



CityZen

megaCITY - Zoom for the Environment

Collaborative Project

7th Framework Programme for Research and Technological Development

Cooperation, Theme 6:

Environment (including Climate Change)

Grant Agreement No.: 212095

Deliverable D1.2.3, type O (article)

Publication on scale errors in global model simulations

based on comparison with regional/Lagrangian model results

and also observations

Due date of deliverable: project month 23

Actual submission date: project month 34

Start date of project: 1 September 2008

Duration: 36 months

Name of lead beneficiary for this deliverable:

INERIS

Scientist(s) responsible for this deliverable:

A. Colette

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Publication on scale errors in global model simulations based on comparison with regional/Lagrangian model results and also observations

This deliverable is constituted of two distinct studies that contribute to the documentation of model performances in terms of megacity air quality modelling across a wide range of scales. One of them has recently been submitted to the ACP Special Issue on megacities.

1. Comparison of global and regional models to observations

In Annex A we provide a copy of a paper by Colette et al. recently submitted for publication in the Megacity Special Issue of Atmospheric Chemistry and Physics (see the abstract below). This paper is entitled “Air quality trends in Europe over the past decade: a first multi-model assessment”. It presents the result of a modelling exercise that involved 6 partners of the project deploying the regional Chemistry Transport Models (CTM) Bolchem, Chimere, Emep and Eurad and the global models Mozart and OsloCTM2. A thorough investigation of model performance in capturing AQ trends in the European Hotspot of Emission is presented. Of more specific relevance to the present deliverable, the paper includes a section (part 4.) devoted to the comparison of Regional and Global CTMs to surface air quality observations. As such it addresses the following objective of D1.2.3: “scale errors in global model simulations based on comparison with regional model results and also observations”.

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

A. Colette, C. Granier, Ø. Hodnebrog, H. Jakobs, A. Maurizi, A. Nyiri, B. Bessagnet, A. D'Angiola, M. D'Isidoro, M. Gauss, F. Meleux, M. Memmesheimer, A. Mieville, L. Rouil, F. Russo, S. Solberg, F. Stordal, F. Tampieri

We discuss the capability of current state-of-the-art chemistry and transport models to reproduce air quality trends and inter annual 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 (NO₂, O₃ and PM₁₀). 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 O₃ is more challenging. Apart from the long term trend, the modelled monthly variability 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 years, hence supporting further emission reductions strategies.

2. Comparison of global and Lagrangian models to observations

In Annex B, a report by Bouarar et al. discusses the capability of the global CTM Mozart to reproduce the pollution plume of the hotspot of anthropogenic emission constituted by the cluster of megacities in the North-Eastern United States. The availability of an outstanding observational dataset obtained during the ICAART 2004 campaign makes this region one of the best documented megacity air pollution plume worldwide, hence supporting the choice of this case study. Using a significant number of airborne transects, Bouarar et al. can thus document the global model performance in terms megacity of pollution plume export. Building upon previous Lagrangian modelling study it also covers the issue of Eulerian versus Lagrangian modelling approaches.

Report on assessment of errors in global model simulations based on comparison of an Eulerian model with observations

I. Bouarar, K. Law, C. Granier, L.K. Emmons

Results from the chemistry transport model MOZART-4 have been compared to in-situ measurements from the ICARTT campaign in order to assess the ability of the model to reproduce the anthropogenic pollution observed in the lower troposphere over North America. During this campaign, which took place in summer 2004, strongly polluted air masses with high values of CO, O₃, HNO₃, NO_x and VOCs were sampled by the P3 aircraft in the lower troposphere over the US north-east coast. The comparison of MOZART-4 results against observations shows that the model is not able to reproduce correctly the high pollutant concentrations measured in the lower troposphere due to deficiencies in exchange of pollutants between the BL and the free troposphere and uncertainties in wet deposition of soluble species, in particular HNO₃, although uncertainties in emissions, in particular NO_x and VOCs, could also influence the model performance. Due to these discrepancies the model also underestimates the pollutant concentrations sampled on 20 and 21 July in the polluted plume identified by Methven et al. (2006) as a case of anthropogenic plume transported at low altitudes from North America toward Europe across the North Atlantic. The model does capture the decrease in pollutant concentrations in the plume between 20 and 21 July but severely underestimates measured concentrations in particular on 20 July. The Lagrangian photochemical trajectory model used by Real et al. (2008), which was initialized by the P3 measurements on 20 July, is able to reproduce the evolution of the chemical composition in the plume by taking into account mixing between the plume and its environment and introducing a more realistic treatment of wet deposition in the model. Therefore, MOZART-4 results could be in better agreement with the observations if the measured concentrations on 20 July are correctly reproduced, in particular HNO₃ which was severely underestimated due, probably, to strong wet deposition in the model. Furthermore, Real et al. (2008) attributed the high O₃ values measured in the plume in the first two days to the high HNO₃ levels on 20 July which produced

NO_x and, thus, O_3 through photolysis. While pollutant values decreased also on 22 July, modeled concentrations increased compared to the previous days. These increases suggest the mixing of the initial plume with air masses influenced by more recent emissions is overestimated in the global model. This is consistent with Real et al. (2008) who showed uncertainties in the calculated plume trajectories on 22 July. Therefore, uncertainties in MOZART-4 parameterizations regarding wet deposition and exchange between the BL and the free troposphere, in addition to uncertainties in emissions, result in underestimation of initial pollutant values in North American anthropogenic plumes and, thus, in underestimation of O_3 and other pollutants export into Europe.

ANNEX A

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

A. Colette (1), C. Granier (2,3,4,5), Ø. Hodnebrog (6), H. Jakobs (7), A. Maurizi (8), A. Nyiri (9), B. Bessagnet (1), A. D'Angiola (2), M. D'Isidoro (8, now at 10), M. Gauss (9), F. Meleux (1), M. Memmesheimer (7), A. Mieville (11), L. Rouil (1), F. Russo (8), S. Solberg (12), F. Stordal (6), F. Tampieri (8)

1. Institut National de l'Environnement Industriel et des Risques (INERIS), Verneuil-en-Halatte, France, augustin.colette@ineris.fr
2. Université Paris 6 Pierre et Marie Curie, CNRS-INSU, LATMOS-IPSL, Paris, France
3. NOAA Earth System Research Laboratory, Boulder, CO, USA
4. Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA
5. Max Planck Institute for Meteorology, Hamburg, Germany
6. Universitetet i Oslo (UiO), Oslo, Norway
7. Rhenish Institute for Environmental Research at the University of Cologne (FRIUUK), Köln, Germany
8. Institute of Atmospheric Sciences and Climate, Consiglio Nazionale Delle Ricerche, Bologna, Italy
9. Meteorologisk institutt (met.no), Oslo, Norway
10. ENEA, Bologna, Italy
11. Laboratoire d'Aerologie, Toulouse, France
12. Norsk Institutt for Luftforskning (NILU), Oslo, Norway

Abstract

We discuss the capability of current state-of-the-art chemistry and transport models to reproduce air quality trends and inter annual 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 (NO₂, O₃ and PM₁₀). 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 O₃ is more challenging. Apart from the long term trend, the modelled monthly variability 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 years, hence supporting further emission reductions strategies.

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 were further 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 now 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 economical 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 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 trend in the Western European pollution hotspots are investigated in Section 0, the modelling setup is presented in Section 0 and a short model evaluation is discussed in Section 0. The discussion of the capability of the models involved to capture observed trends is detailed in Section 0 and the inter-annual variability is addressed in Section 0. Section 0 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

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 (Konovalov et al., 2010) or tropospheric profiles (Thouret et al., 1998; Logan et al., 1999).

Also, we focus only on ozone (O₃), nitrogen dioxide (NO₂) and particulate matter with a diameter smaller than 10µm (PM₁₀). Since these basic compounds have been regulated for several years, they are widely monitored, so that we can com-

pile a significant dataset of stations offering a good coverage (including in urban areas) over the past 10 years.

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 megacities. (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 (SO_{4p}), total Nitrate (NO_{3t}=NO_{3p}+HNO_{3g}) and total Ammonia (NH_{4t}=NH_{4p}+NH_{3g}) (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-year 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.

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 years 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 12W to 30E and 35N to 65N) and the Benelux region (1W-8E, 48N-54N) 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.

Observed trends

The trends observed at each of the selected stations are displayed in Figure 1 for NO₂, O₃ and PM₁₀. 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 years long (in the best case), it was considered unnecessary 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 NO₂ 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 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 µg/m³/year, respectively. We find an absolute majority of European stations 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 O₃ 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 O₃ trend at UB, SB and RB sites is 0.37, 0.27 and 0.05 µg/m³/year, respectively.

The proportion of sites where the O₃ trend is positive is 30.8% when considering daily means but this number drops to 18.5% when considering O₃ 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.

The number of PM₁₀ monitoring stations that pass the filtering described in section 0 is by far lower than for O₃ or NO₂. For instance in France PM₁₀ 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, PM₁₀ concentrations are systematically decreasing, thanks to the air quality regulation enforced during the past decade. However, the trend of total PM₁₀ 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 PM₁₀ 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 section 0).

Sensitivity of the estimated trend to the filtering

In Section 0, we required somewhat arbitrarily that selected records should cover at least 8 years 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 years 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 stringiest criteria of a minimum coverage of 10 year 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 year 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 PM₁₀ (O₃ and NO₂ not represented for concision purposes) trends in the Benelux area, taking as reference the distribution of de-seasonalised trends covering at least 8 years over the 1998-2007 decade. It ap-

pears that central quantiles appear 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 years.

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 similitude 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 years. It can be seen that for all three pollutants, using 7, 8 or 9 years as a threshold yields similar distributions. But using 5, 6, or 10 years 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.

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 although like most other emissions inventories it suffers from a number of limitations well documented in the literature (Vestreng et al., 2009; Jonson et al., 2006; Konovalov et al., 2006). 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. accepted for publication in Climatic Change).

A couple of minor changes were 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.

For each grid point of the inventory we fitted a linear least square regression for the total emissions of PM, NO_x, and non-methane volatile organic compounds, as well as the NMVOC/NO_x ratio and plotted the map of the slope on Figure 3. NO_x emissions decreased throughout Europe, except on the ship tracks. NMVOC decreased also, except in Poland. And the trend in the ratio NMVOC/NO_x shows some interesting patterns with regards to the modelled trends of O₃ that will be

discussed later. Note that the trend of primary PM emission is much more variable geographically.

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.

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

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>.

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.

Eurad

The Eurad model (Jakobs et al., 2002; Memmesheimer et al., 2004; Memmesheimer et al., 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>.

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 and M, 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).

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 O₃, NO₂ and PM₁₀ 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, 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 4a displays the mean seasonal cycles (monthly values based on 10 years 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.

Nitrogen dioxide

All models exhibit a negative NO₂ bias when compared to suburban stations. This feature was expected as we used at best a 50km spatial resolution. The small bias of modelled NO₂ 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_x removal (which would be corroborated in the strong difference on total PM₁₀ discussed below in Section 0). 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 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 Figure 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).

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 NO₂ 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 (Figure 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 O₃ 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.

Ox

The Ox (=NO₂+O₃) climatology (global average over 10 years) is displayed on Figure 5. By filtering out the titration impact of NO_x on O₃ levels, this quantity gives an insight into the degree of photochemical activity of the models. Bolchem appears as one of the least photochemically active models (spatial and temporal global average of 62.6 ±6.2 µg /m³), and to a lesser extent Mozart is also in the lower part of the sample (63.9±9.8µg /m³). OsloCTM2 (70.4±13.9µg /m³), Emep (72.8±11.1µg /m³) and Eurad (74.7±14.9µg /m³) exhibit more similar figures while Chimere (80.8±10.0µg /m³) 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 Ox concentrations above the Mediterranean. Note also the strong influence of O₃ dry deposition schemes as shown by the sharp land/sea gradient.

Particulate Matter

PM₁₀ scores (Table 4) are not available in global model outputs which usually calculate BC/OC rather than total particulate matter. PM₁₀ correlations are much lower than for NO₂ or O₃; 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 (Figure 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.

Nitrate, Ammonium and Sulphate

The overestimation of NH₄t and NO₃t mentioned above for Bolchem and Eurad can be seen on Figure 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 H₂O₂.

5. Modelled trends

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

Nitrogen dioxide

The modelled trend of NO₂ over the whole of Europe is shown in Figure 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 (Figure 3). By contrast NO₂ 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 NO₂ 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 NO₂.

Before proceeding to the quantitative assessment of model performances, a visual comparison of the modelled (Figure 7) and observed (Figure 1) geographical patterns of these trends suggests that the models are quite successful in capturing NO₂ 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 Figure 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 Section 0 in terms of NO₂ model performances. It also shows that Eurad and Bolchem behave very similarly at the beginning of the decade, while the NO₂ decrease by the end of the period is much stronger in Eurad. All other models exhibit very similar behaviours.

While the composite on Figure 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 Figure 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 Section 0. This figure unambiguously 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 NO₂, the inter-annual 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 NO₂ trend is measured (according to the Mann –Kendall test, see Section 0) – i.e. 105 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, 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 (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 $-0.05 \mu\text{g}/\text{m}^3/\text{yr}$, $\sigma = 0.11$. In France all models underestimate the bias (the average difference between the modelled minus observed trends is $0.67 \mu\text{g}/\text{m}^3/\text{yr}$, $\sigma = 0.09$). Such country-level discrepancies – consistently produced by all 6 models – are pointing towards inaccuracies in the national inventory (in which the decreasing trend of NO₂ emission is milder than what was actually observed). However that this is contradictory with the results of (Konovalov 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.

Ozone

The maps of ozone trends are provided in Figure 9. When compared to emissions (Figure 3) and NO_x concentrations trends (Figure 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 O₃ in the Southern UK, Benelux and Germany. This behaviour relates to the switch from a VOC-sensitive towards a more NO_x-sensitive regime (Beekmann and Vautard, 2010) because of the sharp decrease of NO_x 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 (Section 0), 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_x emissions, see Figure 3) does not yield a stronger O₃ production.

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 NO₂ trends discussed above (Section 0) it is thus more cautious to leave this country out of the O₃ trend analysis.

Over Northern Italy, the modelled geographical patterns are highly variable, as well as emission 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 (Figure 10) is large, the models perform decently considering that 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.

Particulate Matter

The modelled PM10 trends obtained by the regional CTMs are displayed on Figure 11. A widespread decrease of PM10 is modelled over most of Europe, except for Spain and France. More curious 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 (Figure 3). The increase in PM over the north Atlantic simulated by Emep results probably from a meteorological change which has an impact on sea-salts (as this feature also appears in the constant emission simulations, see Section 0).

The decreasing trend is not reflected in the composite on Panel (a) of Figure 12 because this composite is influenced by Czech and Spanish stations where an increase is observed. Panel (b) of Figure 12 confirms that positive trends are virtually not captured by any model (without distinction of the countries: all stations are displayed on Figure 12b) thus questioning the role of anthropogenic emissions (Section 0). Assuming that this mismatch is not due to a model shortcoming (that shows to perform well elsewhere), either such trends are 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 PM10 with a fraction of significant trends with correct sign of 65, 62, 68, and 71% respectively for Bolchem, Chimere, Emep and Eurad.

6. Inter-annual 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 inter-annual 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 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 inter-annual variability we also consider the standard deviation of the residuals between de 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, Figure 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 Section 0 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 O3 variability is driven by the seasonal cycle. And once that cycle has been removed, the remaining variability (inter-annual) 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 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-years 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 Section 0, the inter-annual 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 Figure 14, Figure 15, and Figure 16 for NO₂, O₃ and PM₁₀ respectively. When the absolute value of this metric exceeds 1, the role of emission reduction on the modelled trend can be considered as more important than the inter-annual meteorological variability over the 1998-2007 decade.

