characteristic is attributed to the reactivity of the chemical mechanism. A couple of peculiar features could not be explained such as the wintertime secondary maximum modelled by EURAD and the summertime secondary minimum of EMEP. We checked however that these features were not induced by a single event and found that they were recurrent every year over the decade.

#### 4.3 O<sub>x</sub>

The O<sub>x</sub> (=NO<sub>2</sub> + O<sub>3</sub>) climatology (global average over 10 yr) is displayed on Fig. 5. By filtering out the titration impact of NO<sub>x</sub> on O<sub>3</sub> levels, this quantity gives an insight into the degree of photochemical



**Fig. 4.** Seasonal cycles based on 10 yr of daily mean values for  $\text{NO}_2$  (top),  $\text{O}_3$  (middle), and  $\text{PM}_{10}$  (bottom) observed (black) and modelled by BOLCHEM (orange), CHIMERE (green), EMEP (red), EURAD (blue), OSLOCTM2 (cyan) and MOZART (violet). Solid lines for urban background, dashed: suburban, and dotted: rural AIRBASE stations.

activity of the models. BOLCHEM appears as one of the least photochemically active models (spatial and temporal global average of  $62.6 \pm 6.2 \mu\text{g m}^{-3}$ ), and to a lesser extent MOZART is also in the lower part of the sample ( $63.9 \pm 9.8 \mu\text{g m}^{-3}$ ). OSLOCTM2 ( $70.4 \pm 13.9 \mu\text{g m}^{-3}$ ), EMEP ( $72.8 \pm 11.1 \mu\text{g m}^{-3}$ ) and EURAD ( $74.7 \pm 14.9 \mu\text{g m}^{-3}$ ) exhibit more similar figures while CHIMERE ( $80.8 \pm 10.0 \mu\text{g m}^{-3}$ ) is the most active. Note that the spatial variability is high as shown on the maps as well as in the standard deviation given in brackets above.

Hence these global averages are not representative of the photochemical activity over populated areas, where only CHIMERE, EURAD and OSLOCTM2 can be considered as more active. All models but BOLCHEM show very high  $\text{O}_x$  concentrations above the Mediterranean. Note also the strong influence of  $\text{O}_3$  dry deposition schemes as shown by the sharp land/sea gradient.

#### 4.4 Particulate matter

$\text{PM}_{10}$  scores (Table 4) are not available in global model outputs which usually calculate BC/OC rather than total particulate matter.  $\text{PM}_{10}$  correlations are much lower than for  $\text{NO}_2$  or  $\text{O}_3$ ; which is a commonplace feature in such studies. Biases are consistently negative but slightly lower in magnitude for BOLCHEM and EURAD. We will see below that this could be due to a compensation of errors, the bias for ammonium, nitrate and sulphate being quite high for these models. Again, the seasonal cycle (Fig. 4c) is much more pronounced for BOLCHEM and EURAD than for CHIMERE and EMEP, the first two models are subsequently better compared with urban and suburban stations, while the latter two are more representative of rural stations.

#### 4.5 Nitrate, ammonium and sulphate

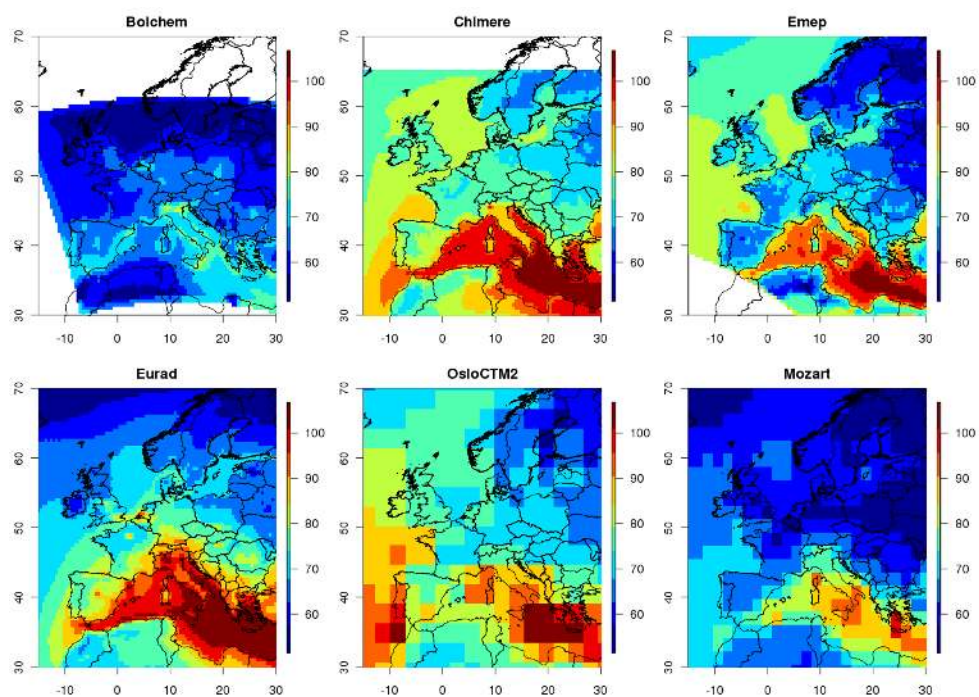
The overestimation of  $\text{NH}_4^+$  and  $\text{NO}_3^-$  mentioned above for BOLCHEM and EURAD can be seen on Fig. 6. EMEP, CHIMERE and OSLOCTM2 have a lower bias compared to the EMEP observations, and the seasonal cycle is quite synchronous with the observations for the last two. The seasonal cycle of MOZART is however slightly stronger. Gaseous sulphur dioxide is well captured by EMEP and CHIMERE but EURAD and BOLCHEM produce a strong overestimation as well as a too strong seasonal cycle. Performances in terms of particulate sulphate are very variable, the best seasonal cycle being that of the EMEP model, while EURAD and OSLOCTM2 exhibit a too strong seasonal cycle attributed by Berglen et al. (2004) to missing oxidation pathways in wintertime, especially by  $\text{H}_2\text{O}_2$ .

### 5 Modelled trends

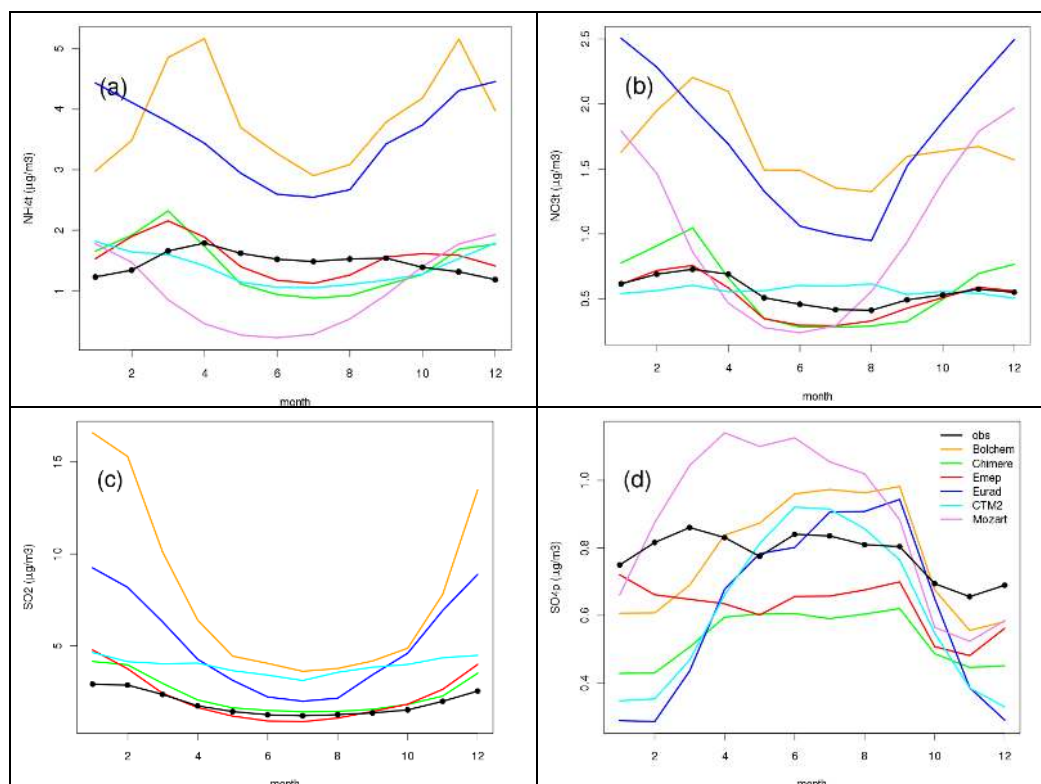
The capability of chemistry transport models to capture the observed trends of major atmospheric pollutants is discussed in this section.