The patterns of NO₂ NRT are widely consistent with NO₂ emission changes (Figure 3) because NO₂ 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 Section 0. 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 (Section 0) that the observed and modelled trends of O₃ in Europe during the 1998-2007 period are slightly positive over European megacities where the confidence on NO₂ 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 O₃ 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 O₃ increase modelled by Emep and Eurad for the CTRL simulation (Figure 9) so that the patterns of NRT are less pronounced for these models on Figure 15.

In the CST simulation, PM₁₀ concentrations exhibit very small trends except in France where a slight positive trend is captured by all models and over the North Atlantic where Eme_p shows an increase of sea-salt (already mentioned in Section 0). The NRT patterns on Figure 16 are thus very close to the modelled trends on Figure 11, except in France where the decrease is stronger, and over the North Atlantic where the positive trend in Eme_p results vanishes. Perhaps the most surprising finding is a relatively similar trend for PM₁₀ in Eurad results in the CST and CTRL simulations, yielding milder patterns on Figure 16 compared to Figure 11 (see e.g. the absence of a negative trend North of Morocco and Algeria on Figure 16). Otherwise most models show that the order of magnitude of the decrease of PM₁₀ due to anthropogenic emissions management reaches or exceeds the natural variability over most of Europe.

8. Conclusion

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 inter-annual variability of air pollution in major anthropogenic emission hotspots in Europe. A special attention is given to the cluster of European megacities 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 years 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 inter-annual 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 profiles in the emissions inventories. We also found that the models exhibited various degree of photochemical activity, hence leading to quite variable O₃ modelling skills. The performances of the RCTM to model aerosols could be divided in two broad types of behaviour: small bias in total PM₁₀ due to an overestimation of ammonium nitrate, or a strong negative PM₁₀ 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 aggre-

gated 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 study involving global models. Downwards trends of NO₂ 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. PM₁₀ trends were also quite well captured, although the validation could not be as quantitative because of the relative lack of long term measurements. O₃ 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 O₃ 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 O₃ over the Benelux hotspot of emissions. Given the photochemical regimes dominating there, more ambitious VOC 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 NO₂ 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 NO₂ background levels. Consequently, this result suggests that ambitious environmental policies have a beneficial impact on NO₂ 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).

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 lead to localised increases of ozone, especially 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 inter-annual variability.

9. 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.

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, *Atmospheric Environment*, 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, 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, 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., Lioussé, C., Moukhtar, S., Pun, B., Seigneur, C., and Schulz, M.: Regional modeling of carbonaceous aerosols over Europe—focus on secondary organic aerosols, *Journal of Atmospheric Chemistry*, 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, 2003.
- Braniš, M.: Long Term Trends in Concentration of Major Pollutants (SO₂, CO, NO, NO₂, O₃ and PM₁₀) in Prague – Czech Republic (Analysis of Data Between 1992 and 2005), *Water, Air, & Soil Pollution: 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, *Meteorology and Atmospheric Physics*, 53, 137-153, 1994.
- CAFE: 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 Tzagatakis, I.: Trends in NO_x and NO₂ 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, *Atmospheric Environment. 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., Dumitrean, P., and Kent, A.: Photochemical ozone formation in north west Europe and its control, *Atmospheric Environment*, 37, 1983-1991, 2003.
- 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, 2010.
- Favez, O., Cachier, H., Sciare, J., and Le Moullec, Y.: Characterization and contribution to PM_{2.5} of semi-volatile aerosols in Paris (France), *Atmospheric Environment*, 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, *Atmospheric Environment*, 37, 1503-1519, 2003.
- Giorgi, F., Jones, C., and Arsar, 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.
- 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, 2006.
- Harrison, R. M., Stedman, J., and Derwent, D.: New Directions: Why are PM₁₀ concentrations in Europe not falling?, *Atmospheric Environment*, 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, 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., 2005.
- Hollander, M., and Wolfe, D. A.: Nonparametric Statistical Methods., in 2 ed., John Wiley & Sons, New York, 27-33; 68-75, 1999.
- Honoré, C., Menut, L., Bessagnet, B., Meleux, F., Rouil, 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, 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, 2005.

- Jakobs, H., Tilmes, S., Heidegger, A., Nester, K., and Smiatek, G.: Short-Term Ozone Forecasting with a Network Model System during Summer 1999, *Journal of Atmospheric Chemistry*, 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, 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, *Journal of Atmospheric Chemistry*, 55, 167-187, 2006.
- Kendall, M. G.: Rank Auto Correlation Methods, 4 ed., Griffin, 1976.
- Konovalov, I. B., Beekmann, M., Richter, A., and Burrows, J. P.: Inverse modelling of the spatial distribution of NO_x emissions on a continental scale using satellite data, *Atmos. Chem. Phys.*, 6, 1747-1770, 2006.
- 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., Burrows, J. P., and Hilboll, A.: Multi-annual changes of NO_x emissions in megacity regions: nonlinear trend analysis of satellite measurement based estimates, *Atmos. Chem. Phys.*, 10, 8481-8498, 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, *Atmospheric Environment. Part A. General Topics*, 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, 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, Oslo, 47-62, 2004.
- Memmesheimer, M., Friese, E., Ebel, A., Jakobs, H. J., Feldmann, H., Kessler, C., Piekorz, and G.: Long-term simulations of particulate matter in Europe on different scales using sequential nesting of a regional model, 1-2, Inderscience Enterprises, Geneva, SUISE, 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, 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, *Atmospheric Environment*, 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., Liousse, 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, *Atmospheric Environment*, 43, 5268-5350, 2009.
- 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., Hueglin, C., Staehelin, J., and Prevot, 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, 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. Cycles*, 22, GB2002, 2008.

- Sillman, S.: The relation between ozone, NO_x and hydrocarbons in urban and polluted rural environments, *Atmospheric Environment*, 33, 1821-1845, 1999.
- Simpson, D., Andersson-Skold, Y., and Jenkin, M. E.: Updating the chemical scheme for the EMEP MSC-W oxidant model : current status, The Norwegian Meteorological Institute, EMEP, Oslo, 1993.
- Simpson, D.: Biogenic emissions in Europe 2. Implications for ozone control strategies, *J. Geophys. Res.*, 100, 22891-22906, 1995.
- 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, 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, *Atmospheric Environment*, 42, 1777-1789, 2008.
- Thouret, V., Marenco, A., Logan, J. A., Nédélec, P., and Grouhel, C.: Comparisons of ozone measurements from the MOZIC 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 parameterization in large-scale models, 8, American Meteorological Society, Boston, MA, ETATS-UNIS, 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.
- 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, *Atmospheric Environment*, 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, 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, *Atmospheric Environment*, 43, 4822-4832, 2009.
- Vestreng, V., Breivik, K., Adams, M., Wagener, A., Goodwin, J., Rozovskaya, 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_x emissions in Europe with focus on road transport control measures, *Atmos. Chem. Phys.*, 9, 1503-1520, 2009.
- Wesely, M. L.: Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models, *Atmospheric Environment* (1967), 23, 1293-1304, 1989.

	Europe			Benelux		
	Total available	8 years with 75% annual coverage	Passed visual screening	Total available	8 years with 75% annual coverage	Passed visual screening
O3	1855	717	705	339	166	162
NO2	1997	669	649	354	160	158
PM10	1533	166	164	252	30	30
SO2	56	38	37	5	0	0
SO4p	54	41	37	3	3	3
NO3t	41	24	22	0	0	0
NH4t	37	25	21	0	0	0

Table 1 : Number of available in-situ surface records obtained from the AIRBASE repository (O3, NO2, and PM10) or the EMEP network (SO2, SO4p, NO3t, and NH4t) before and after applying the quality check criteria, and for both the whole European domain and the Benelux hotspot.

	NO2			O3			PM10		
	# sta.	p-val.	Median trend (µg/m3/yr)	# sta.	p-val.	Median trend (µg/m3/yr)	# sta.	p-val.	Median trend (µg/m3/yr)
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

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-years threshold), and the median of the distribution of the trends at all stations.

		BOLCHEM	CHIMERE	EMEP	EURAD	OsloCTM2	Mozart 4
Chemical Mechanism	Gas-phase	SAPRC90 (Carter, 1990)	Melchior 2 reduced: 44 species, 120 reactions (Lattuati, 1997).	EMEP (Simpson et al., 1993), with subsequent updates (Simpson et al., 2011)	RADM2, RACMMIM (Stockwell et al., 1997; Geiger et al., 2003; Karl et al., 2006)	QSSA solver (Hesstvedt et al., 1978) (Berntsen and Isaksen, 1997) for ozone/NO _x /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 (0.5deg)	Europe (0.5deg)	Europe (50km)	Europe (50km)	Global (T42, ~2.81deg)	Global (T63, ~1.8deg)
	Number of vertical levels (min, agl-)	16 (40m agl- 500hPa)	8 (30m agl- 500hPa)	20 (90m agl - 100hPa)	23 (20m agl- 100 hPa)	60 (8m agl. - 0.1 hPa)	28

	max)						
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).	O3: observation-based climatology modulated by potential vorticity, latitude and interannually. (Logan, 1999)	O3: observation-based climatology modulated interannually. (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 ₂ 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 pers. comm.)	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 pers. comm.)

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

D1.2.3

FP7 – CityZen

Deliverable



	NO2			O3			PM10		
	R	Bias ($\mu\text{g}/3$)	RMSE ($\mu\text{g}/3$)	R	Bias ($\mu\text{g}/3$)	RMSE ($\mu\text{g}/3$)	R	Bias ($\mu\text{g}/3$)	RMSE ($\mu\text{g}/3$)
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			

Table 4 : Model performances at AIRBASE suburban station computed over 10years on the basis of daily means.

	Europe	AT	CH	CZ	DE	FR	NL
Avg (σ)	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	105	14	6	7	35	12	5

Table 5 : Fraction of sites where the sign of the NO2 trend is correctly captured by the models (average – avg - and standard deviation – σ – 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).

	AT	BE	CH	CZ	DE	ES	
Avg (σ)	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 (σ)	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

Table 6 : Percentage of underestimation in the modelled median inter-annual variability (average – avg - and standard deviation – σ) at all stations of a given country and across all models. The inter-annual 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).

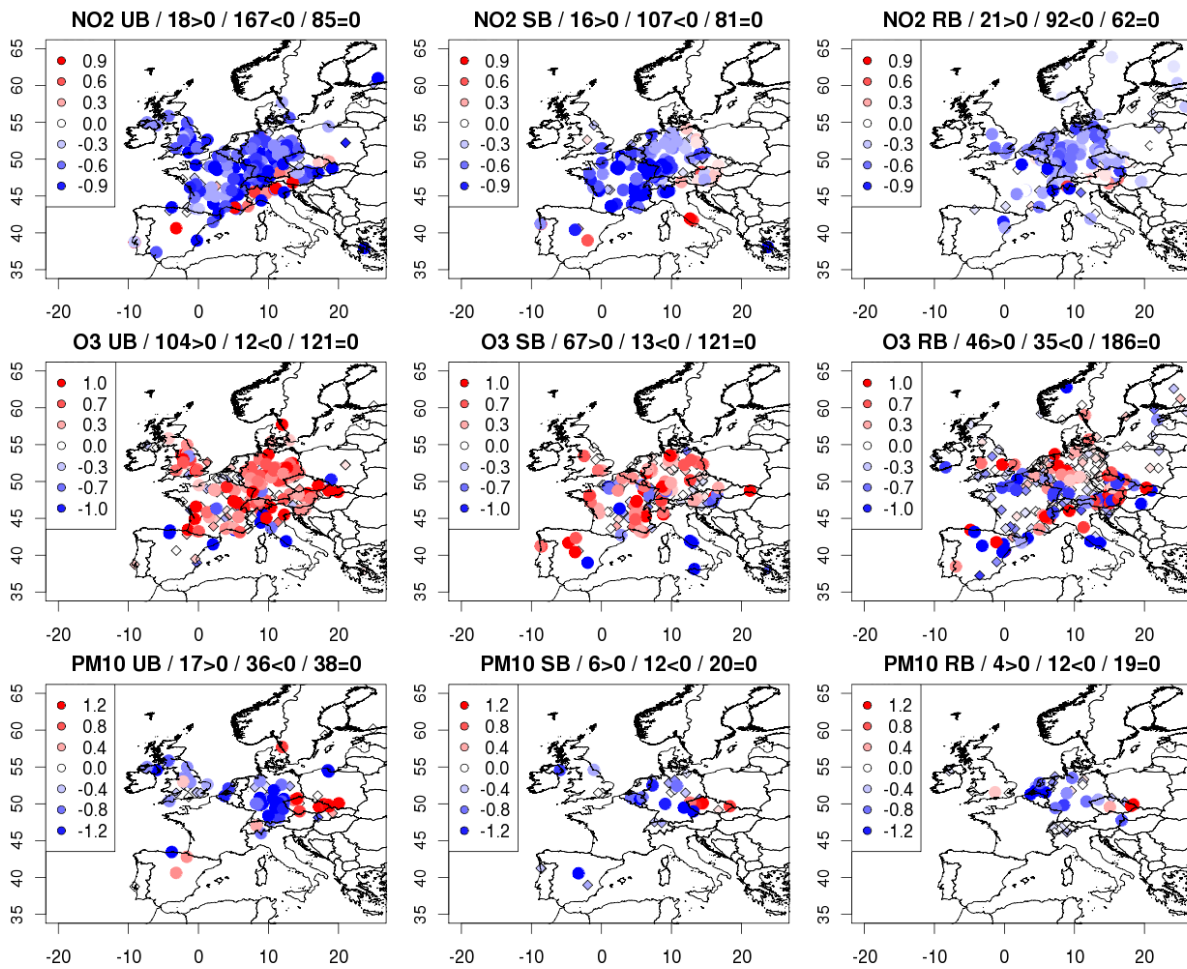


Figure 1 : Trends of daily means of NO₂, O₃, and PM₁₀ (µg/m³/yr) observed at urban background (UB), suburban background (SB) and rural background (RB) airbase stations. Stations where a statistically significant trend is observed are shown with a large dot a small diamond is used otherwise. The title of each panel also provides the number of stations with a positive, negative or null (not significant) slope.