#### 5.1 Nitrogen dioxide

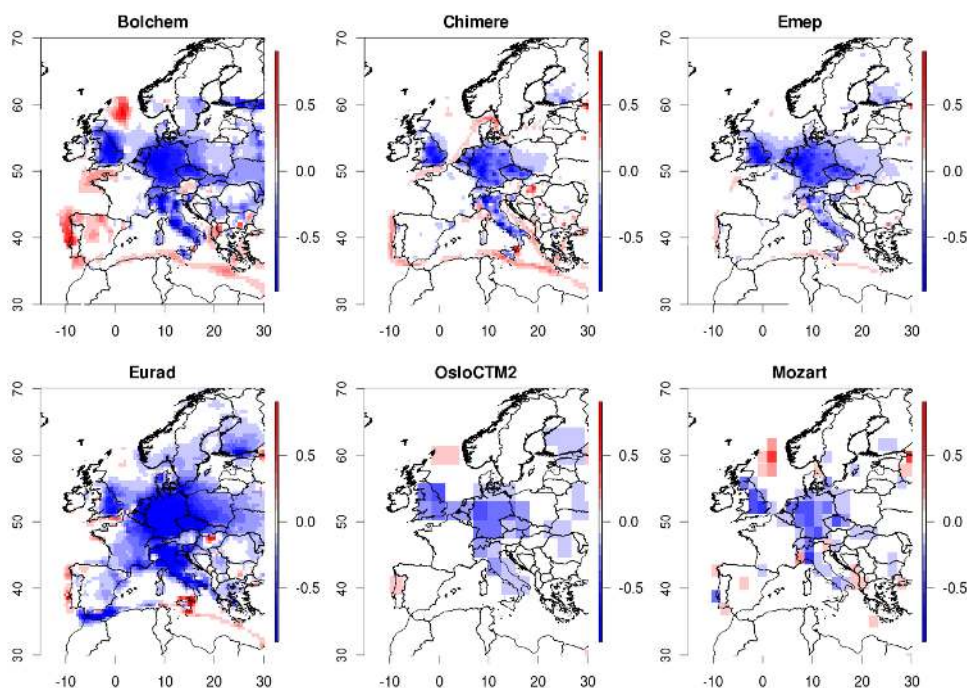
The modelled trend of  $\text{NO}_2$  over the whole of Europe is shown in Fig. 7 for each model. The main feature is a pronounced decrease over most of Western Europe (more specifically United Kingdom, Germany, Benelux and Italy) except France and Spain, reflecting the trend of primary emission reductions reported in the inventory (Fig. 3). By contrast  $\text{NO}_2$



**Fig. 5.** Average modelled  $O_x$  ( $\mu\text{g m}^{-3}$ ) [ $=\text{NO}_2 + \text{O}_3$ ] fields over 10 yr for each model.



**Fig. 6.** Same as Fig. 4 for total  $\text{NH}_{4t}$  (a), total  $\text{NO}_{3t}$  (b),  $\text{SO}_2$  (c) and  $\text{SO}_{4p}$  (d) recorded at EMEP background stations.



**Fig. 7.** Modelled  $\text{NO}_2$  trend ( $\mu\text{g m}^{-3} \text{yr}^{-1}$ ) for each CTM and at each grid point computed on the basis of monthly means of daily means over the 1998–2007 period with a linear least square fit of de-seasonalised values.

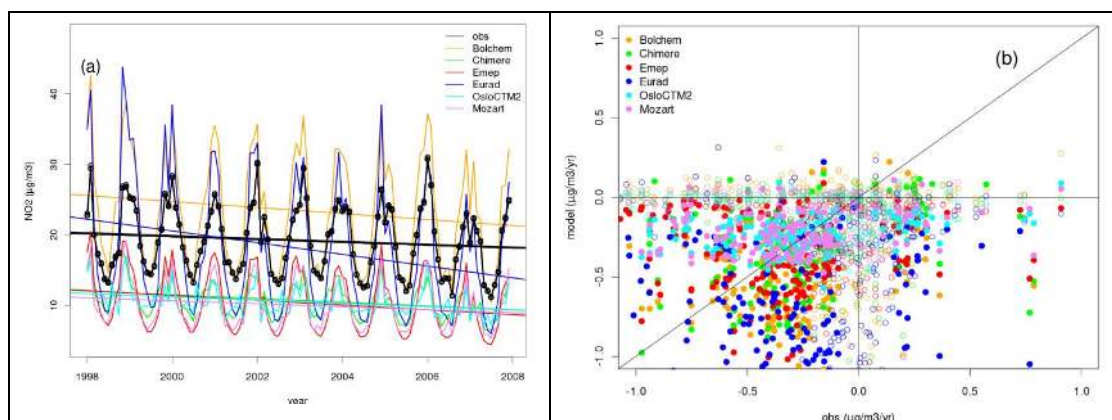
tends to increase over the main ship tracks. These dominating patterns are consistently captured by all models. An exception is seen in EURAD which calculates a wider extent of the  $\text{NO}_2$  decreasing trend (especially in France), even reaching the ship track north of Morocco and Algeria. The use of identical anthropogenic emissions rules out the evolution of ship emissions to explain this feature. Meteorology is probably a dominating factor here as the PBL depth (not shown) appears to exhibit a positive trend in the EURAD simulation in that area, explaining the increased dilution of  $\text{NO}_2$ .

Before proceeding to the quantitative assessment of model performances, a visual comparison of the modelled (Fig. 7) and observed (Fig. 1) geographical patterns of these trends suggests that the models are quite successful in capturing  $\text{NO}_2$  trends, especially in the UK, Germany, Benelux and Czech Republic. The lack of decrease or slight increase over Spain, Poland and Austria is reproduced as well as the more noisy behaviour over Italy. However, the models seem to underestimate the trends in France. The fact that these patterns match quite well with national boundaries suggests that total emissions reported to EMEP at the national level may play a significant role here, as will be confirmed below.

A more detailed comparison of modelled versus observed trends is provided in Fig. 8. The composite time series on panel (a) consists in an average of all monthly time series observed and modelled at AIRBASE background rural and background suburban stations. It reflects some of the findings discussed in Sect. 4.1 in terms of  $\text{NO}_2$  model performances.

It also shows that EURAD and BOLCHEM behave very similarly at the beginning of the decade, while the  $\text{NO}_2$  decrease by the end of the period is much stronger in EURAD. All other models exhibit very similar behaviours.

While the composite on Fig. 8a offers a visual picture of the trend, it consists in an average of stations spread across the whole of Europe, hence aggregating different trends. Panel (b) of Fig. 8 shows the scatter between observed and modelled trends (defined as the slope of the de-seasonalised monthly mean time series) at each individual station. Such a result requires that each individual record is sufficiently reliable to assess a trend, which could only be achieved with the subset of long term time series presented in Sect. 2.2. This figure un-ambiguously shows that the correlation between modelled and observed trends is not perfect. Even if all the models used in the present study obtain decent scores in capturing  $\text{NO}_2$ , the interannual trend appears to be more challenging and most points are located quite a distance away from the 1-1 line on that scatter plot. Nevertheless the sign of the trend seems to be quite well captured at most locations; a hit-rate metric (percentage of sites where the sign of the trend is captured by the model) for model performance is thus preferred to a quantitative correlation. When considering only stations where a significant  $\text{NO}_2$  trend is measured (according to the Mann-Kendall test, see Sect. 2.3) – i.e. 235 background suburban and rural sites – the sign of the trend is well captured at 68, 72, 81, 80, 70 and 67 % of the stations for BOLCHEM, CHIMERE, EMEP, EURAD,



**Fig. 8.** (a) European-wide composite of modelled and observed monthly means of NO<sub>2</sub> trend ( $\mu\text{g m}^{-3}$ ) at the air quality monitoring stations of background suburban and rural type. The straight line shows the best linear least square fit. (b) Scatter plot of modelled and observed trend (computed as linear least square of the de-seasonalised time series, in  $\mu\text{g m}^{-3} \text{ yr}^{-1}$ ) at each individual station. Sites where a significant slope is computed are marked with a filled symbol.

**Table 5.** Fraction of sites where the sign of the NO<sub>2</sub> trend is correctly captured by the models (average – avg – and standard deviation –  $\sigma$  – of the individual fraction correct of each model) for the countries where a significant trend is observed at 5 stations at least (number of selected stations – nst – provided on the last row).