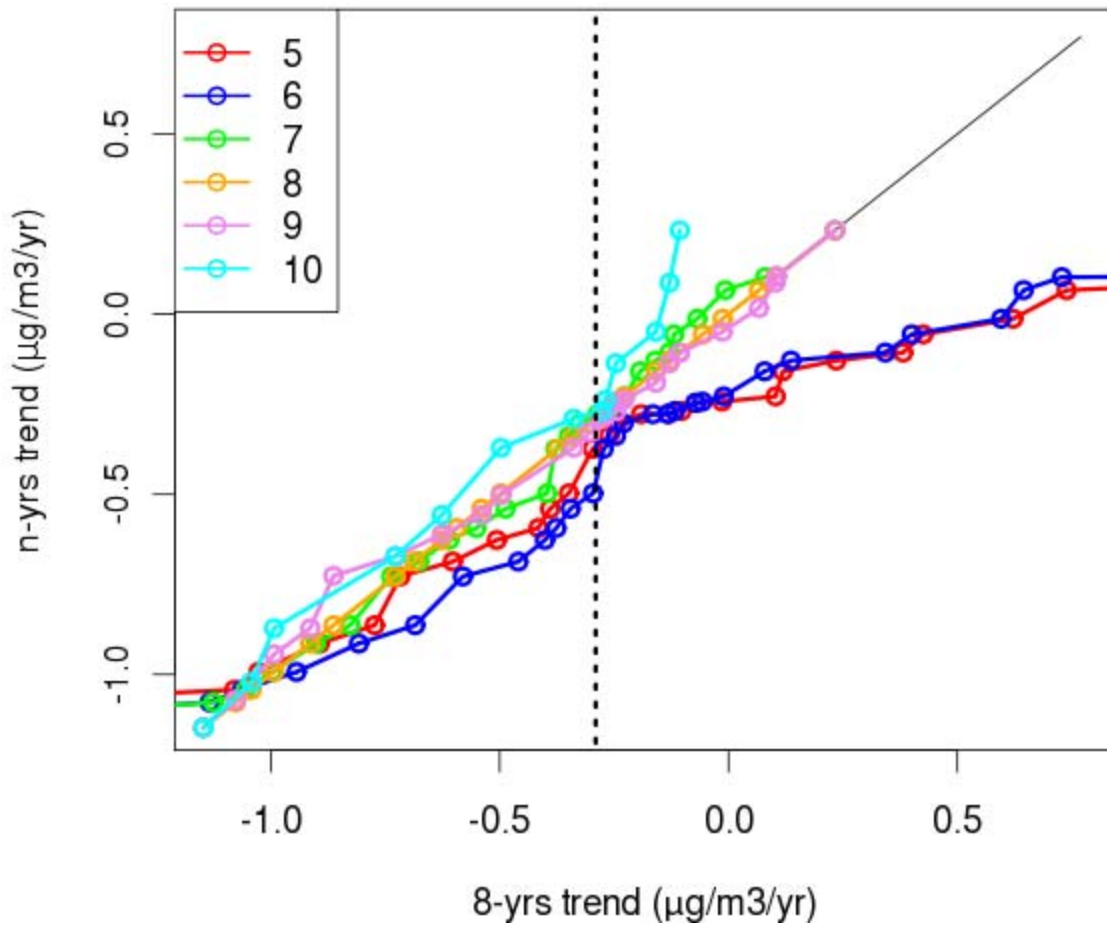


Figure 2 : Comparison of the distributions of PM10 trends ($\mu\text{g}/\text{m}^3/\text{yr}$) in Benelux, depending on the threshold of minimum number of years used in the data quality check procedure (8 years being used as a reference on the x-axis). The vertical dashed line shows the median of the reference distribution.

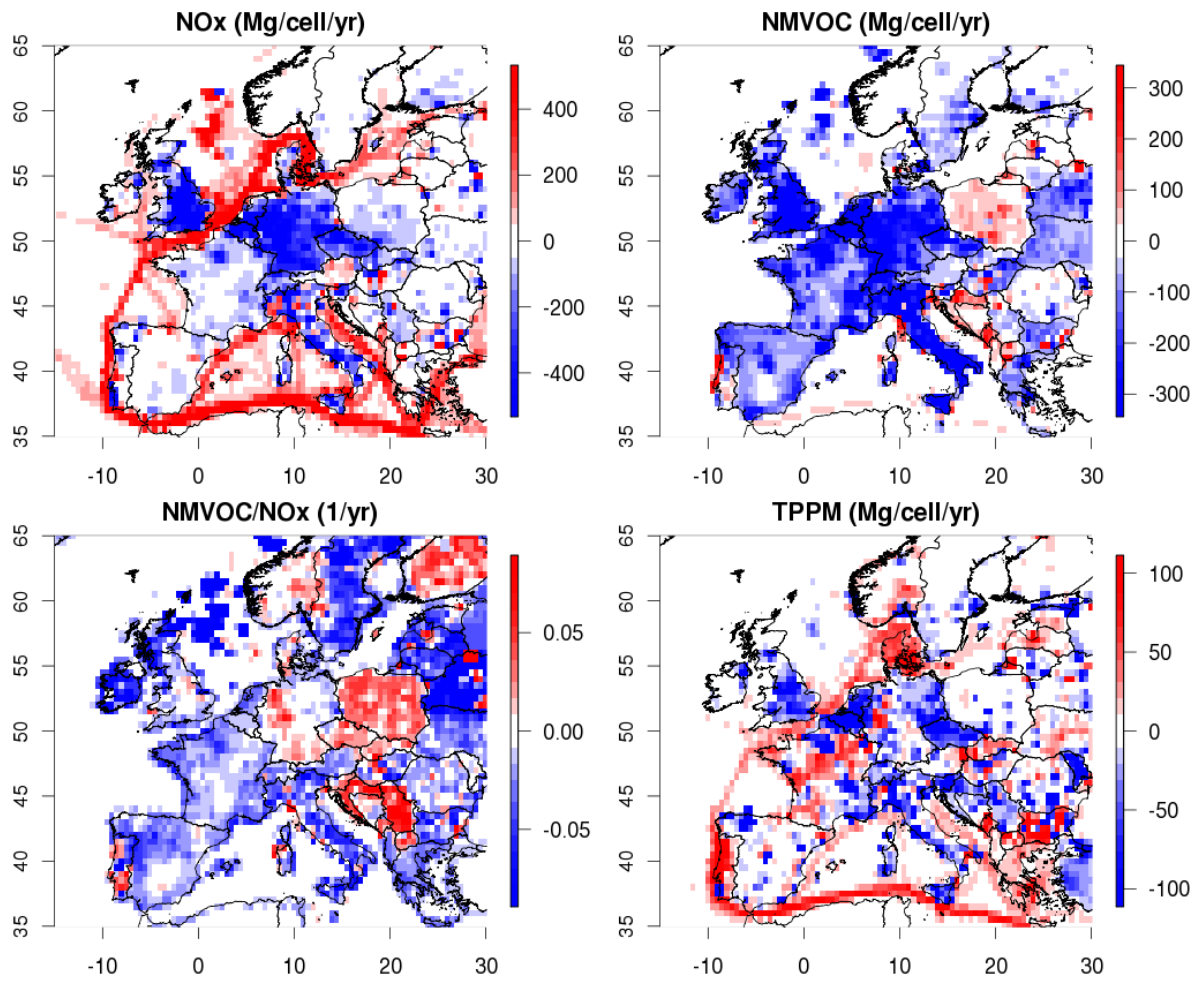


Figure 3 : Map of EmeP expert emissions trends (linear least square fit of annual totals) over 1998-2007 for NO_x, NMVOC, NMVOC/NO_x and total primary PM (TPPM). Units are Mg/yr except for NMVOC/NO_x (yr⁻¹).

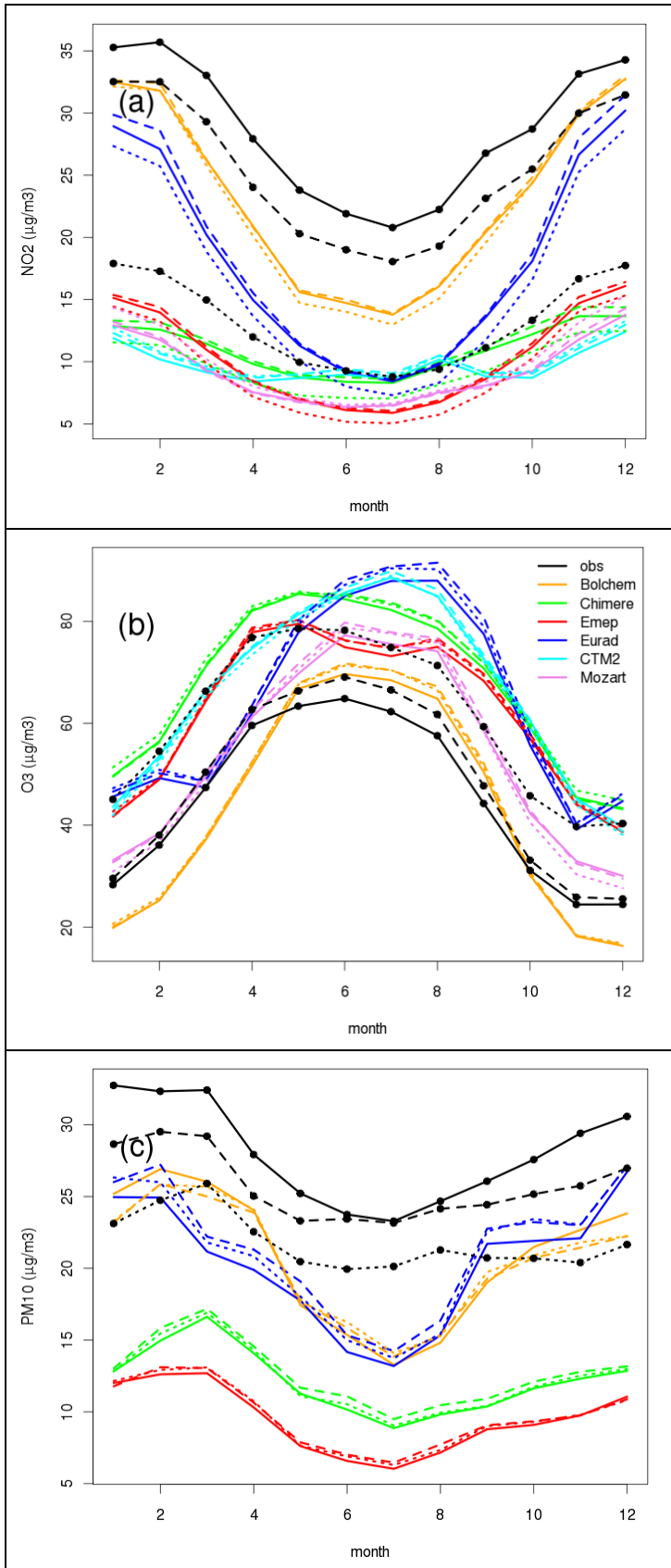


Figure 4 : Seasonal cycles based on 10 years of daily mean values for NO2 (top), O3 (middle), and PM10 (bottom) observed (black) and modelled by Bolchem (orange), Chimere (green), Emep (red), Eurad (blue), OsloCTM2 (cyan) and Mozart (violet). Solid lines for urban background, dashed: suburban, and dotted: rural AIRBASE stations.

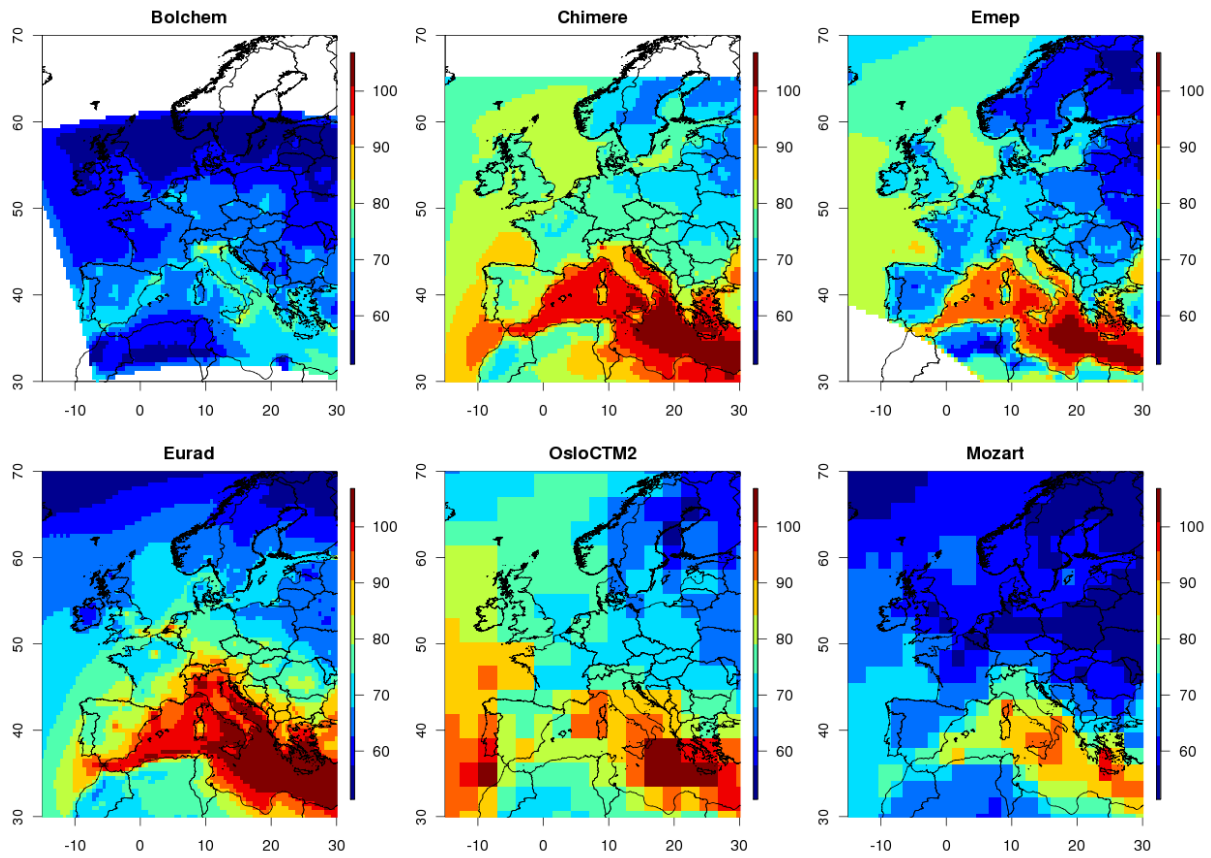


Figure 5 : Average modelled Ox ($\mu\text{g}/\text{m}^3$) [=NO₂+O₃] fields over 10 years for each model.

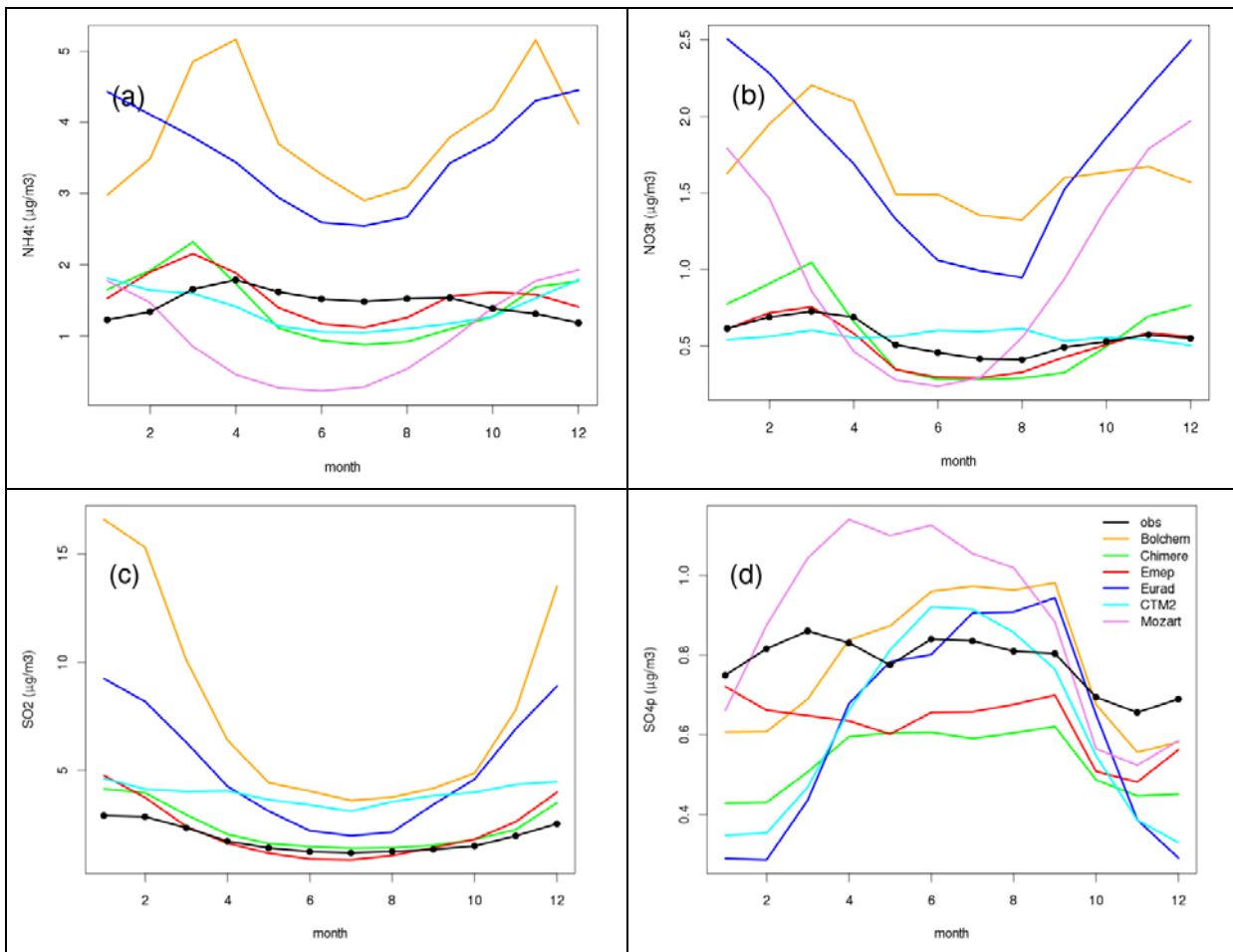


Figure 6 : Same as Figure 4 for total NH₄t (a), total NO₃t (b), SO₂ (c) and SO₄p (d) recorded at Emep background stations.

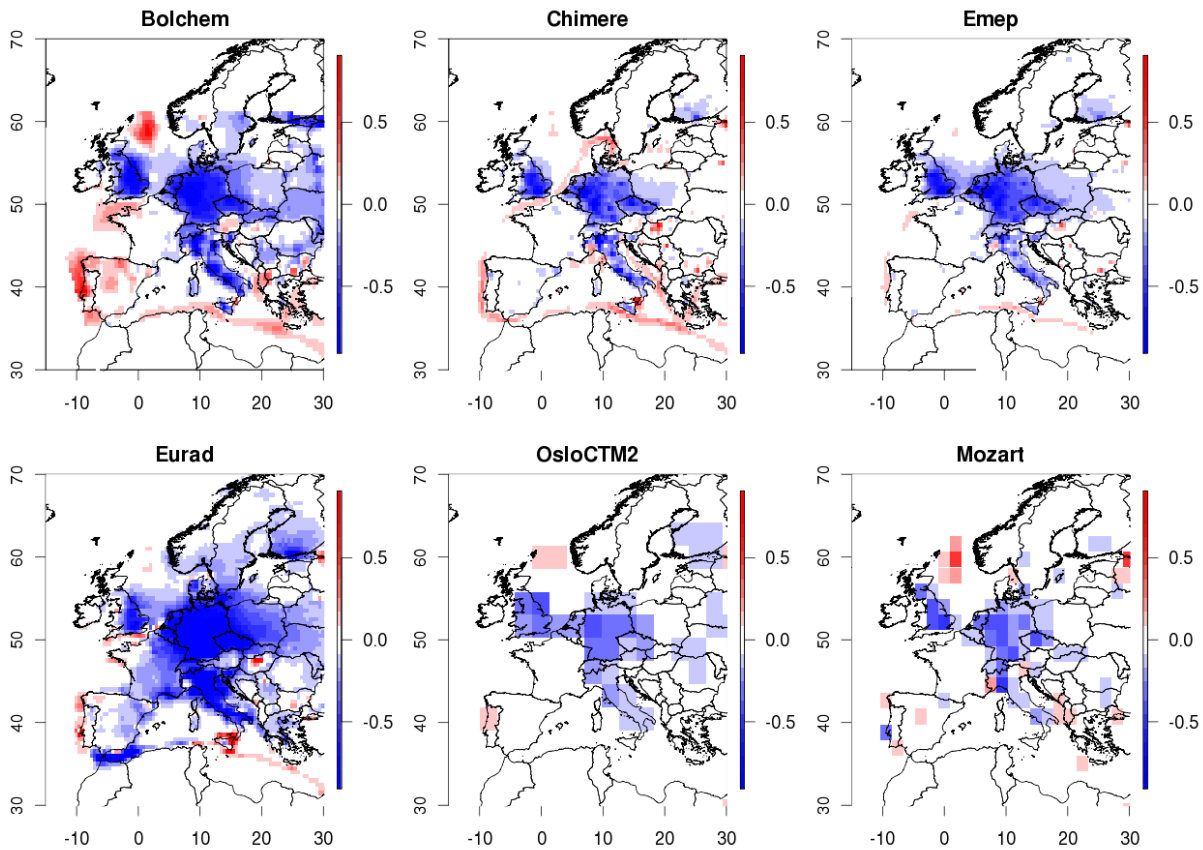


Figure 7: Modelled NO₂ trend (µg/m³/yr) for each CTM and at each grid point computed on the basis of monthly means of daily means over the 1998-2007 period with a linear least square fit of de-seasonalised values.

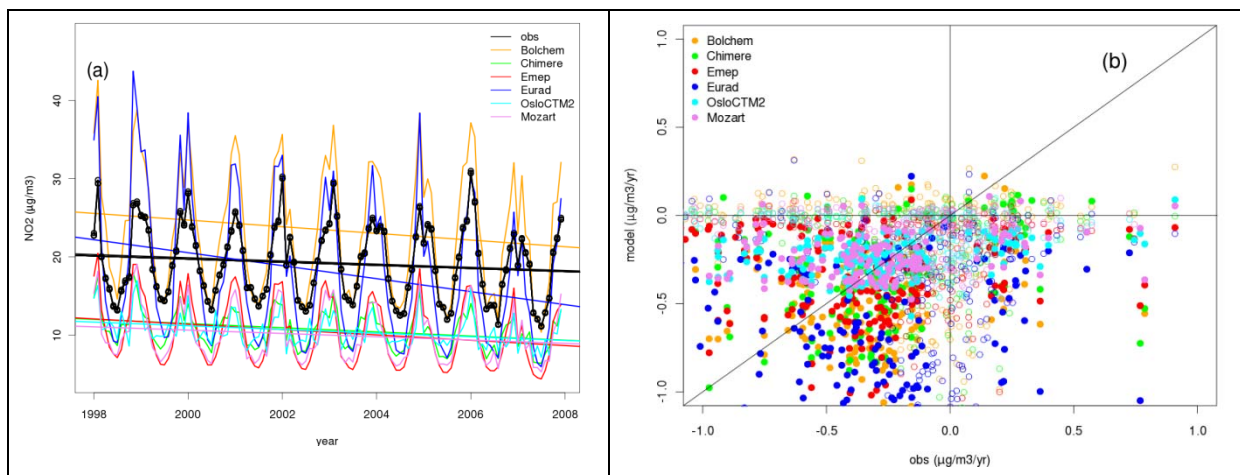


Figure 8: (a) European-wide composite of modelled and observed monthly means of NO₂ trend (µg/m³) at the air quality monitoring stations of background suburban and rural type. The straight line shows the best linear least square fit. (b) scatter plot of modelled and observed trend (computed as linear least square of the de-seasonalised time series, in µg/m³/yr) at each individual station. Sites where a significant slope is computed are marked with a filled symbol.

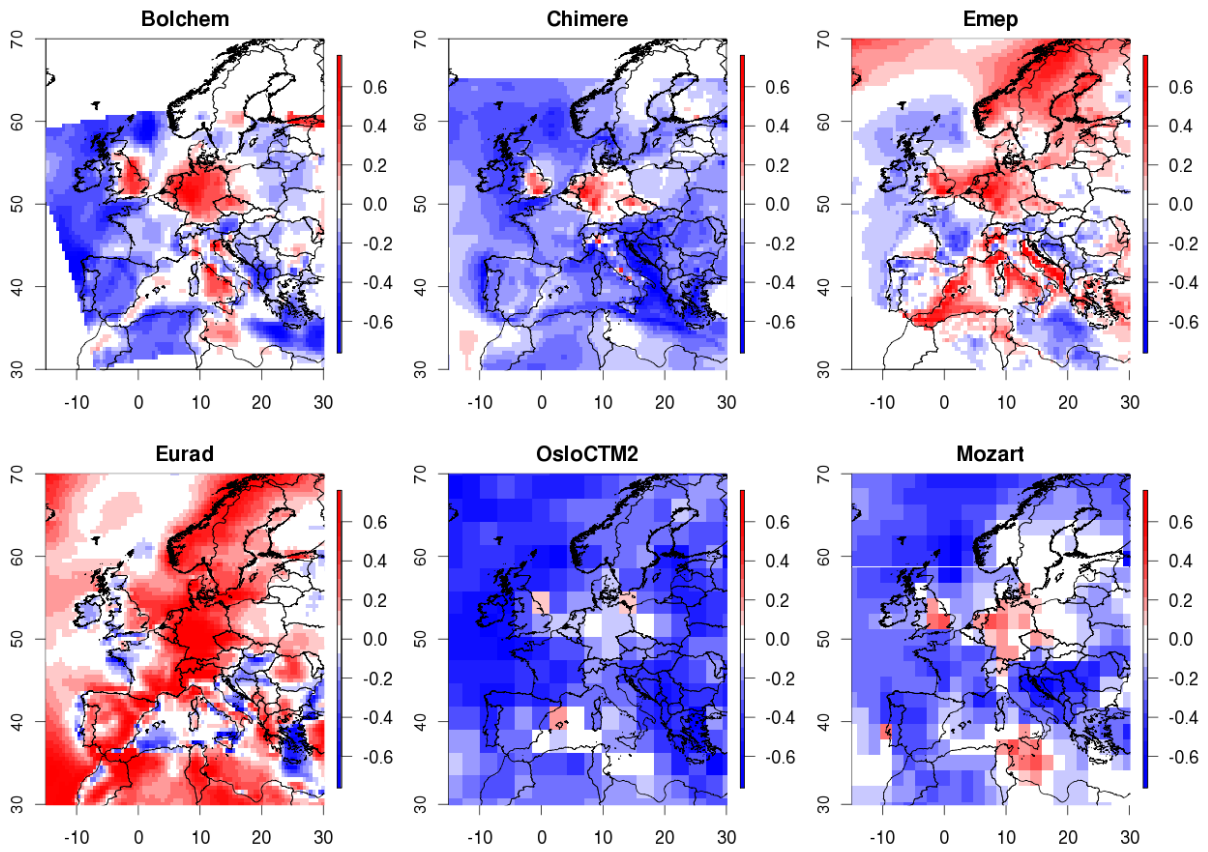


Figure 9 : Same as Figure 7 for O₃.

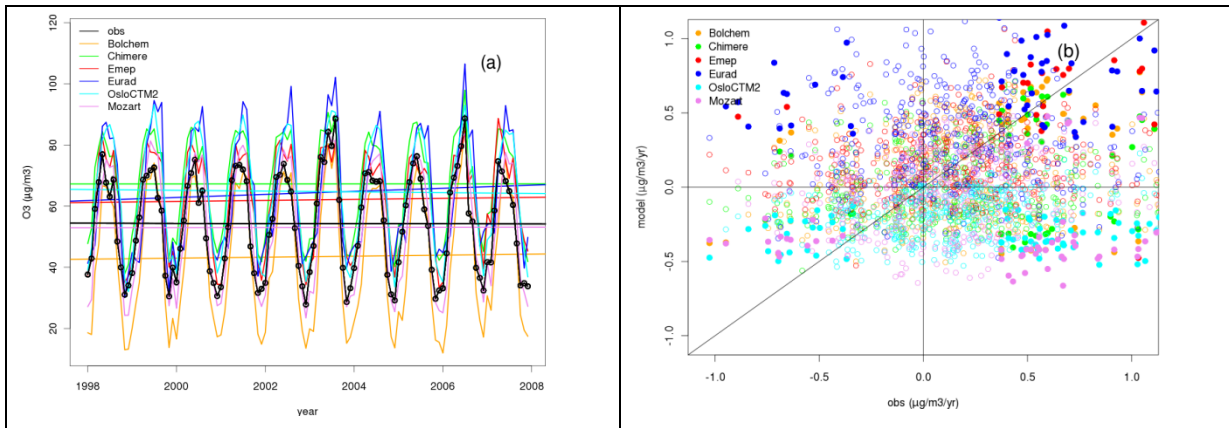


Figure 10 : Same as Figure 8 for O₃.

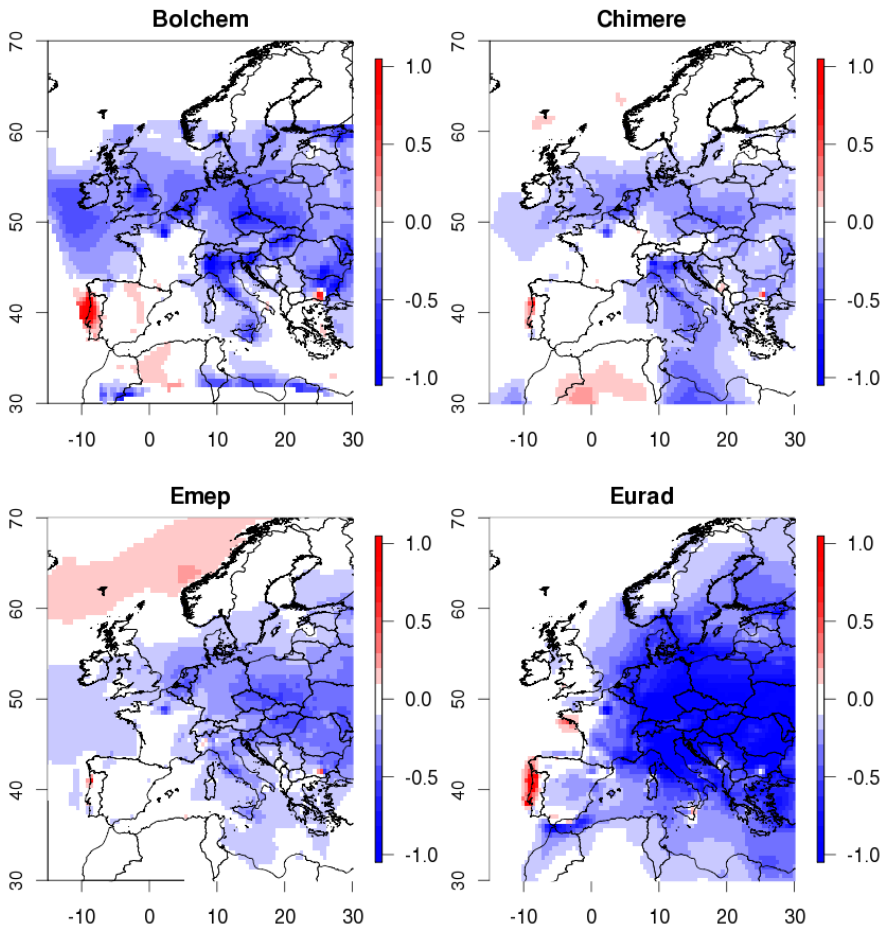


Figure 11 Same as Figure 7 for PM10.