	Europe	AT	CH	CZ	DE	FR	NL
Avg ( $\sigma$ )	0.73 (0.06)	0.56 (0.07)	1 (0)	0.86 (0)	0.8 (0.05)	0.39 (0.29)	0.8 (0.25)
nst	235	14	6	7	35	12	5

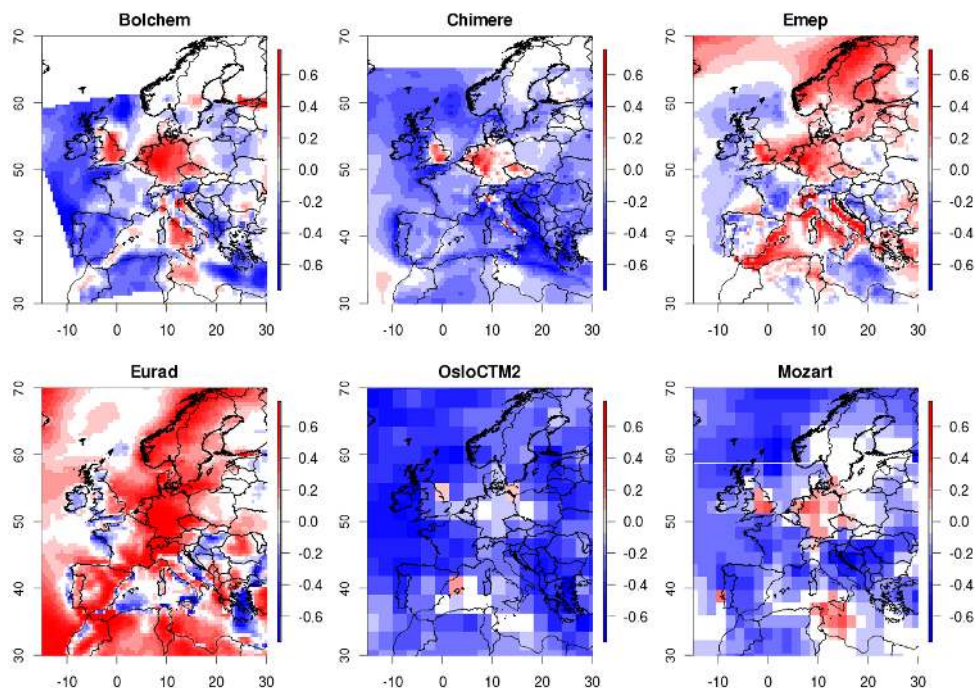
OSLOCTM2 and MOZART, respectively. That is a good overall performance of 73 % ( $\sigma = 6$  %) on average across all models. This indicator varies widely on the country-level basis for rural sites (Table 5), the scores are much worse for all models in France and Austria. In Austria the trends are small in magnitude, making it more challenging to capture the sign correctly, this is illustrated by the spread of the distribution of model minus observed trend bias: average  $-0.05 \mu\text{g m}^{-3} \text{ yr}^{-1}$ ,  $\sigma = 0.11$ . In France all models underestimate the bias (the average difference between the modelled minus observed trends is  $0.67 \mu\text{g m}^{-3} \text{ yr}^{-1}$ ,  $\sigma = 0.09$ ). Such country-level discrepancies – consistently produced by all 6 models – are pointing towards inaccuracies in the national inventory (in which the decreasing trend of NO<sub>2</sub> emission is milder than what was actually observed). However that this is contradictory with the results of (Kononov and Beekmann, 2008) who compared satellite-derived trends and EMEP inventories and found a good agreement for France. It should be noted that they focused on a different time period (1996–2005) and also a different version of the EMEP expert emissions.

## 5.2 Ozone

The maps of ozone trends are provided in Fig. 9. When compared to emissions (Fig. 3) and NO<sub>x</sub> concentrations trends (Fig. 7) these maps should be interpreted in terms of photochemical regimes. The fact that we include results of six distinct CTMs also gives a robust insight into the model uncertainty, and the comparison of model versus observed trends can be used to infer the most reliable behaviour.

The strongest pattern is an increase of daily O<sub>3</sub> in the Southern UK, Benelux and Germany. This behaviour relates to the switch from a VOC-sensitive towards a more NO<sub>x</sub>-sensitive regime (Beekmann and Vautard, 2010; von Schneidmesser et al., 2010) because of the sharp decrease of NO<sub>x</sub> emissions not accompanied by a significant reduction of VOCs (Sillman, 1999). It is worth noting that this feature is produced by all models (even the global models, although the signal in OSLOCTM2 is milder) and is also detected in the observations (Sect. 2.3), hence demonstrating the robustness of this statement.

On the contrary, Poland seems to have switched to a VOC-sensitive regime from the beginning of the period since the increased VOC emissions (with little changes of NO<sub>x</sub> emissions, see Fig. 3) does not yield a stronger O<sub>3</sub> production.



**Fig. 9.** Same as Fig. 7 for O<sub>3</sub>.

Over France, the observed trend is very noisy for suburban and rural background stations. It is thus difficult to identify which model is doing best. Given the higher uncertainty on NO<sub>2</sub> trends discussed above (Sect. 5.1) it is thus more cautious to leave this country out of the O<sub>3</sub> trend analysis.

Over Northern Italy, the modelled geographical patterns are highly variable, as well as trends in observations. This apparently noisy behaviour is thus quite plausible in this area dominated by very stagnant meteorological conditions.

The very different behaviour in the Mediterranean region is interesting as it highlights the much larger model uncertainty in this area. However the lack of measurements prevents us from concluding on the most reliable trends.

Overall, although the scatter between modelled and observed trends (Fig. 10) is large, the models perform decently considering that ozone precursor emissions are very uncertain over relatively large areas. Considering only sites where a significant trend is observed the percentage of RB and SB stations where the sign of the trend is correctly captured is 58, 58, 66, 71, 39, and 51 % for BOLCHEM, CHIMERE, EMEP, EURAD, OSLOCTM2 and MOZART respectively.

### 5.3 Particulate matter

The modelled PM<sub>10</sub> trends obtained by the regional CTMs are displayed on Fig. 11. A widespread decrease of PM<sub>10</sub> is modelled over most of Europe, except for Spain, Portugal and France. More peculiar features include localised increases over Bulgaria and part of Portugal that can be related

to changes in the trends of total primary particulate matter in the EMEP inventory (Fig. 3). The increase in PM over the north Atlantic simulated by EMEP results probably from a meteorological change which has an impact on sea-salt emissions (as this feature also appears in the constant emission simulations, see Sect. 7).

The decreasing trend is not reflected in the composite on Panel (a) of Fig. 12 because this composite is influenced by Czech and Spanish stations where an increase is observed. Panel (b) of Fig. 12 confirms that positive trends are virtually not captured by any model (without distinction of the countries: all stations are displayed on Fig. 12b) thus questioning the role of anthropogenic emissions (Sect. 2.3). The fact that models perform well elsewhere shows that this mismatch is not due to a model shortcoming. Such trends are thus either inappropriately reported in the EMEP inventory or the observed trends are induced – in part – by classes of emissions not adequately included in the inventory (wildfires, domestic wood burning, or re-suspension of terrigenous particulate matter). Nevertheless, apart from the Czech Republic and Spain, we can conclude that the models are quite successful at capturing the trend of PM<sub>10</sub> with a fraction of significant trends with correct sign of 65, 62, 68, and 71 % respectively for BOLCHEM, CHIMERE, EMEP and EURAD.

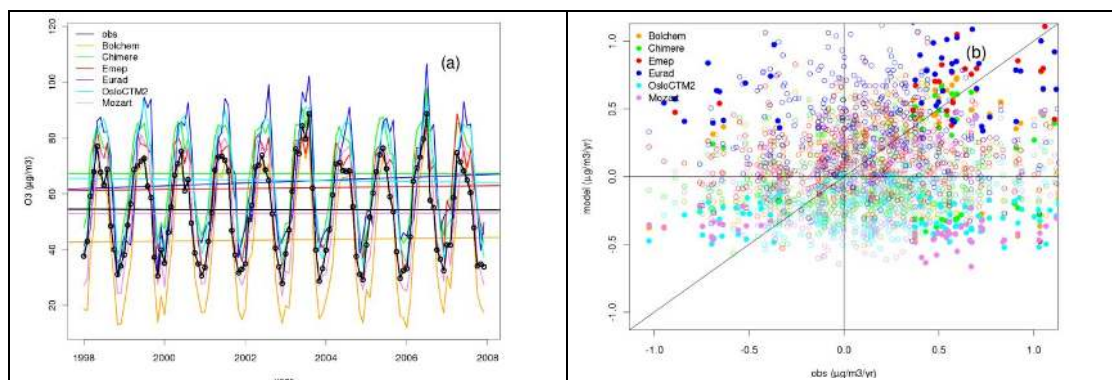


Fig. 10. Same as Fig. 8 for O<sub>3</sub>.