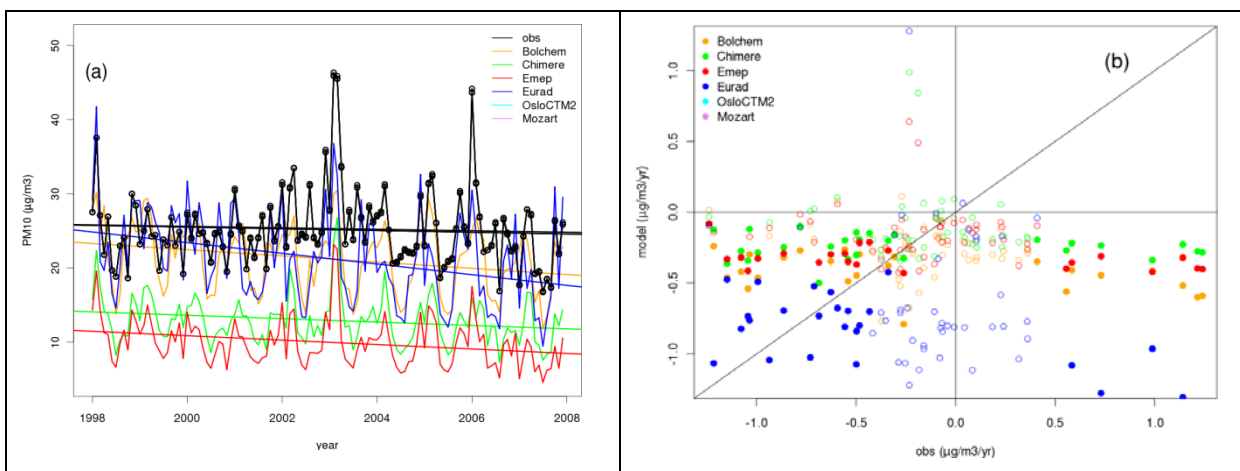


Figure 12 Same as Figure 8 for PM10.

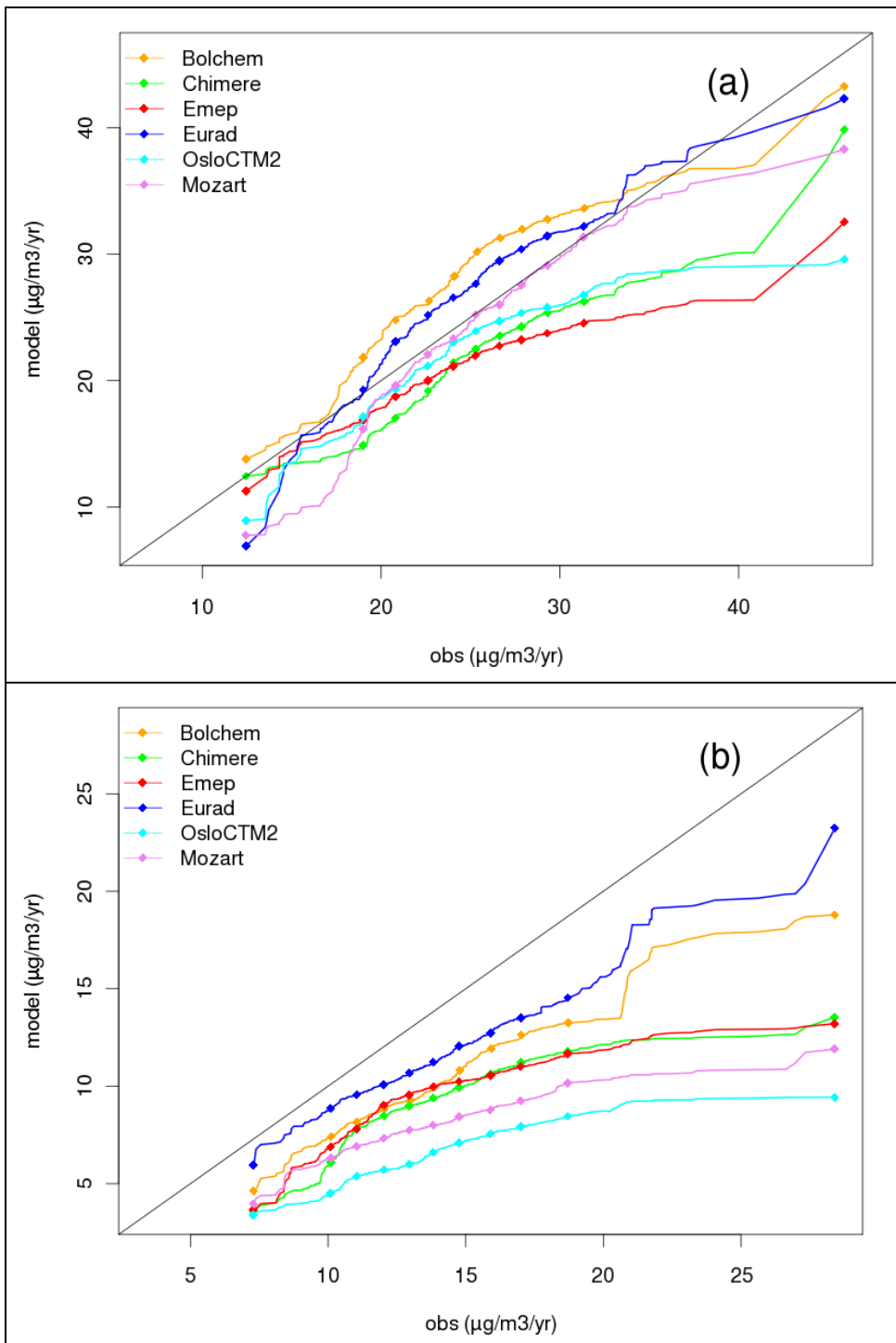


Figure 13 : Top panel: quantile-quantile plot of the standard deviation of the residuals of monthly mean O₃, once the linear trend has been removed, observations (x-axis) being used as a reference. In the bottom panel, the seasonal cycle has also been removed. The dots indicate the percentiles by multiples of 10 (0, 10th, 20th, ..., 100th).

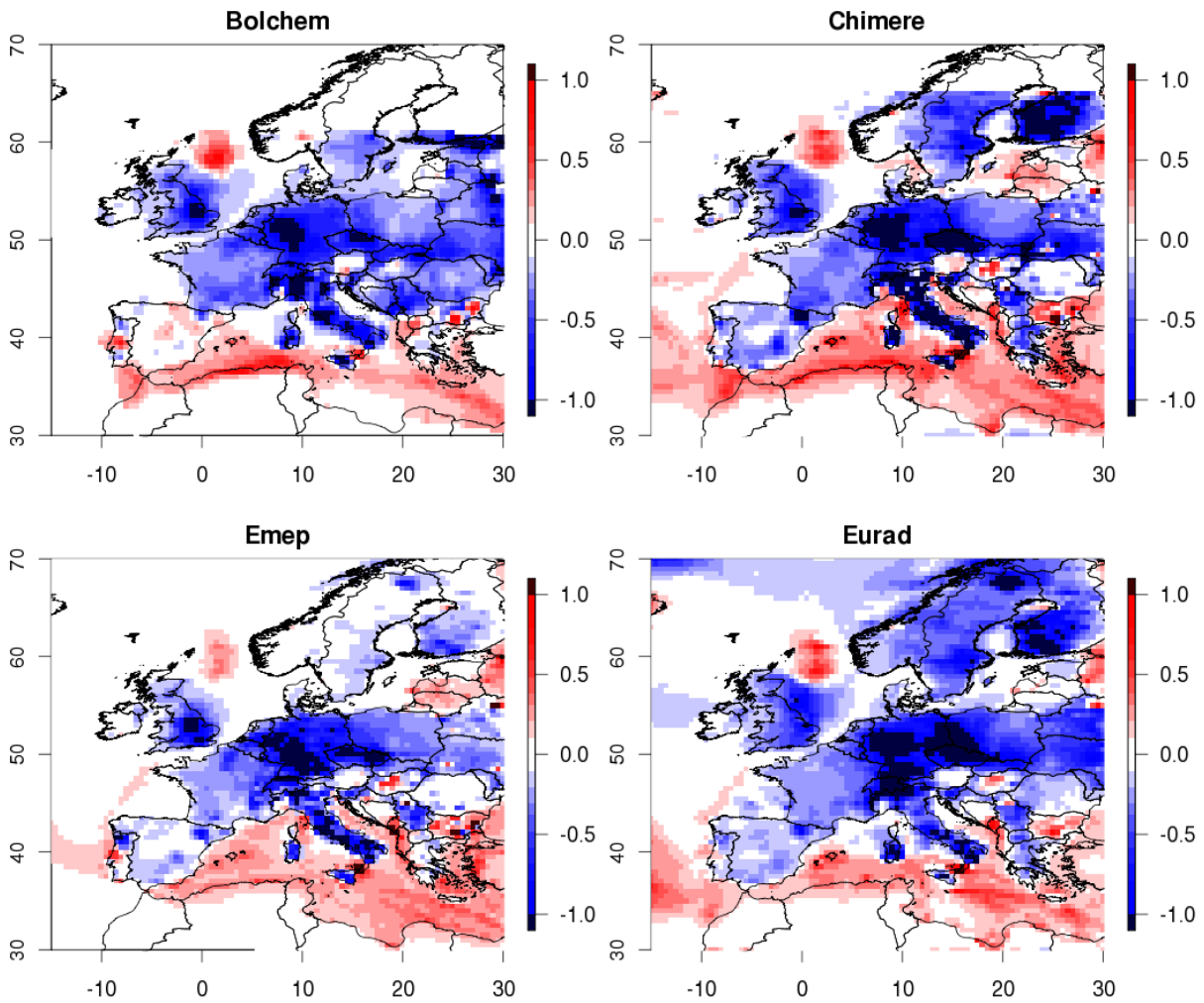


Figure 14: Trend of NO₂ due to the anthropogenic emission evolution alone (linear least square fit of the difference between the reference run and a simulation with constant – 2007 – emissions), normalised by the inter-annual meteorological variability (standard deviation of the simulation with constant emissions).

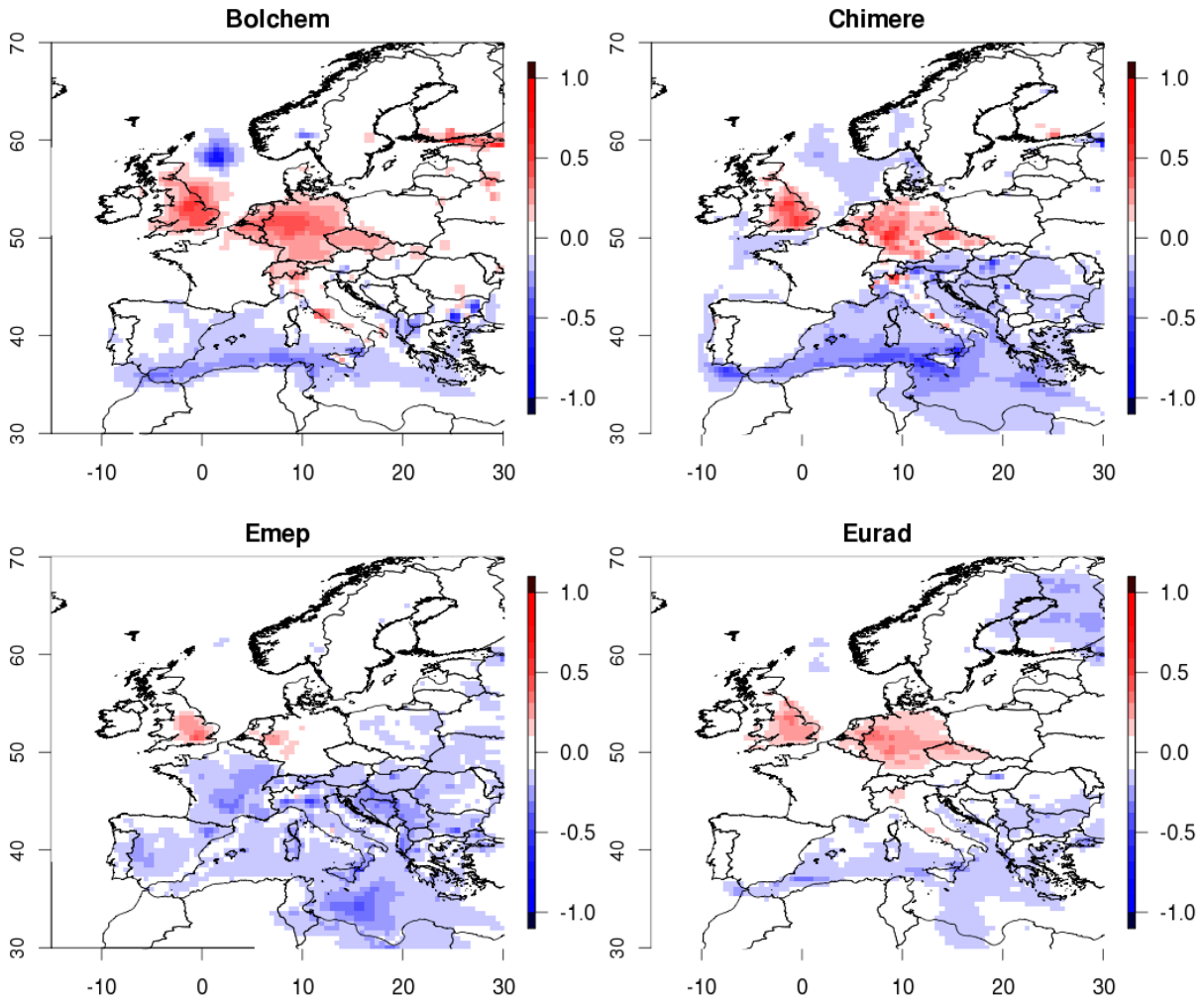


Figure 15 : Same as Figure 14 for O3.

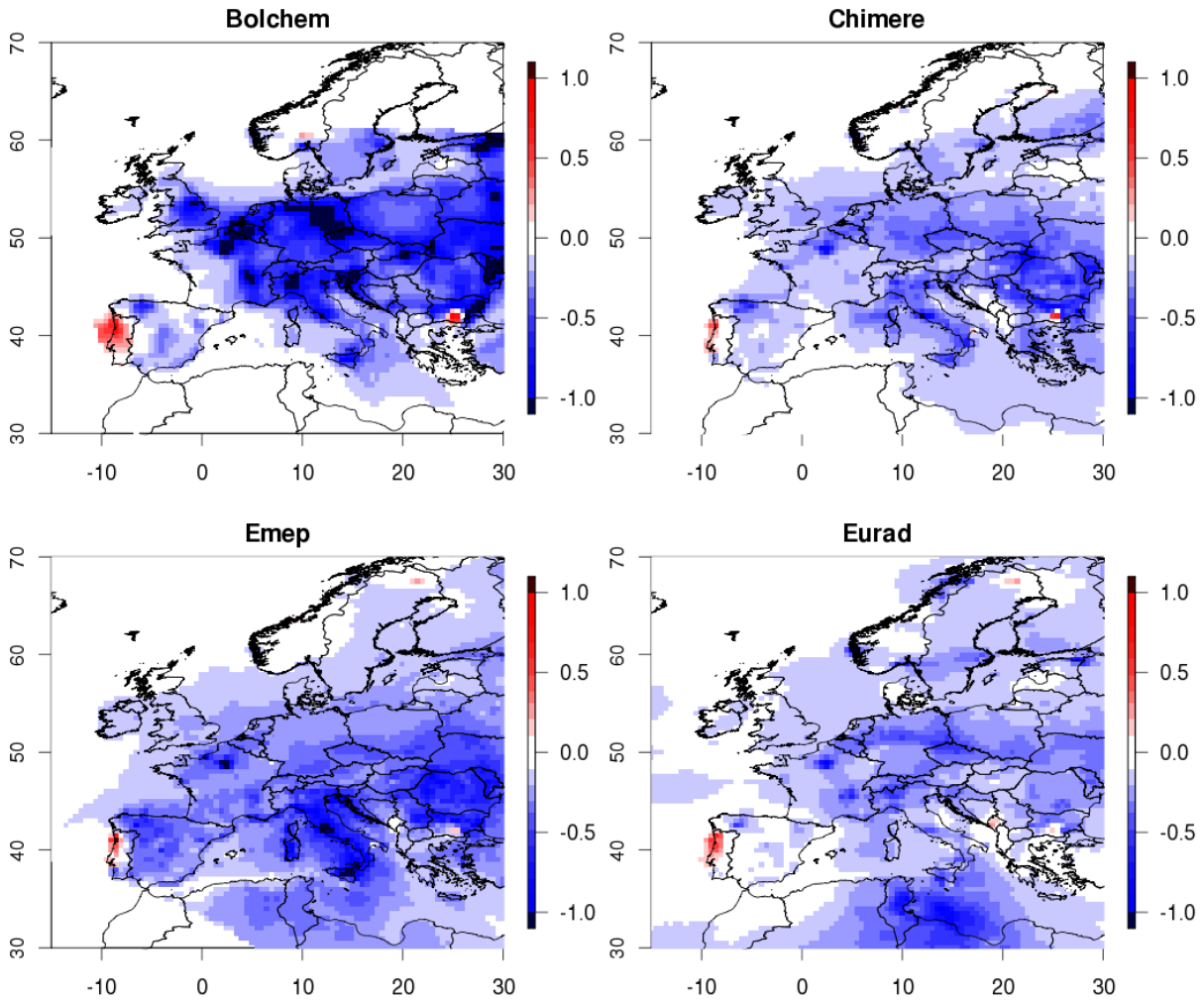


Figure 16 : Same as Figure 14 for PM10.

ANNEX B

Report on assessment of errors in global model simulations based on comparison of an Eulerian model with observations

I. Bouarar¹, K. Law¹, C. Granier¹, L.K. Emmons²

¹ UPMC Univ. Paris 06, Univ. Versailles Saint-Quentin, CNRS/INSU, LATMOS/IPSL, Paris, France

² National Center for Atmospheric Research, Boulder, CO, USA

1. Introduction

It is recognized that in order to investigate air quality over Europe we should take into account impacts from long-range transport of pollutants from other continents, in particular from North America. Several studies have shown that air masses crossing the Atlantic are highly polluted due to large biomass burning and anthropogenic emissions over North America (Stohl et al., 2002; Owen et al., 2006). Recent studies focused on the impact of North American anthropogenic pollution on O₃ over Europe (Li et al., 2002; Methven et al., 2006; Real et al., 2008). These studies showed that low level transport of such pollution can lead to important impacts on European boundary layer (BL hereafter) O₃. They reported several polluted plumes measured over the west coast of Europe with North American origins.

In this study, we analyse the capability of the chemistry transport model MOZART-4 to reproduce the anthropogenic pollutants measured in the lower troposphere during the ICARTT (International Consortium for Atmospheric Research on Transport and Transformation) campaign which took place in summer 2004 between North America and Europe (Fehsenfeld et al., 2006). One of the main objectives of this project was to investigate anthropogenic pollution transport from North America. We present an evaluation of MOZART-v4 performance during the first few days downwind from emission regions through comparison with in-situ data measured by the NOAA-P3 aircraft (P3 hereafter) and by the MOZAIC (Measurement of OZone and water vapour by Airbus In-service aircraft) program (Marenco et al., 1998) as well as with satellite CO measurements from the MOPITT (Measurement Of Pollution In The Troposphere) instrument (Deeter et al., 2004). We also examine the ability of the model to reproduce the concentrations of CO, NO_x, O₃, HNO₃ and volatile organic compounds (VOCs) measured in polluted plumes sampled by the P3. The 6-hourly average concentrations from the model have been spatially interpolated to the location of the P3 and MOZAIC observations before comparison. Model performance is also compared to results from Lagrangian model runs performed by Real et al. (2008).

2. Case study

In this study, we focus on the anthropogenic pollution measured onboard the NOAA P3 aircraft in the lower troposphere over the north-east coast of the US on 3 successive days (20, 21 and 22 July 2004) during the ICARTT campaign. The P3 flight tracks on these days are shown in Figure 1. Details about the aircraft instruments and the general strategy of ICARTT are summarized by Fehsenfeld et al. (2006). The aircraft measurements were made near New York on 20 July. On 21 and 22 July, the aircraft flew downwind over Gulf of Maine and Nova Scotia. Using Lagrangian trajectories and hydrocarbon fingerprint analysis, Methven et al. (2006) identified an anthropogenic polluted plume which was sampled by the P3 in the lower troposphere on the three successive days. Figure 1 also shows three forward trajectories initialised in the Lagrangian plume sampled by the P3 on each of the 3 days. These trajectories indicate that the sampled air parcels were transported towards Europe. Methven et al. (2006) showed that this plume was also intercepted by the DLR Falcon aircraft over the west of Ireland on 25 July. The anthropogenic plume sampled by the P3 is therefore an example of export of North American pollution toward Europe at low altitudes. Previous studies on summertime meteorology over North America (e.g. Li et al., 2005) reported the important role of cold fronts crossing the north-east of the US during summer on low level transport of North American pollution. Real et al. (2008) analysed the meteorological situation during the ICARTT campaign and also showed that an important outflow was observed below 3 km over the P3 flight region after the passage of a cold front. They also showed that this outflow, which transported air from the New York region to the north-east, was followed by strong eastward flow that occurred over the North Atlantic between 22 July and 25 July. These winds allowed to North American pollution to be exported to Europe.

3. The MOZART-4 model

MOZART-4 (Model of Ozone And Related Tracers, version 4) is a global chemistry transport model developed jointly at several institutions in model development (Emmons et al., 2010). The model was driven by meteorology from the NCAR reanalysis of the National Centers for Environmental Prediction (NCEP) forecasts (Kalnay et al., 1996). The model simulation was performed with horizontal resolution of approximately $2.8^\circ \times 2.8^\circ$ (approximately 280 km) and 28 vertical levels extending from the surface to around 2 hPa. It was run with a chemical mechanism which includes 85 gas-phase species, 12 bulk aerosol compounds, 39 photolysis and 157 gas-phase reactions. Monthly average biogenic emissions are taken from the POET (Precursors of Ozone and their Effects in the Troposphere) inventory (Granier et al., 2005). Biogenic emissions of isoprene and monoterpenes are based on the Model Emissions of Gases and Aerosols from Nature (MEGAN) (Guenther et al., 2006). Emissions of NO from soil and fertilizer use are based, respectively, on the Yienger and Levy (1995) and Bouwman et al. (2002) inventories. NO emissions from lightning depend on cloud top height (Price and Rind, 1992; Price et al., 1997) and their vertical distribution is determined according to Pickering et al. (1998) and DeCaria et al. (2005). Most anthropogenic emissions are taken from the POET inventory. Other anthropogenic emissions are from Bond et al. (2004) (for black and organic carbon) and EDGAR databases (<http://www.mnp.nl/edgar/>) (for SO₂ and NH₃). Anthropogenic emissions over Asia are taken from the Regional Emission inventory for Asia (REAS) (Ohara et al., 2007). Monthly average biomass burning emissions for 2004 were taken from the GFEDv2 database (Van der Werf et al., 2006). Emmons et al. (2010) provide a more detailed description of chemical and physical processes in the MOZART-4 version used in this study.

4. Results

Figure 2 shows a comparison of model results interpolated along the P3 flight tracks against CO, NO_x, HNO₃ and PAN measurements on 20, 21 and 22 July. High CO concentrations (more than 350 ppbv) were measured below 2 km on 20 July. Enhanced NO_x (more than 20 ppbv), HNO₃ (more than 20 ppbv) and PAN (more than 2 ppbv) were also measured at the same altitudes. The measured values decreased on the next two days but remained high on 21 July. The model is in good agreement with the observations above 3 km but fails to reproduce the enhanced CO, NO_x, HNO₃ and PAN concentrations seen below 2 km. The concentrations measured near the surface (below 300 m) are overestimated. Comparison with average CO measurements from MOPITT for July 2004 at 800 hPa (Figure 3) also shows that modeled CO is underestimated by more than 50 % over the north and north-east US during July 2004. The discrepancy with the P3 and MOPITT observations could therefore be due to deficiencies in exchange of surface emissions between the boundary layer (BL) and the free troposphere in the model. Comparisons for acetone, acetaldehyde and isoprene for 20 July are presented in Figure 4. As in the case of CO and other species presented above, the observed acetone, acetaldehyde and isoprene profiles also show high values in the lower troposphere. The model fails to reproduce the observed enhancements below 2 km and overestimates the concentrations near the surface layer. Comparisons for other VOCs (for example C₂H₄ and C₂H₆) (not shown) also show a similar bias in model results confirming low export of pollutants from the BL into the free troposphere in the model. This discrepancy could also explain the modeled O₃ values (Figure 5) which are underestimated above the surface layer throughout the lower troposphere. Uncertainties in emissions, in particular NO_x and VOCs, could also influence the performance of the model, although the deficiency in uplift of surface pollutants explains, most likely, the bias in model results. Figure 6 displays the simulated CO and O₃ profiles compared to MOZAIC measurements over the urban area of New York on 20 and 21 July. Consistent with the P3 observations, MOZAIC measurements also show enhanced CO (~ 300 ppbv) and O₃ (~ 90 ppbv) below 2 km due to influence of anthropogenic emissions. As in the comparisons with the P3 measurements, the model fails to reproduce the enhanced CO and O₃ while surface concentrations are overestimated. Comparison with MOZAIC measurements made over other urban areas in North America (e.g. Washington and Montreal) (not shown) also showed similar features in the model profiles: overestimation of surface concentrations and underestimation in other levels up to 2-3 km.

Figure 7 shows the simulated concentrations of CO, NO_x, HNO₃, PAN, C₂H₆ and O₃ and the mean concentrations measured by the P3 on 20, 21 and 22 July in the polluted plume identified by Methven et al. (2006). As discussed previously, the highest values in the plume were observed on 20 July. Measured CO, PAN, C₂H₆ and O₃ values decreased slightly on 21 July (-8%, -24%, 9% and -0.7% respectively). Strongest decreases occurred however in the NO_x and HNO₃ concentrations (-46% and -66% respectively). Concentrations also decreased on 22 July, compared to the previous days, but increased for some VOCs (e.g. C₂H₆, C₂H₄ and acetone) (not shown). This could be due to mixing with other air masses influenced by recent emissions as suggested by Real et al. (2008). The model underestimates the measurements in the plume on 20 and 21 July. Measured concentrations of CO, NO_x, PAN, C₂H₆ and O₃ on these two days are underestimated by more than 85 ppbv, 0.25 ppbv, 450 pptv, 600 pptv and 35 ppbv respectively. Measured HNO₃ are severely underestimated by around 13.5 ppbv on 20 July and 4.5 ppbv on 21 July. This suggests a higher wet deposition in the model, in particular on 20 July. Since the simulated NO_x and PAN values are also underestimated, the bias in HNO₃ could also be due to low production from NO_x oxidation. Simulated values on 22 July show an increase in the concentrations, while the chemical evolution in the plume shows that all pollutant concentrations, except some VOCs, decreased compared to the previous two days. These increases in the modeled concentrations on 22 July may be explained by mixing with other polluted air masses, as reported above, linked to the Real et al. (2008) finding concerning uncertainties in plume transport.

Figure 8 shows the chemical evolution of the polluted plume calculated by Real et al. (2008) using a Lagrangian photochemical trajectory model. Their model, initialized with the P3 measurements on 20 July, is in much better agreement with the measurements compared to MOZART-4 and in general reproduces correctly the decreases in the pollutant concentrations during transport. However, despite the underestimation of pollutant concentrations, MOZART-4 also captures the decrease in the observed values between 20 and 21 July. Moreover, as is the case in the observations, modeled NO_x and HNO_3 show large decreases on 21 July (-58%). Model results would be improved if concentrations were simulated correctly in the initial plume on 20 July. Using several sensitivity studies, Real et al. (2008) analyzed the processes influencing the evolution of pollutants during transport at low altitudes from North America. They demonstrated that the maintenance of high O_3 levels in the plume in the first couple of days is mainly due to high HNO_3 photolysis on 20 July leading to NO_x and, thus, to O_3 production. Underestimation of HNO_3 in the model could therefore also explain the underestimation of O_3 and other species, such as NO_x , in addition to the bias in export of pollutants from the BL into the free troposphere reported previously. Real et al. (2008) concluded that the evolution of pollutant plumes when crossing the Atlantic towards Europe at low altitudes is mainly driven by photochemical and deposition processes, in particular wet deposition of HNO_3 and other soluble species. Underestimation of concentrations in such plumes over North America may lead therefore to underestimation of O_3 and other pollutants exported to Europe. Thus, the discrepancies in MOZART-4 discussed here may lead to an underestimation in the export of O_3 to the global free troposphere and downwind continents.

5. Conclusions

Results from the chemistry transport model MOZART-4 have been compared to in-situ measurements from the ICARTT campaign in order to assess the ability of the model to reproduce the anthropogenic pollution observed in the lower troposphere over North America. During this campaign, which took place in summer 2004, strongly polluted air masses with high values of CO , O_3 , HNO_3 , NO_x and VOCs were sampled by the P3 aircraft in the lower troposphere over the US northeast coast. The comparison of MOZART-4 results against observations shows that the model is not able to reproduce correctly the high pollutant concentrations measured in the lower troposphere due to deficiencies in exchange of pollutants between the BL and the free troposphere and uncertainties in wet deposition of soluble species, in particular HNO_3 , although uncertainties in emissions, in particular NO_x and VOCs, could also influence the model performance. Due to these discrepancies the model also underestimates the pollutant concentrations sampled on 20 and 21 July in the polluted plume identified by Methven et al. (2006) as a case of anthropogenic plume transported at low altitudes from North America toward Europe across the North Atlantic. The model does capture the decrease in pollutant concentrations in the plume between 20 and 21 July but severely underestimates measured concentrations in particular on 20 July. The Lagrangian photochemical trajectory model used by Real et al. (2008), which was initialized by the P3 measurements on 20 July, is able to reproduce the evolution of the chemical composition in the plume by taking into account mixing between the plume and its environment and introducing a more realistic treatment of wet deposition in the model. Therefore, MOZART-4 results could be in better agreement with the observations if the measured concentrations on 20 July are correctly reproduced, in particular HNO_3 which was severely underestimated due, probably, to strong wet deposition in the model. Furthermore, Real et al. (2008) attributed the high O_3 values measured in the plume in the first two days to the high HNO_3 levels on 20 July which produced NO_x and, thus, O_3 through photolysis. While pollutant values decreased also on 22 July, modeled concentrations increased compared to the previous days. These increases suggest the mixing of the initial plume with air masses influenced by more recent

emissions is overestimated in the global model. This is consistent with Real et al. (2008) who showed uncertainties in the calculated plume trajectories on 22 July. Therefore, uncertainties in MOZART-4 parameterizations regarding wet deposition and exchange between the BL and the free troposphere, in addition to uncertainties in emissions, result in underestimation of initial pollutant values in North American anthropogenic plumes and, thus, in underestimation of O₃ and other pollutants export into Europe.