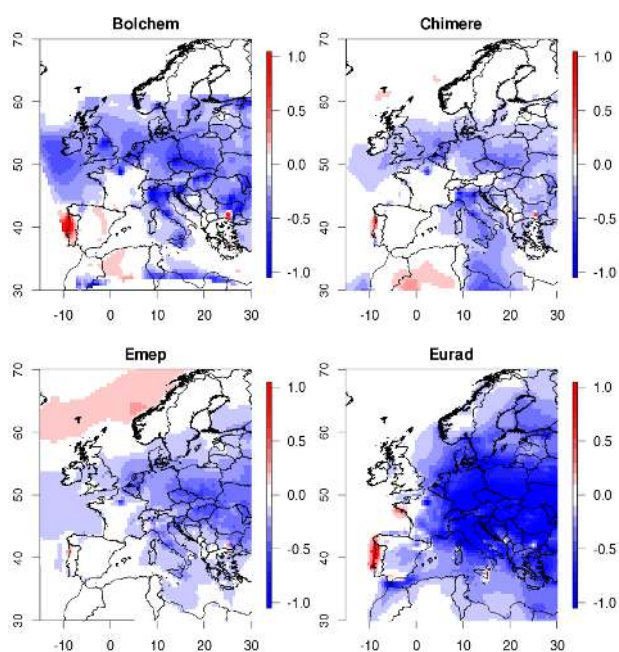


Fig. 11. Same as Fig. 7 for PM<sub>10</sub>.

## 6 Interannual variability

One of the scopes of the present study is to prepare future air quality projections, and hence to assess the skill of the models in capturing the interannual variability. To reach this goal, we discussed above their capability to reproduce past trends. In the present Section we focus on the spread around that trend, i.e. the year to year variability.

For both AIRBASE measurements (background suburban and rural stations exclusively) and the model output interpolated at the measurement sites, we compute the residual between the time series of ozone and their linear least-square fit. The standard deviation of these residuals is thus a proxy of the temporal variability in addition to the long

term changes. Note however that at this stage the seasonal variability is included in that metric of the interannual variability. Hence, in order to investigate exclusively the interannual variability we also consider the standard deviation of the residuals between the de-seasonalised time series and their linear least-square fit. At each station we obtain two standard deviations (monthly and de-seasonalised). For each model, Fig. 13 shows the quantile-quantile distribution of these two proxies, the reference (x-axis) being the distribution of the observations. The dots are equally spaced by quantiles of multiples of 10 percent. So that for example, the 5th dot represents the median of the variability in the observations (x-axis) and in the model (y-axis).

From the top panel, it appears that the month-to-month variability (once the long term trend is removed) is very well captured by all models (except at sites where the variability is very high, outside of the 10%–90% percentile ranges). The characteristics of the modelled seasonal cycles discussed in Sect. 4.2 are reflected here: EMEP and CHIMERE showing less variability in the higher quantiles.

The results on the bottom panel are not as good. A large part of the monthly O<sub>3</sub> variability is driven by the seasonal cycle. And once that cycle has been removed, the remaining variability (interannual) is more challenging to capture. Here the quantile-quantile plot shows that all models underestimate the variability compared to the observations. The median is underestimated by 28.8, 30.7, 26.3, 17.6, 53.8, and 40.3% by BOLCHEM, CHIMERE, EMEP, EURAD, OSLOCTM2, and MOZART respectively. When split by country (Table 6), it appears that this performance is very variable according to the country, similarly to the estimate of the trend in Table 5. It is therefore likely that an underestimation of the year-to-year changes of anthropogenic emissions could be partly responsible for the inability of the models to capture the observations.

Nevertheless, if such models are used for the projections of future changes, it will be essential to investigate the relevance of implementing quantile-matching corrections (Panofsky

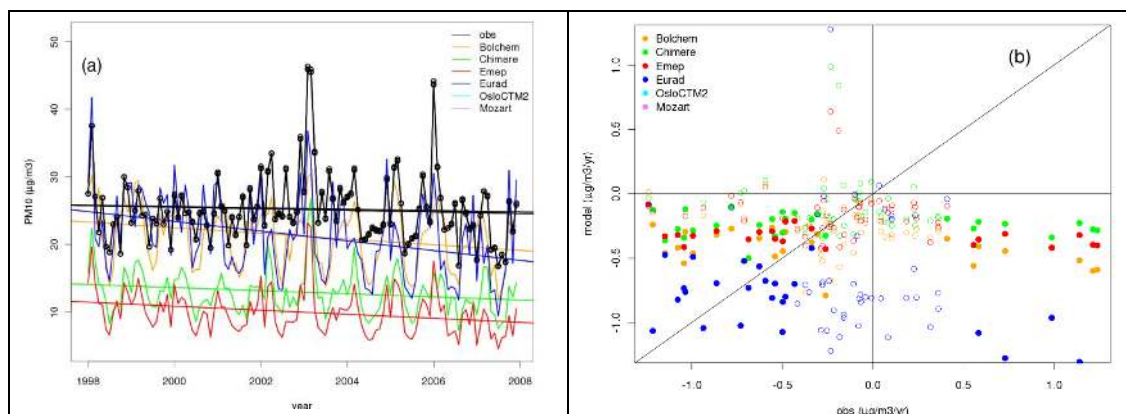


Fig. 12. Same as Fig. 8 for  $\text{PM}_{10}$ .

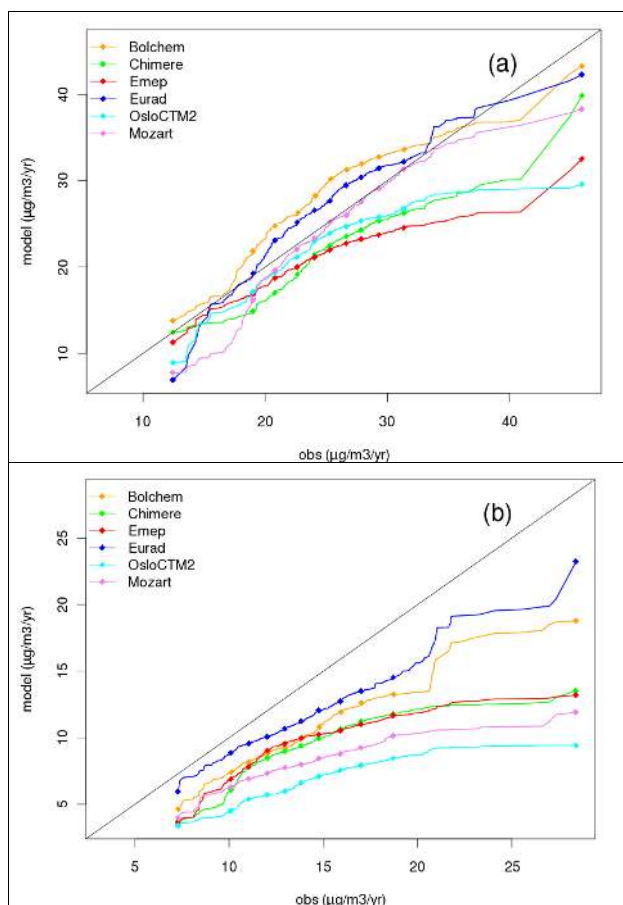


Fig. 13. Top panel: quantile-quantile plot of the standard deviation of the residuals of monthly mean  $\text{O}_3$ , once the linear trend has been removed, observations (x-axis) being used as a reference. In the bottom panel, the seasonal cycle has also been removed. The dots indicate the percentiles by multiples of 10 (0, 10th, 20th, ..., 100th).

and Brier, 1968; Li et al., 2010) to account for this underestimation of the remaining variability.

## 7 Anthropogenic emission reduction versus natural meteorological variability

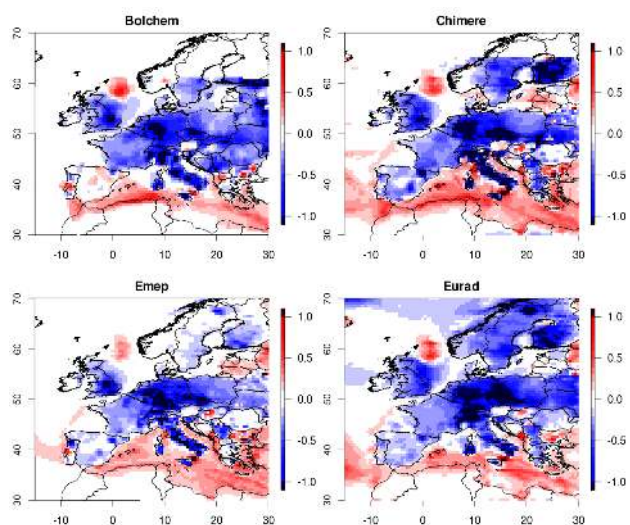
Each of the four regional chemistry transport models repeated the 10-yr simulation using constant emissions. The emissions of the last year of the decade (2007) were considered more reliable and therefore chosen for this experiment. The comparison of the trend modelled with constant (CST) and time-varying (CTRL) emissions can be used to infer the respective role of meteorological variability and anthropogenic emissions changes on the modelled concentrations of major pollutants.

We make use of a normalised relative trend (NRT): a quantitative metric defined in EEA (2009). This metric is the ratio between the trend brought about by the anthropogenic emission changes divided by the meteorological variability:

- At each grid point, the difference CTRL minus CST annual means is computed. The trend of this difference is directly related to emission changes. Assuming no trends of any factor besides anthropogenic emissions changes, this quantity would be positively correlated with the anthropogenic emissions changes.
- The meteorological variability is estimated as the standard deviation of the simulation with constant emissions. Although, as we discussed in Sect. 6, the interannual variability might be underestimated, these model simulations with constant emissions represent the only available proxy to estimate the specific impact of meteorology.

In both cases, these quantities are computed using annual values. The map of the ratio obtained for each models are displayed on Fig. 14, Fig. 15, and Fig. 16 for  $\text{NO}_2$ ,  $\text{O}_3$  and  $\text{PM}_{10}$  respectively. When the absolute value of this metric exceeds





**Fig. 14.** Trend of  $\text{NO}_2$  due to the anthropogenic emission evolution alone (linear least square fit of the difference between the reference run and a simulation with constant – 2007 – emissions), normalised by the interannual meteorological variability (standard deviation of the simulation with constant emissions).

1, the role of emission reduction on the modelled trend can be considered as more important than the interannual meteorological variability over the 1998–2007 decade.

The patterns of  $\text{NO}_2$  NRT are widely consistent with  $\text{NO}_2$  emission changes (Fig. 3) because  $\text{NO}_2$  concentrations are directly influenced by primary emissions. The areas where all four models consistently identify a consistent decrease of NRT higher than unity are: the greater London area (UK), the Ruhr (Germany), Benelux, the Czech Republic and Italy. At this stage it is important to recall that the present discussion involves exclusively models, it is thus essential to go back to our assessment of the validity of the modelled trends against observations in Sect. 5.1. In Table 5, we provided a quantification of model performances in reproducing the trends on a country-level basis. We found that all models were quite successful in the Czech Republic, Germany, and the Netherlands and performing less well in France. Unfortunately, most other European countries did not offer an appropriate monitoring network to be included in the comparison. Nevertheless, we can be quite confident in the behaviour of the models in Germany, Benelux and the Czech Republic and thus conclude that the significant NRT identified there is a robust finding.

We saw before (Sect. 5.2) that the observed and modelled trends of  $\text{O}_3$  in Europe during the 1998–2007 period are slightly positive over European megacities where the confidence on  $\text{NO}_2$  trends is higher (UK, Benelux, Germany). These positive trends are however small and usually not significant in the CTRL simulation, and this is even more true for the CST simulation. Nevertheless, and interestingly

enough, EMEP and EURAD seem to capture a positive  $\text{O}_3$  trend in the CST experiment, reflecting either a direct impact of temperature changes over that period, or a reinforced role of biogenic emissions in these models (indirectly related to temperature changes). The consequence is a modulation of the widespread  $\text{O}_3$  increase modelled by EMEP and EURAD for the CTRL simulation (Fig. 9) so that the patterns of NRT are less pronounced for these models on Fig. 15.

In the CST simulation,  $\text{PM}_{10}$  concentrations exhibit very small trends except in France where a slight positive trend is captured by all models and over the North Atlantic where EMEP shows an increase of sea-salt (already mentioned in Sect. 5.3). The NRT patterns on Fig. 16 are thus very close to the modelled trends on Fig. 11, except in France where the decrease is stronger, and over the North Atlantic where the positive trend in EMEP results vanishes. Perhaps the most surprising finding is a relatively similar trend for  $\text{PM}_{10}$  in EURAD results in the CST and CTRL simulations, yielding milder patterns on Fig. 16 compared to Fig. 11 (see e.g. the absence of a negative trend North of Morocco and Algeria on Fig. 16). Otherwise most models show that the order of magnitude of the decrease of  $\text{PM}_{10}$  due to anthropogenic emissions management reaches or exceeds the natural variability over most of Europe.

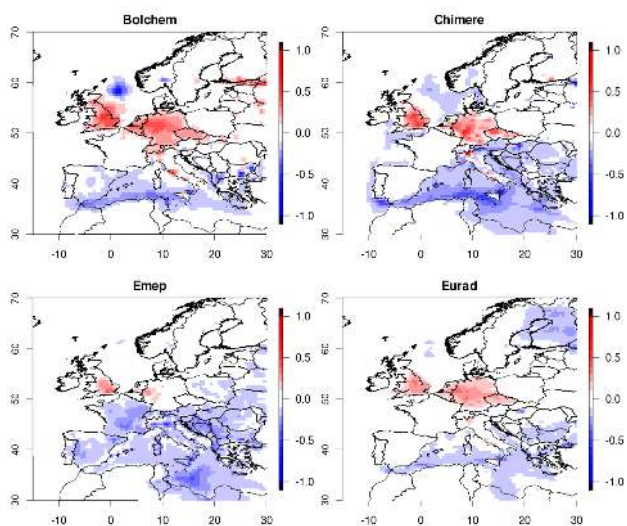
## 8 Conclusions

This paper contributes to the assessment of the capacity of state-of-the-art regional and global chemistry transport models (RCTM and GCTM) to capture the interannual variability of air pollution in major anthropogenic emission hotspots in Europe. A special attention is given to the cluster of large European cities in Northern France, Southern United Kingdom, Benelux and Western Germany. The purpose of the study is to investigate past modelled trends in order to demonstrate the potential and limitations of existing models for assessing the impact of future air pollution control strategies. To address these points a coordinated numerical experiment covering a period of 10 yr and involving six modelling groups was conducted. It is the first time that the air quality modelling community performs a modelling exercise covering such a time scale.

A model evaluation was performed to understand the respective strengths and weaknesses of the models. Although the scope of the study was focused on trends and interannual variability, it was also the opportunity to propose a multi-annual model evaluation. The most striking result is the consistency of model performances between regional and global chemistry-transport models induced by the scope of the study (focused on daily mean scores rather than on hourly or peak values, Valari and Menut, 2008) and the use of a common emission inventory. Another interesting conclusion in terms of scale errors regards the dissimilarity of seasonal cycles amongst RCTMs, given that they rely on identical seasonal

**Table 6.** Percentage of underestimation in the modelled median interannual variability (average – avg - and standard deviation –  $\sigma$ ) at all stations of a given country and across all models. The interannual variability is estimated as the distribution of residuals of the de-seasonalised residuals of the linear fit of monthly time series. Only countries where at least 5 stations are available are shown (number of selected stations – nst – provided on the last row).

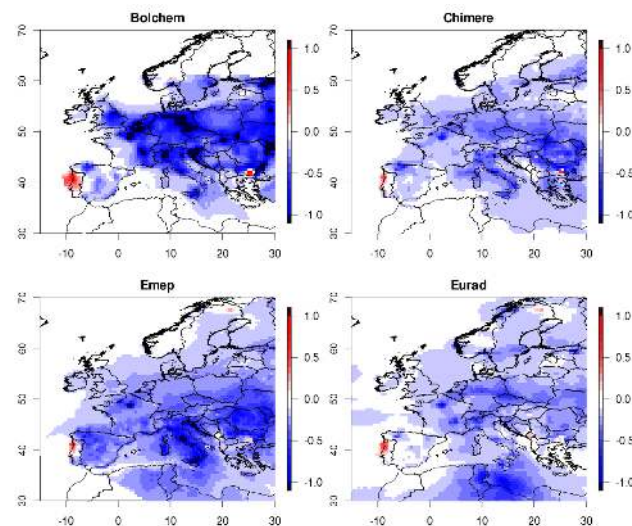
	AT	BE	CH	CZ	DE	ES	
Avg ( $\sigma$ )	34.1 (14.1)	15.8 (10.1)	20.7 (10.95)	23.4 (13.65)	34.3 (16.0)	30.4 (16.0)	
nst	45	15	7	13	62	17	
	FI	FR	GB	IT	NL	NO	PL
Avg ( $\sigma$ )	17.6 (22.6)	42.4 (12.8)	49.3 (13.6)	48.98 (14.3)	47.7 (13.8)	34.4 (11.5)	31.9 (12.8)
nst	5	43	14	8	11	5	7



**Fig. 15.** Same as Fig. 14 for  $O_3$ .

profiles in the emissions inventories. We also found that the models exhibited various degree of photochemical activity, hence leading to quite variable  $O_3$  modelling skills. The performances of the RCTM to model aerosols could be divided in two broad types of behaviour: small bias in total  $PM_{10}$  due to an overestimation of ammonium nitrate, or a strong negative  $PM_{10}$  bias. We conclude that the ensemble of models implemented here covers a wide envelope of behaviours. This leads to a higher confidence in the representativeness of this set of models, and shows that they reflect well the modelling capacities of the atmospheric chemistry modelling community.