References

- Bond, T. C., D. G. Streets, K. F. Yarber, S. M. Nelson, J. - H. Woo, and Z. Klimont, A technology - based global inventory of black and organic carbon emissions from combustion, *J. Geophys. Res.*, 109, D14203, doi:10.1029/2003JD003697, 2004.
- Bouwman, A. F., L. J. M. Boumans, and N. H. Batjes, Modeling global annual N₂O and NO emissions from fertilized fields, *Global Biogeochem.Cycles*, 16, 1080, doi:10.1029/2001 GB001812, 2002.
- DeCaria, A. J., K. E. Pickering, G. L. Stenchikov, and L. E. Ott, Lightning - generated NOX and its impact on tropospheric ozone production: A three - dimensional modeling study of a Stratosphere - Troposphere Experiment: Radiation, Aerosols and Ozone (STERA0 - A) thunderstorm, *J. Geophys. Res.*, 110, D14303, doi:10.1029/2004JD005556, 2005.
- Deeter M. N., L. K. Emmons, D. P. Edwards, J. C. Gille, and J. R. Drummond, Vertical resolution and information content of CO profiles retrieved by MOPITT, *Geophys. Res. Lett.*, 31, L15112, doi:10.1029/2004GL020235, 2004.
- Emmons, L. K., S. Walters, P. G. Hess, J.-F. Lamarque, G. G. Pfister, D. Fillmore, C. Granier, A. Guenther, D. Kinnison, T. Laepple, J. Orlando, X. Tie, G. Tyndall, C. Wiedinmyer, S. L. Baughcum, and S. Kloster, *Geosci. Model Dev.*, 3, 43-67, 2010.
- Fehsenfeld, F. C., G. Ancellet, T. S. Bates, A. H. Goldstein, R. M. Hardesty, R. Honrath, K. S. Law, A. C. Lewis, R. Leaitch, S. McKeen, J. Meagher, D. D. Parrish, A. A. P. Pszenny, P. B. Russell, H. Schlager, J. Seinfeld, R. Talbot, R. Zbinden, International Consortium for Atmospheric Research on Transport and Transformation (ICARTT): North America to Europe: Overview of the 2004 summer field study, *J. Geophys. Res.*, 111, D23S01, doi:10.1029/2006JD007829, 2006.
- Granier, C., A. Guenther, J. Lamarque, A. Mieville, J. Muller, J. Olivier, J. Orlando, G. Petron, G. Thyndall and S. Wallens, POET, a database of surface emissions of ozone precursors, available at: <http://accent.aero.jussieu.fr/contact.php>, 2005.
- Guenther, A., T. Karl, P. Harley, C. Wiedinmyer, P. I. Palmer, and C. Geron, Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), *Atmos. Chem. Phys.*, 6, 3181-3210, 2006.
- Kalnay, E., M. Kanamitsu, R. Kistler, et al.: The NCEP/NCAR 40 Year Reanalysis Project, *B. Am. Meteorol. Soc.*, 77, 437-471, 1996.
- Li, Q., D. J. Jacob, R. Park, Y. Wang, C. Heald, R. Hudman, R. Yantosca, R. Martin, and M. Evans, Transatlantic transport of pollution and its effects on surface ozone in Europe and North America, *J. Geophys. Res.*, 107, D13, doi:10.1029/2001JD001422, 2002.

- Marenco, A., V. Thouret, P. Nédélec, H. Smit, M. Helten, D. Kley, F. Karcher, P. Simon, K. Law, J. Pyle, G. Poschmann, R. Von Wrede, C. Hume, T. Cook, Measurement of ozone and water vapor by Airbus in - service aircraft: The MOZAIC airborne program, An overview, *J. Geophys. Res.*, 103, D19, doi:10.1029/98JD00977, 1998.
- Methven, J., S. R. Arnold, A. Stohl, M. J. Evans, M. Avery, K. Law, A. C. Lewis, P. S. Monks, D. D. Parrish, C. E. Reeves, H. Schlager, E. Atlas, D. R. Blake, H. Coe, J. Crosier, F. M. Flocke, J. S. Holloway, J. R. Hopkins, J. McQuaid, R. Purvis, R. Rappenglück, H. B. Singh, N. M. Watson, L. K. Whalley, P. I. Williams, Establishing Lagrangian connections between observations within air masses crossing the Atlantic during the International Consortium for Atmospheric Research on Transport and Transformation experiment, *J. Geophys. Res.*, 111, D23S62, doi:10.1029/2006JD007540, 2006.
- Ohara, T., H. Akimoto, J. Kurokawa, N. Horii, K. Yamaji, X. Yan, and T. Hayasaka, An Asian emission inventory of anthropogenic emission sources for the period 1980–2020, *Atmos. Chem. Phys.*, 7, 4419-4444, 2007.
- Owen, R. C., O. R. Cooper, A. Stohl, and R. E. Honrath, An analysis of the mechanisms of North American pollutant transport to the central North Atlantic lower free troposphere, *J. Geophys. Res.*, 111, D23S58, doi:10.1029/2006JD007062, 2006.
- Pickering, K., Y. Wang, W. - K. Tao, C. Price, and J. - F. Müllern, Vertical distributions of lightning NO_x for use in regional and global chemical transport models, *J. Geophys. Res.*, 103, D23, doi:10.1029/98JD02651, 1998.
- Price, C. and Rind, D., A Simple Lightning Parameterization for Calculating Global Lightning Distributions, *J. Geophys. Res.*, 97, D9, doi:10.1029/92JD00719, 1992.
- Price, C., J. Penner, and M. Prather, NO_x from lightning 1. Global distribution based on lightning physics, *J. Geophys. Res.*, 102, D5, doi:10.1029/96JD03504, 1997.
- Real, E., K. S. Law, H. Schlager, A. Roiger, H. Huntrieser, J. Methven, M. Cain, J. Holloway, J. A. Neuman, T. Ryerson, F. Flocke, J. de Gouw, E. Atlas, S. Donnelly, and D. Parrish, *Atmos. Chem. Phys.*, 8, 7737-7754, 2008.
- Stohl, A, S. Eckhardt, C. Forster, P. James, and N. Spichtinger, On the pathways and timescales of intercontinental air pollution transport, *J. Geophys. Res.*, 107, 4684, doi:10.1029/2001JD001396, 2002.
- van der Werf, G. R., J. T. Randerson, L. Giglio, G. J. Collatz, P. S. Kasibhatla, and A. F. Arellano Jr., Interannual variability in global biomass burning emissions from 1997 to 2004, *Atmos. Chem. Phys.*, 6, 3423-3441, 2006.
- Yienger, J. and Levy, H., Empirical model of global soil - biogenic NO_x emissions, *J. Geophys. Res.*, 100, D6, doi:10.1029/95JD00370, 1995.

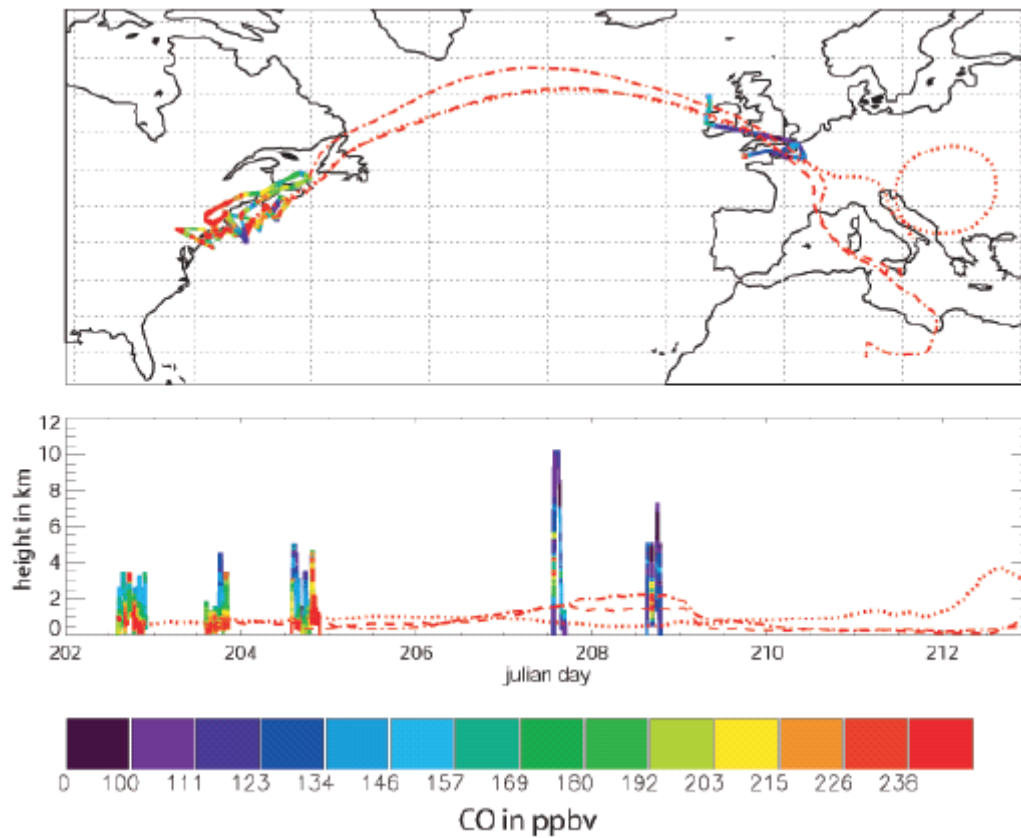


Fig. 1: Flight tracks of P3 aircraft on 20, 21, and 22 July 2004 and Falcon flights on 25 and 26 July 2004, coloured with measured CO. Three forward trajectories initialised in the 3 Lagrangian samplings of the P3 are shown for 20 (dotted), 21 (dashed) and (dashed-dotted) July 2004, respectively. Figure from Real et al. (2008).

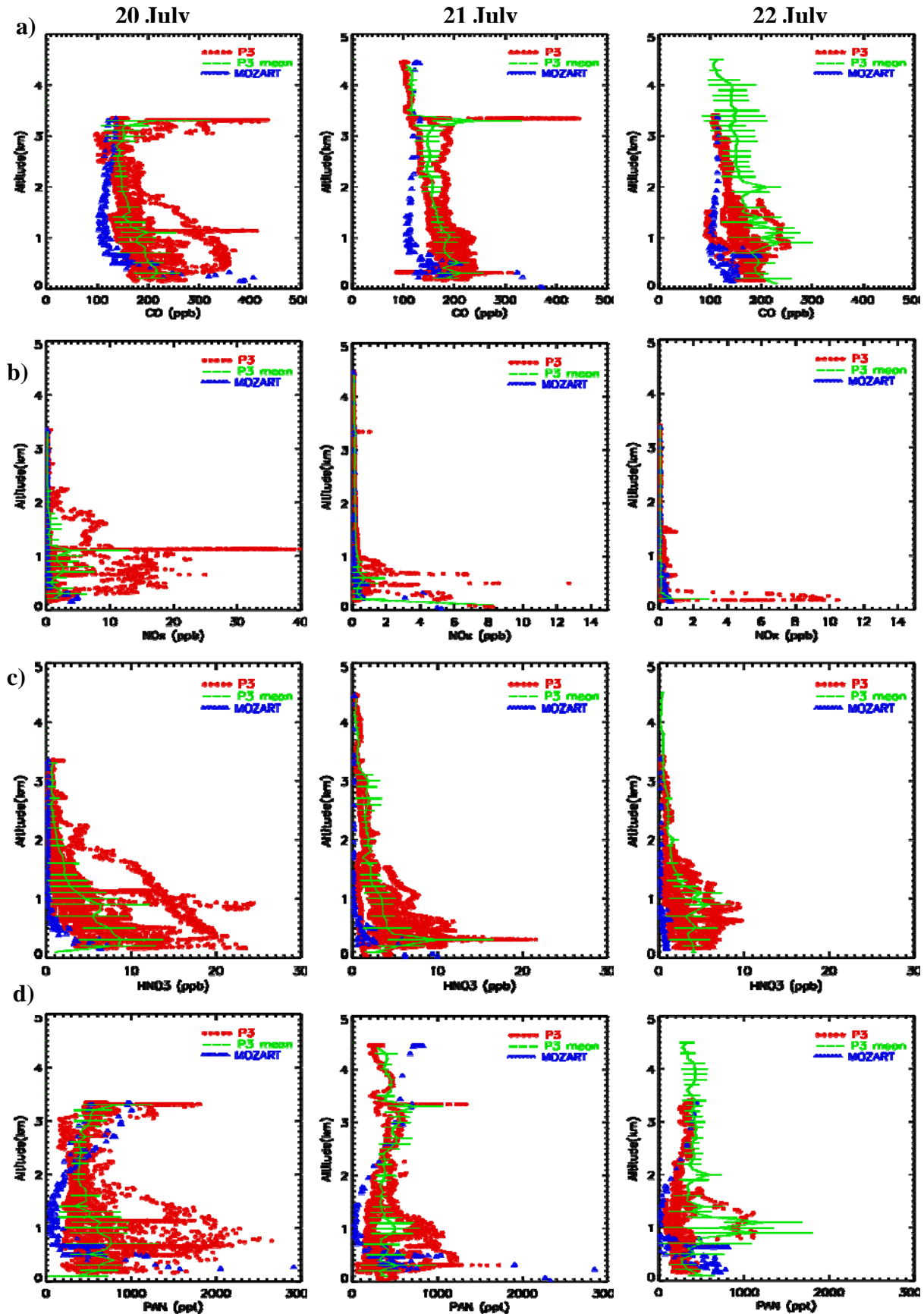


Fig. 2: Comparison of simulated (blue) and observed (red) vertical profiles of CO (a), NO_x (b), HNO₃ (c) and PAN (d) by the P3 aircraft on 20, 21 and 22 July 2004. The green profiles represent the mean observed values (with horizontal bars for standard deviations).

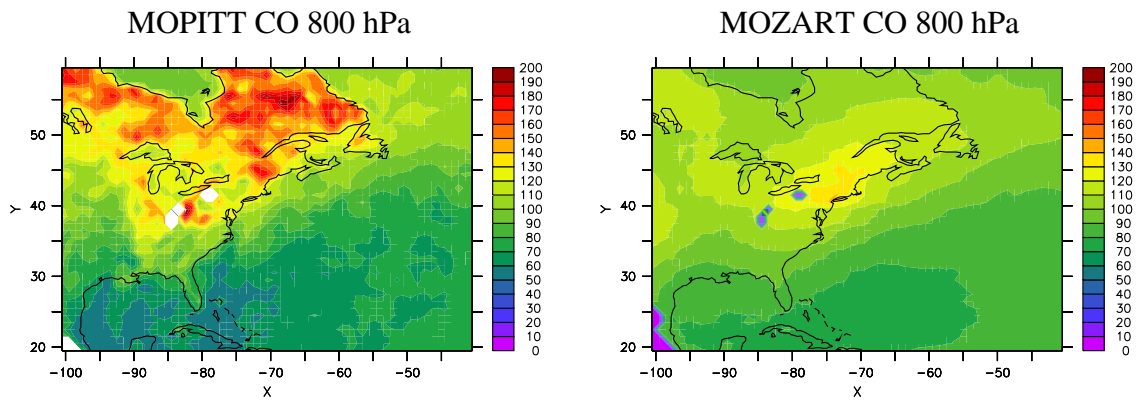


Fig. 3: Comparison between MOPITT and MOZART-4 CO (ppbv) at 800 hPa averaged for July 2004.