The CTMs proved to be quite successful in capturing the decreasing trend of primary pollutants, especially in the emission hotspot areas around the Benelux region. Note that we focused here exclusively on background stations and on aggregated metrics such as daily and monthly means. The results might have been substantially different at urban or traffic monitoring sites or when investigating peak values, but such proxies were considered irrelevant in a multi-model



**Fig. 16.** Same as Fig. 14 for  $PM_{10}$ .

study involving global models. Downwards trends of  $NO_2$  were successfully captured at 73 % of the stations on average for all models. Important mismatches were systematically modelled (e.g. France) pointing towards caveats in the emissions inventory.  $PM_{10}$  trends were also quite well captured, although the validation could not be as quantitative because of the relative lack of long term measurements.  $O_3$  trends turned out to be much more challenging to reproduce, partly because the trends are small in magnitude during the period under consideration. Nevertheless, the models capture the trend in the majority of stations and we could discuss  $O_3$  evolution in terms of photochemical regimes. As suggested elsewhere (Beekmann and Vautard, 2010), it is found that the  $NO_x$ -reduction policy yields moderate increases in  $O_3$  over the Benelux hotspot of emissions. Given the  $NO_x$ -saturated photochemical regimes dominating there, the titration of  $NO_x$  on  $O_3$  dominates and more ambitious  $NO_x$  reduction measures could be considered in future policies.

We also devoted a special focus on the modelled temporal variability (apart from the linear trend mentioned above). It appears that the variability of the residual between the monthly means and the linear trend is well reproduced. However, this variability is heavily influenced by the seasonal cycle. Hence the capacity of the models to capture this variability does not reflect their performance in reproducing year-to-year changes. Once the seasonal cycle is removed, the inter-annual variability is less well modelled. This result clearly shows that caution needs to be taken when using these models to assess future air quality variability.

In a last part, the respective role of meteorology and anthropogenic emission changes is addressed by comparing model simulations with constant emissions. We find out that the magnitude of the anthropogenic NO<sub>2</sub> decrease exceeds the natural variability over most of Europe. This demonstrates that emission reduction strategies enforced over the past decade led to the reduction of NO<sub>2</sub> background levels. Consequently, this result suggests that ambitious environmental policies have a beneficial impact on NO<sub>2</sub> ambient concentrations, even if this effect was not as large as expected when the emission control strategies were decided (partly because of an increased proportion of diesel engines and a subsequent change in the NO/NO<sub>2</sub> ratio).

To summarize, the trend assessment conducted here shows that reductions of anthropogenic emissions of nitrogen oxides and particulate matter effectively lead to reductions of atmospheric loading of primary constituents. However, the insufficient efforts on volatile organic compounds in areas exposed to a VOC-sensitive photochemical regime associated to a decrease of NO<sub>x</sub> titration of O<sub>3</sub> in NO<sub>x</sub> saturated areas lead to localised increases of ozone, especially sensitive over the most urbanised areas. The model assessment proved that the models were efficient at capturing the trend of primary species but the more limited magnitude of ozone changes was more challenging to reproduce. Nevertheless we conclude that these models capture most of the important features to justify their implementation for future projections of air quality provided that enough attention is given to their underestimation of interannual variability.

*Acknowledgements.* The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 212095 (CityZen). Anthropogenic emissions were provided by EMEP and complemented for global models in the framework of the MACC FP7 Project (grant agreement no. 218793). Surface observations were obtained through the AIRBASE (EEA) and EBAS (NILU) repositories. Large scale reanalyses used to drive the mesoscale meteorological models were provided by ECMWF and NCEP/NCAR.

Edited by: L. Molina

## References

- Ackermann, I. J., Hass, H., Memmesheimer, M., Ebel, A., Binkowski, F. S., and Shankar, U.: Modal aerosol dynamics model for Europe: development and first applications, *Atmos. Environ.*, 32, 2981–2999, 1998.
- Beekmann, M. and Vautard, R.: A modelling study of photochemical regimes over Europe: robustness and variability, *Atmos. Chem. Phys.*, 10, 10067–10084, doi:10.5194/acp-10-10067-2010, 2010.
- Berglen, T. F., Berntsen, T. K., Isaksen, I. S. A., and Sundet, J. K.: A global model of the coupled sulfur/oxidant chemistry in the troposphere: The sulfur cycle, *J. Geophys. Res.*, 109, D19310, doi:10.1029/2003JD003948, 2004.
- Berntsen, T. K. and Isaksen, I. S. A.: A global three-dimensional chemical transport model for the troposphere 1. Model description and CO and ozone results, *J. Geophys. Res.*, 102, 21239–21280, 1997.
- Bessagnet, B., Menut, L., Curci, G., Hodzic, A., Guillaume, B., Liousse, C., Moukhtar, S., Pun, B., Seigneur, C., and Schulz, M.: Regional modeling of carbonaceous aerosols over Europe – focus on secondary organic aerosols, *J. Atmos. Chem.*, 61, 175–202, 2008.
- Binkowski, F. S. and Roselle, S. J.: Models-3 Community Multiscale Air Quality (CMAQ) model aerosol component 1. Model description, *J. Geophys. Res.*, 108, 4183, doi:10.1029/2001JD001409, 2003.
- Braniš, M.: Long Term Trends in Concentration of Major Pollutants (SO<sub>2</sub>, CO, NO, NO<sub>2</sub>, O<sub>3</sub> and PM<sub>10</sub>) in Prague – Czech Republic (Analysis of Data Between 1992 and 2005), *Water Air Soil Pollut. Focus*, 8, 49–60, 2008.
- Buzzi, A., Fantini, M., Malguzzi, P., and Nerozzi, F.: Validation of a limited area model in cases of mediterranean cyclogenesis: Surface fields and precipitation scores, *Meteorol. Atmos. Phys.*, 53, 137–153, 1994.
- CAFE: Communication from the Commission to the Council and the European parliament : Thematic Strategy on air pollution SEC(2005)1132 et SEC(2005)1133 European Commission, Brussels, 2005.
- Carslaw, D., Beevers, S., Westmoreland, E., Williams, M., Tate, J., Murrells, T., Stedman, J., Li, Y., Grice, S., Kent, A., and Tsagatakis, I.: Trends in NO<sub>x</sub> and NO<sub>2</sub> emissions and ambient measurements in the UK, Defra, London, 2011.
- Carter, W. P. L.: A detailed mechanism for the gas-phase atmospheric reactions of organic compounds, *Atmos. Environ.*, Part A. General Topics, 24, 481–518, 1990.
- Derwent, R. G., Jenkin, M. E., Saunders, S. M., Pilling, M. J., Simmonds, P. G., Passant, N. R., Dollard, G. J., Dumitrescu, P., and Kent, A.: Photochemical ozone formation in north west Europe and its control, *Atmos. Environ.*, 37, 1983–1991, 2003.
- EC: Directive 2001/42/EC of the European Parliament and of the Council of 27 June 2001 on the assessment of the effects of certain plans and programmes on the environment, 2001.
- EEA: Assessment of ground-level ozone in EEA member countries, with a focus on long-term trends, European Environment Agency, Copenhagen, 56, 2009.
- Emmons, L. K., Walters, S., Hess, P. G., Lamarque, J.-F., Pfister, G. G., Fillmore, D., Granier, C., Guenther, A., Kinnison, D., Laepple, T., Orlando, J., Tie, X., Tyndall, G., Wiedinmyer, C., Baughcum, S. L., and Kloster, S.: Description and evaluation of

- the Model for Ozone and Related chemical Tracers, version 4 (MOZART-4), *Geosci. Model Dev.*, 3, 43–67, doi:10.5194/gmd-3-43-2010, 2010.
- Endresen, Ø., Sørsgård, E., Behrens, H. L., Brett, P. O., and Isaksen, I. S. A.: A historical reconstruction of ships' fuel consumption and emissions, *J. Geophys. Res.*, 112, D12301, doi:10.1029/2006JD007630, 2007.
- Eyring, V., Isaksen, I. S. A., Berntsen, T., Collins, W. J., Corbett, J. J., Endresen, O., Grainger, R. G., Moldanova, J., Schlager, H., and Stevenson, D. S.: Transport impacts on atmosphere and climate: Shipping, *Atmos. Environ.*, 44, 4735–4771, 2010.
- Favez, O., Cachier, H., Sciare, J., and Le Moulllec, Y.: Characterization and contribution to PM<sub>2.5</sub> of semi-volatile aerosols in Paris (France), *Atmos. Environ.*, 41, 7969–7976, 2007.
- Geiger, H., Barnes, I., Bejan, I., Benter, T., and Spittler, M.: The tropospheric degradation of isoprene: an updated module for the regional atmospheric chemistry mechanism, *Atmos. Environ.*, 37, 1503–1519, 2003.
- Giorgi, F., Jones, C., and Asrar, G. R.: Addressing climate information needs at the regional level: the CORDEX framework, *WMO Bulletin*, 58, 175–183, 2009.
- Graf, H.-F., Feichter, J., and Langmann, B.: Volcanic sulfur emissions: Estimates of source strength and its contribution to the global sulfate distribution, *J. Geophys. Res.*, 102, 10727–10738, 1997.
- Granier, C., Lamarque, J. F., Mieville, A., Muller, J. F., Olivier, J., Orlando, J., Peters, J., Petron, G., Tyndall, G., and Wallens, S.: POET: A Database of Surface Emissions of Ozone Precursors, Paris, 2005.
- Granier, C., Bessagnet, B., Bond, T., D'Angiola, A., Denier van der Gon, H., Frost, G., Heil, A., Kaiser, J., Kinne, S., Klimont, Z., Kloster, S., Lamarque, J.-F. o., Liousse, C., Masui, T., Meleux, F., Mieville, A., Ohara, T., Raut, J.-C., Riahi, K., Schultz, M., Smith, S., Thompson, A., van Aardenne, J., van der Werf, G., and van Vuuren, D.: Evolution of anthropogenic and biomass burning emissions of air pollutants at global and regional scales during the 1980–2010 period, *Climatic Change*, 109, 163–190, 2011.
- Grell, G. A., Duddhia, J., and Stauffer, D.: A description of the fifth-generation Penn State/Ncar meso-scale model (MM5), National Center of Atmospheric Research, Boulder, 1994.
- Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), *Atmos. Chem. Phys.*, 6, 3181–3210, doi:10.5194/acp-6-3181-2006, 2006.
- Harrison, R. M., Stedman, J., and Derwent, D.: New Directions: Why are PM<sub>10</sub> concentrations in Europe not falling?, *Atmos. Environ.*, 42, 603–606, 2008.
- Hauglustaine, D. A., Hourdin, F., Jourdain, L., Filiberti, M. A., Walters, S., Lamarque, J. F., and Holland, E. A.: Interactive chemistry in the Laboratoire de Météorologie Dynamique general circulation model: Description and background tropospheric chemistry evaluation, *J. Geophys. Res.*, 109, D04314, doi:10.1029/2003JD003957, 2004.
- Hesstvedt, E., Hov, Ö., and Isaksen, I. S. A.: Quasi-steady-state approximations in air pollution modeling: Comparison of two numerical schemes for oxidant prediction, *International Journal of Chemical Kinetics*, 10, 971–994, 1978.
- Hipel, K. W. and McLeod, A. I.: *Time Series Modelling of Water Resources and Environmental Systems*, Elsevier, Amsterdam, 2005.
- Hollander, M. and Wolfe, D. A.: *Nonparametric Statistical Methods*, in 2 ed., John Wiley & Sons, New York, 27–33, 68–75, 1999.
- Holtzlag, A. A. M., De Bruijn, E. I. F., and Pan, H. L.: A High Resolution Air Mass Transformation Model for Short-Range Weather Forecasting, 118, 1561–1575, doi:10.1175/1520-0493(1990)118<1561:AHRAMT>2.0.CO;2, 1990.
- Honoré, C., Menut, L., Bessagnet, B., Meleux, F., Rouïl, L., Vautard, R., Poisson, N., Peuch, V. H., and Carlos Borrego and Eberhard, R.: Chapter 3.4 PREV'AIR: A platform for air quality monitoring and forecasting, in: *Developments in Environmental Sciences*, Elsevier, Burlington, 293–300, 2007.
- Isaksen, I. S. A., Zerefos, C., Kourtidis, K., Meleti, C., Dalsøren, S. B., Sundet, J. K., Grini, A., Zanis, P., and Balis, D.: Tropospheric ozone changes at unpolluted and semipolluted regions induced by stratospheric ozone changes, *J. Geophys. Res.*, 110, D02302, doi:10.1029/2004JD004618, 2005.
- Jakobs, H., Tilmes, S., Heidegger, A., Nester, K., and Smiatek, G.: Short-Term Ozone Forecasting with a Network Model System during Summer 1999, *J. Atmos. Chem.*, 42, 23–40, 2002.
- Jonson, J. E., Simpson, D., Fagerli, H., and Solberg, S.: Can we explain the trends in European ozone levels?, *Atmos. Chem. Phys.*, 6, 51–66, doi:10.5194/acp-6-51-2006, 2006.
- Karl, M., Dorn, H. P., Holland, F., Koppmann, R., Poppe, D., Rupp, L., Schaub, A., and Wahner, A.: Product study of the reaction of OH radicals with isoprene in the atmosphere simulation chamber SAPHIR, *J. Atmos. Chem.*, 55, 167–187, 2006.
- Kendall, M. G.: *Rank Auto Correlation Methods*, 4 ed., Griffin, Oxford, 1976.
- Konovalov, I. B. and Beekmann, M.: On the Use of Air Quality Monitoring Networks for the Evaluation of Nitrogen Oxide Emission Inventories, *The Open Atmospheric Science Journal*, 2, 232–248, 2008.
- Konovalov, I. B., Beekmann, M., Richter, A., and Burrows, J. P.: Inverse modelling of the spatial distribution of NO<sub>x</sub> emissions on a continental scale using satellite data, *Atmos. Chem. Phys.*, 6, 1747–1770, doi:10.5194/acp-6-1747-2006, 2006.
- Konovalov, I. B., Beekmann, M., Richter, A., Burrows, J. P., and Hilboll, A.: Multi-annual changes of NO<sub>x</sub> emissions in megacity regions: nonlinear trend analysis of satellite measurement based estimates, *Atmos. Chem. Phys.*, 10, 8481–8498, doi:10.5194/acp-10-8481-2010, 2010.
- Lamb, B., Gay, D., Westberg, H., and Pierce, T.: A biogenic hydrocarbon emission inventory for the U.S.A. using a simple forest canopy model, *Atmos. Environ. Part A*, 27, 1673–1690, 1993.
- Lattuati, M.: Impact des émissions européennes sur le bilan d'ozone troposphérique à l'interface de l'Europe et de l'Atlantique nord: apport de la modélisation lagrangienne et des mesures en altitude, Univ. Pierre et Marie Curie, Paris, 1997.
- Li, H., Sheffield, J., and Wood, E. F.: Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching, *J. Geophys. Res.*, 115, D10101, doi:10.1029/2009JD012882, 2010.
- Logan, J. A.: An analysis of ozonesonde data for the troposphere: Recommendations for testing 3-D models and development of