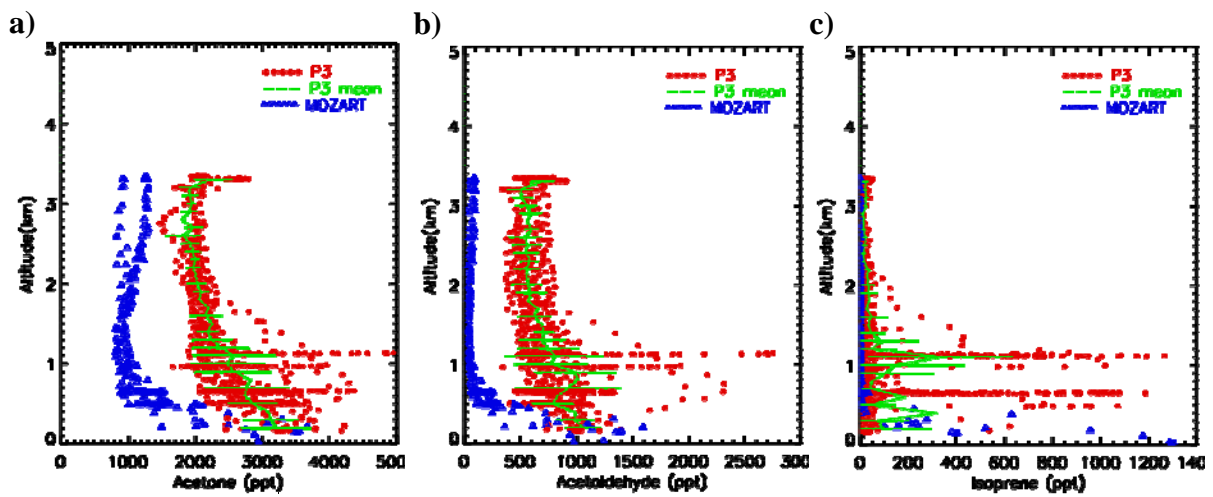


Fig. 4: Same as Figure 2 for acetone (a), acetaldehyde (b) and isoprene (c) in pptv on 20 July 2004.

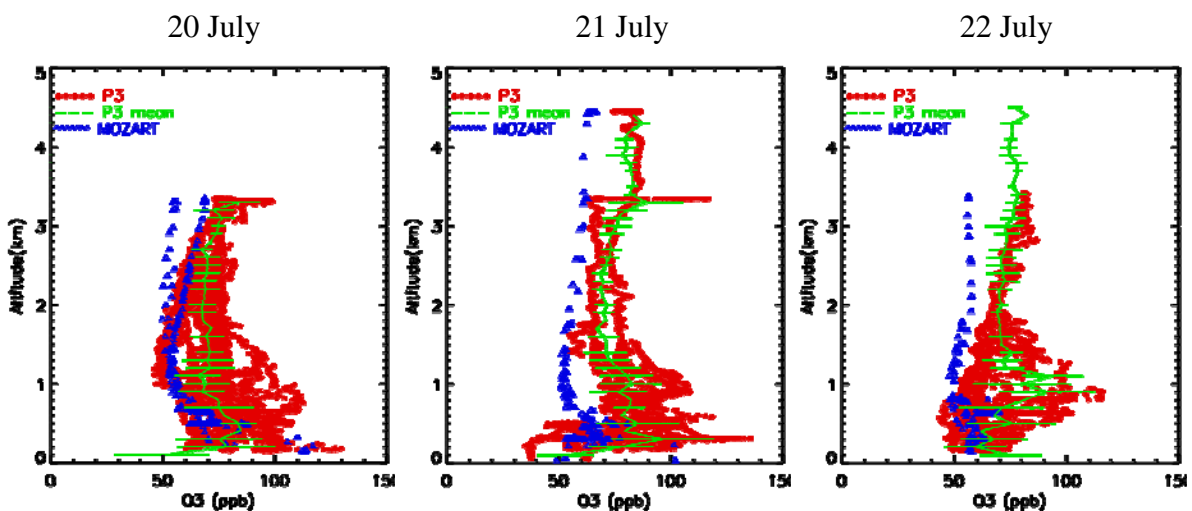


Fig. 5: Same as Figure 2 for O₃ (ppbv).

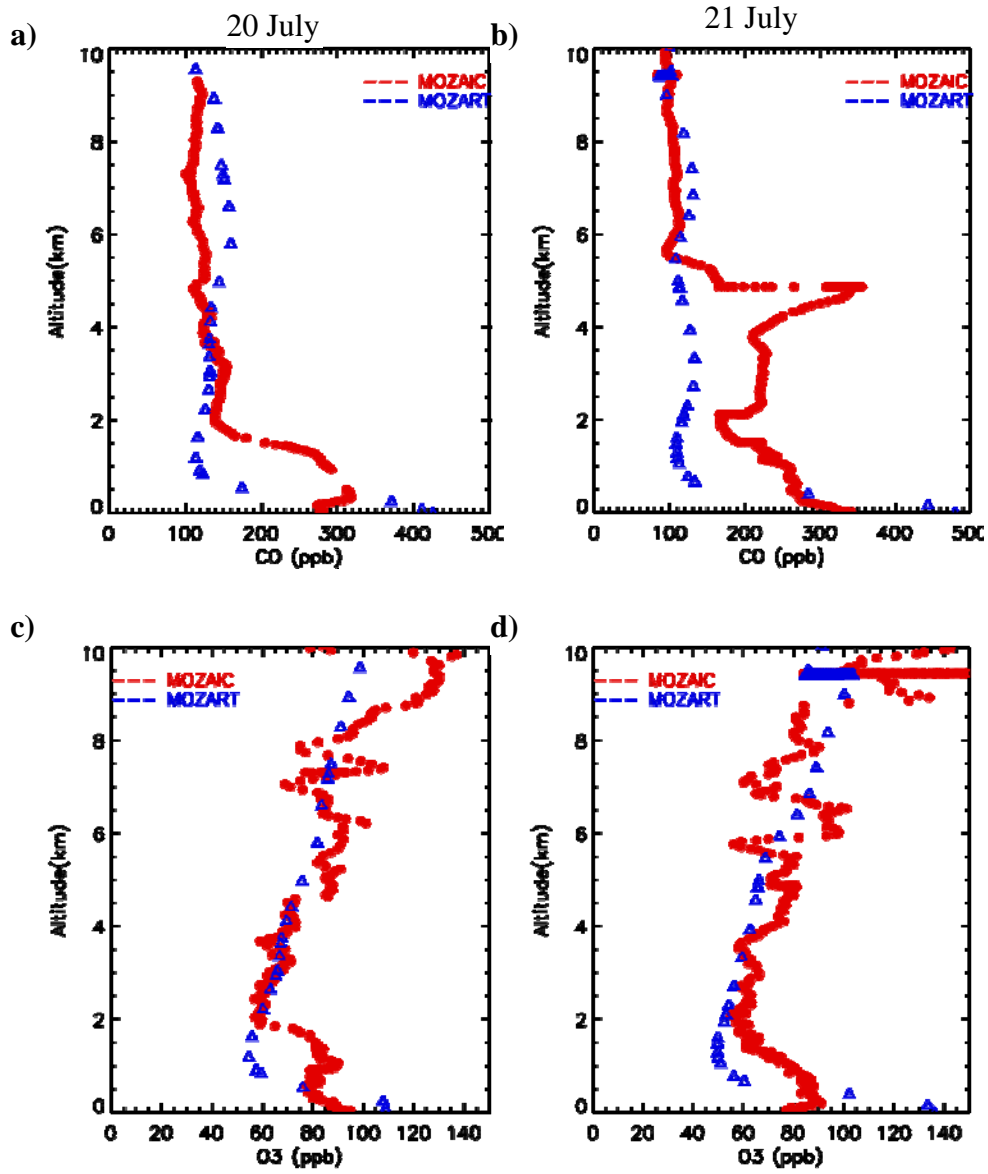


Fig. 6: Comparison of simulated (blue) and observed (red) CO (a, b) and O₃ (c, d) vertical profiles by MOZAIC aircraft over New York on 20 (left) and 21 (right) July 2004.

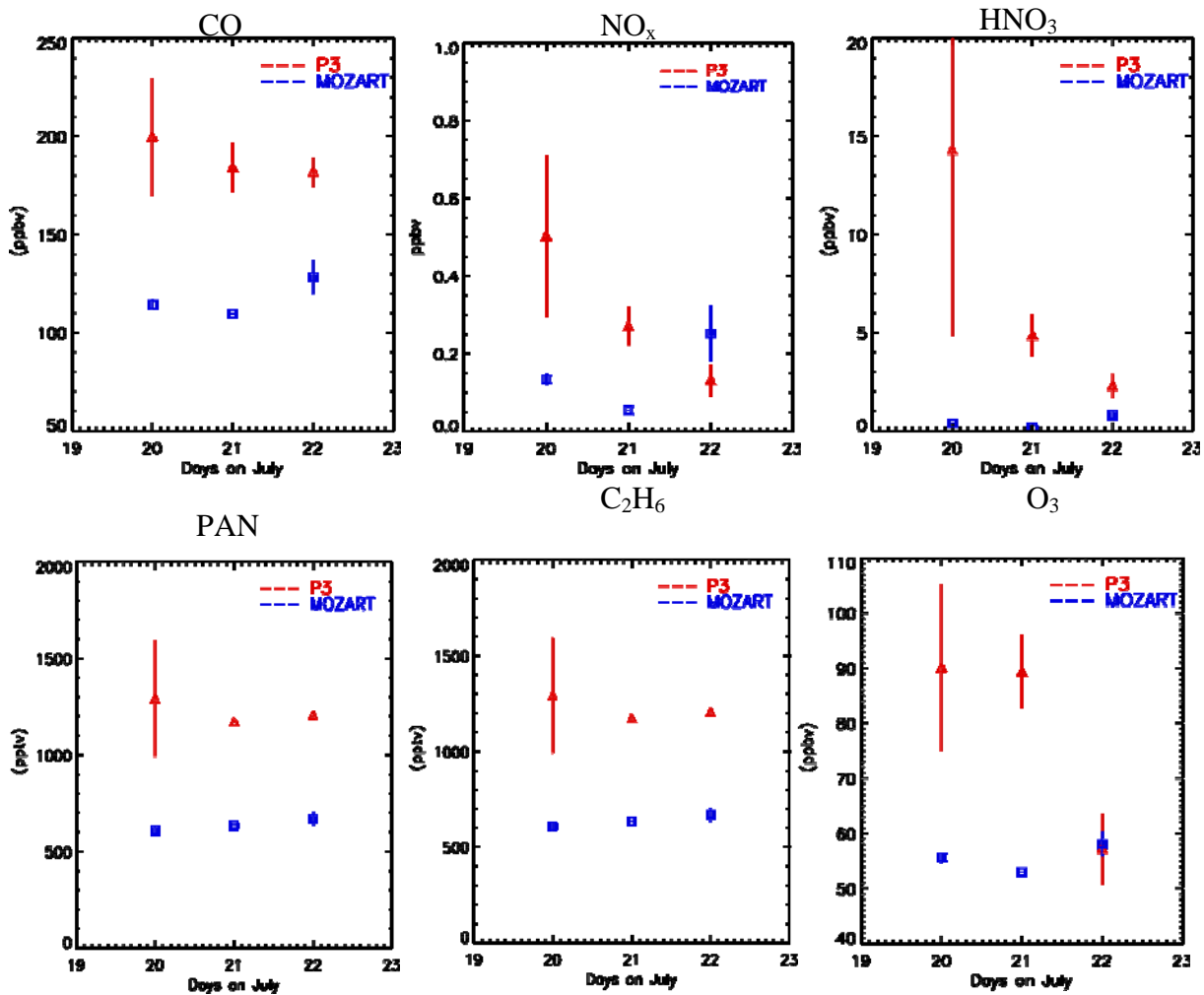


Fig. 7: Comparison of temporal evolution of average simulated (blue) and mean P3 (red) concentrations of CO, NO_x, HNO₃, PAN, C₂H₆ and O₃ between 20 and 22 July 2004. Observed and simulated concentrations were averaged over the time segments corresponding to the time when the P3 anthropogenic plume identified by Methven et al. (2006) was observed on each day. Standard deviations are represented as vertical bars.

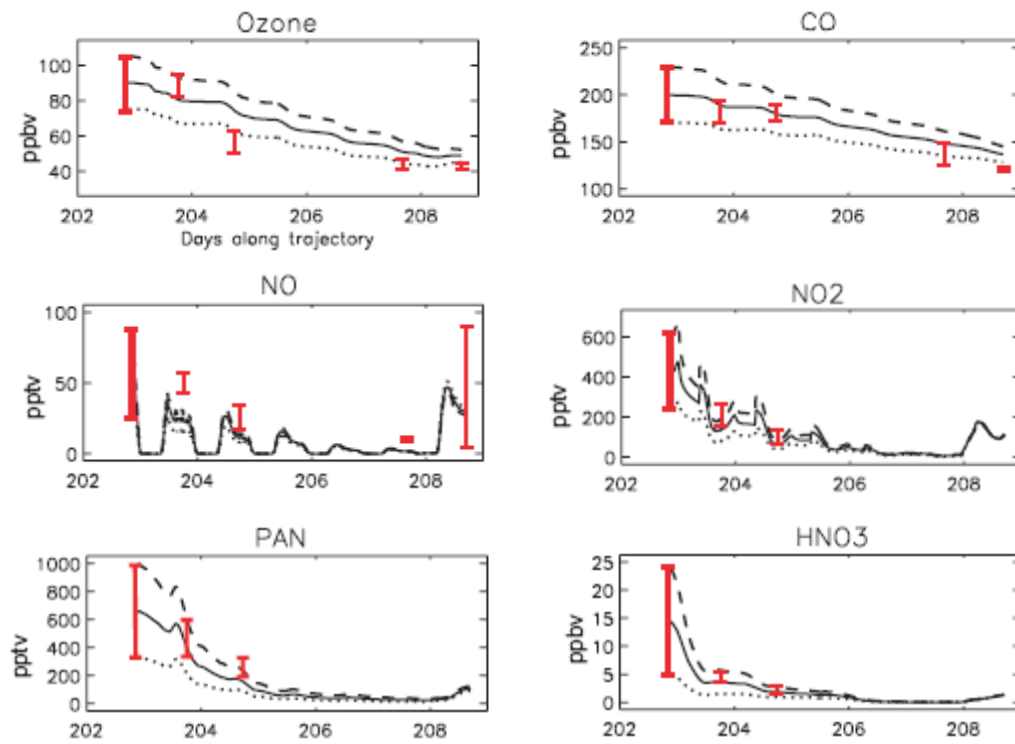


Fig. 8: Same as Figure 7 but for results of the photochemical trajectory model used by Real et al. (2008). The temporal evolution is shown as Julian days corresponding to the period from 19 July to 26 July 2006. Mean P3 observations are represented in red (with standard deviations as vertical bars) and mean modelled concentrations are shown in black continuous lines. Dotted (and dashed) lines correspond to model results obtained when it was initialised with mean + (and -) standard deviation concentrations taken from the P3 plume on 20 July 2004. Figure from Real et al. (2008).