- a gridded climatology for tropospheric ozone, *J. Geophys. Res.*, 104, 16115–16149, 1999.
- Logan, J. A., Megretskaia, I. A., Miller, A. J., Tiao, G. C., Choi, D., Zhang, L., Stolarski, R. S., Labow, G. J., Hollandsworth, S. M., Bodeker, G. E., Claude, H., De Muer, D., Kerr, J. B., Tarasick, D. W., Oltmans, S. J., Johnson, B., Schmidlin, F., Staehelin, J., Viatte, P., and Uchino, O.: Trends in the vertical distribution of ozone: A comparison of two analyses of ozonesonde data, *J. Geophys. Res.*, 104, 26373–26399, 1999.
- Løvblad, G., Tarrasón, L., Tørseth, K., and Dutchak, S.: EMEP Assessment, Part I, European Perspective, met.no, Oslo, 47–62, 2004.
- Memmesheimer, M., Friese, E., Ebel, A., Jakobs, H. J., Feldmann, H., Kessler, C., and Piekorz, G.: Long-term simulations of particulate matter in Europe on different scales using sequential nesting of a regional model, 1–2, Inderscience Enterprises, Geneva, SUISSE, 25 pp., 2004.
- Memmesheimer, M., Wurzler, S., Friese, E., Jakobs, H. J., Feldmann, H., Ebel, A., Kessler, C., Geiger, J., Hartmann, U., Brandt, A., Pfeffer, U., Dorn, H. P., and Carlos Borrego and Eberhard, R.: Chapter 2.8 Long-term simulations of photo-oxidants and particulate matter over Europe with emphasis on North Rhine-Westphalia, in: *Developments in Environmental Sciences*, Elsevier, 158–167, 2007.
- Metzger, S., Dentener, F., Pandis, S., and Lelieveld, J.: Gas/aerosol partitioning: 1. A computationally efficient model, *J. Geophys. Res.*, 107, 4312, doi:10.1029/2001JD001102, 2002.
- Mircea, M., D'Isidoro, M., Maurizi, A., Vitali, L., Monforti, F., Zanini, G., and Tampieri, F.: A comprehensive performance evaluation of the air quality model BOLCHEM to reproduce the ozone concentrations over Italy, *Atmos. Environ.*, 42, 1169–1185, 2008.
- Monks, P. S., Granier, C., Fuzzi, S., Stohl, A., Williams, M. L., Akimoto, H., Amann, M., Baklanov, A., Baltensperger, U., Bey, I., Blake, N., Blake, R. S., Carslaw, K., Cooper, O. R., Dentener, F., Fowler, D., Fragkou, E., Frost, G. J., Generoso, S., Ginoux, P., Grewe, V., Guenther, A., Hansson, H. C., Henne, S., Hjorth, J., Hofzumahaus, A., Huntrieser, H., Isaksen, I. S. A., Jenkin, M. E., Kaiser, J., Kanakidou, M., Klimont, Z., Kulmala, M., Laj, P., Lawrence, M. G., Lee, J. D., Lioussé, C., Maione, M., McFiggans, G., Metzger, A., Mieville, A., Moussiopoulos, N., Orlando, J. J., O'Dowd, C. D., Palmer, P. I., Parrish, D. D., Petzold, A., Platt, U., Pöschl, U., Prévôt, A. S. H., Reeves, C. E., Reimann, S., Rudich, Y., Sellegri, K., Steinbrecher, R., Simpson, D., ten Brink, H., Theloke, J., van der Werf, G. R., Vautard, R., Vestreng, V., Vlachokostas, C., and von Glasow, R.: Atmospheric composition change – global and regional air quality, *Atmos. Environ.*, 43, 5268–5350, 2009.
- Myhre, G., Grini, A., and Metzger, S.: Modelling of nitrate and ammonium-containing aerosols in presence of sea salt, *Atmos. Chem. Phys.*, 6, 4809–4821, doi:10.5194/acp-6-4809-2006, 2006.
- Nenes, A., Pilinis, C., and Pandis, S. N.: ISORROPIA: a new thermodynamic equilibrium model for multiphase multicomponent marine aerosols, *Aquat. Geochem.*, 4, 123–152, 1998.
- Olivier, J., Peters, J., Granier, C., Petron, G., Müller, J. F., and Wallens, S.: Present and future surface emissions of atmospheric compounds, POET Rep. 2/ Eur. Union, Brussels, 2003.
- Ordóñez, C., Mathis, H., Furger, M., Henne, S., Hglin, C., Staehelin, J., and Prévôt, A. S. H.: Changes of daily surface ozone maxima in Switzerland in all seasons from 1992 to 2002 and discussion of summer 2003, *Atmos. Chem. Phys.*, 5, 1187–1203, doi:10.5194/acp-5-1187-2005, 2005.
- Panofsky, H. A. and Brier, G. W.: *Some Applications of Statistics to Meteorology*, The Pennsylvania State University, 224 pp., 1968.
- Prather, M. J.: *Numerical advection by conservation of second-order moments*, 6, American Geophysical Union, Washington, DC, ETATS-UNIS, 1986.
- Schell, B., Ackermann, I. J., Hass, H., Binkowski, F. S., and Ebel, A.: Modeling the formation of secondary organic aerosol within a comprehensive air quality model system, *J. Geophys. Res.*, 106, 28275–28293, 2001.
- Schultz, M. G., Heil, A., Hoelzemann, J. J., Spessa, A., Thonicke, K., Goldammer, J. G., Held, A. C., Pereira, J. M. C., and van het Bolscher, M.: Global wildland fire emissions from 1960 to 2000, *Global Biogeochem. Cy.*, 22, GB2002, doi:10.1029/2007GB003031, 2008.
- Sillman, S.: The relation between ozone, NO<sub>x</sub> and hydrocarbons in urban and polluted rural environments, *Atmos. Environ.*, 33, 1821–1845, 1999.
- Simpson, D.: Biogenic emissions in Europe 2. Implications for ozone control strategies, *J. Geophys. Res.*, 100, 22891–22906, 1995.
- Simpson, D., Andersson-Skold, Y., and Jenkin, M. E.: Updating the chemical scheme for the EMEP MSC-Woxidant model : current status, The Norwegian Meteorological Institute, EMEP, Oslo, 1993.
- Simpson, D., Fagerli, H., Jonson, J., Tsyro, S., Wind, P., and Tuovinen, J.-P.: *The EMEP Unified Eulerian Model. Model Description*, The Norwegian Meteorological Institute, EMEP, Oslo, 2003.
- Simpson, D., Benedictow, A., Berge, H., Fagerli, H., Gauss, M., Jonson, J. E., Nyiri, A., Semeena, V., Steensen, B. M., Tsyro, S., Valdebenito, A., and Wind, P.: *The EMEP MSC-W Chemical Transport Model I: Model Description*, in preparation, 2011.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X. Y., Wang, W., and Powers, J. G.: A Description of the Advanced Research WRF Version 3, NCAR, 2008.
- Søvde, O. A., Gauss, M., Smyshlyaev, S. P., and Isaksen, I. S. A.: Evaluation of the chemical transport model Oslo CTM2 with focus on arctic winter ozone depletion, *J. Geophys. Res.*, 113, D09304, doi:10.1029/2007JD009240, 2008.
- Spiro, P. A., Jacob, D. J., and Logan, J. A.: Global Inventory of Sulfur Emissions With 1° × 1° Resolution, *J. Geophys. Res.*, 97, 6023–6036, 1992.
- Stockwell, W. R., Kirchner, F., Kuhn, M., and Seefeld, S.: A new mechanism for regional atmospheric chemistry modeling, *J. Geophys. Res.*, 102, 25847–25879, 1997.
- Symeonidis, P., Poupkou, A., Gkantou, A., Melas, D., Devrim Yay, O., Pouspourika, E., and Balis, D.: Development of a computational system for estimating biogenic NMVOCs emissions based on GIS technology, *Atmos. Environ.*, 42, 1777–1789, 2008.
- Thouret, V., Marenco, A., Logan, J. A., Nédélec, P., and Grouhel, C.: Comparisons of ozone measurements from the MOZAIC airborne program and the ozone sounding network at eight locations, *J. Geophys. Res.*, 103, 25695–25720, 1998.
- Tiedtke, M.: A comprehensive mass flux scheme for cumulus pa-

- parameterization in large-scale models, 11, 1989.
- Undén, P., Rontu, L., H., J., Lynch, P., and Calvo, J.: HIRLAM-5 Scientific Documentation, HIRLAM-5 project, SMHI, Norrköping, Sweden, 2003.
- UNECE: The 1999 Gothenburg Protocol to Abate Acidification, Eutrophication and Ground-level Ozone UNECE, Gothenburg, Report, 1999.
- Valari M. and Menut, L.: Does increase in air quality models resolution bring surface ozone concentrations closer to reality?, *J. Atmos. Ocean. Tech.*, 25, 1955–1968, doi:10.1175/2008JTECHA1123.1, 2008.
- van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Kasibhatla, P. S., and Arellano Jr., A. F.: Interannual variability in global biomass burning emissions from 1997 to 2004, *Atmos. Chem. Phys.*, 6, 3423–3441, doi:10.5194/acp-6-3423-2006, 2006.
- van Loon, M., Vautard, R., Schaap, M., Bergström, R., Bessagnet, B., Brandt, J., Builtjes, P. J. H., Christensen, J. H., Cuvelier, C., Graff, A., Jonson, J. E., Krol, M., Langner, J., Roberts, P., Rouil, L., Stern, R., Tarrasón, L., Thunis, P., Vignati, E., White, L., and Wind, P.: Evaluation of long-term ozone simulations from seven regional air quality models and their ensemble, *Atmos. Environ.*, 41, 2083–2097, 2007.
- Vautard, R., Szopa, S., Beekmann, M., Menut, L., Hauglustaine, D. A., Rouil, L., and Roemer, M.: Are decadal anthropogenic emission reductions in Europe consistent with surface ozone observations?, *Geophys. Res. Lett.*, 33, L13810, doi:10.1029/2006GL026080, 2006.
- Vautard, R., Schaap, M., Bergström, R., Bessagnet, B., Brandt, J., Builtjes, P. J. H., Christensen, J. H., Cuvelier, C., Foltescu, V., Graff, A., Kerschbaumer, A., Krol, M., Roberts, P., Rouil, L., Stern, R., Tarrason, L., Thunis, P., Vignati, E., and Wind, P.: Skill and uncertainty of a regional air quality model ensemble, *Atmos. Environ.*, 43, 4822–4832, 2009.
- Vestreng, V., Breivik, K., Adams, M., Wagener, A., Goodwin, J., Rozovskkaya, O., and Pacyna, J. M.: Inventory Review 2005, Emission Data reported to LRTAP Convention and NEC Directive, Initial review of HMs and POPs, MSC-W, 2005.
- Vestreng, V., Ntziachristos, L., Semb, A., Reis, S., Isaksen, I. S. A., and Tarrasón, L.: Evolution of NO<sub>x</sub> emissions in Europe with focus on road transport control measures, *Atmos. Chem. Phys.*, 9, 1503–1520, doi:10.5194/acp-9-1503-2009, 2009.
- Vignati, E., Wilson, J., and Stier, P.: M7: An efficient size-resolved aerosol microphysics module for large-scale aerosol transport models, *J. Geophys. Res.*, 109, D22202, doi:10.1029/2003JD004485, 2004.
- von Schneidemesser, E., Monks, P. S., and Plass-Duelmer, C.: Global comparison of VOC and CO observations in urban areas, *Atmos. Environ.*, 44, 5053–5064, 2010.
- Wesely, M. L.: Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models, *Atmos. Environ.*, 23, 1293–1304, 1989.