Geosci. Model Dev. Discuss., 5, 2083–2138, 2012
www.geosci-model-dev-discuss.net/5/2083/2012/
doi:10.5194/gmdd-5-2083-2012

Geoscientific Model Development Discussions

How realistic are air quality hindcasts driven by forcings from climate model simulations?

G. Lacressonnière[1], V.-H. Peuch[2], J. Arteta[1], B. Josse[1], M. Joly[1], V. Marécal[1], D. Saint Martin[1], M. Déqué[1], and L. Watson[1]

[1]GAME/CNRM, Météo-France, URA1357, CNRS – Centre National de Recherches Météorologiques, 42 av. G.Coriolis, 31057 Toulouse, France
[2]ECMWF, Shinfield Park, Reading, UK

Received: 12 July 2012 – Accepted: 17 July 2012 – Published: 31 July 2012

Correspondence to: G. Lacressonnière (gwendoline.lacressonniere@meteo.fr)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Predicting how European air quality could evolve over the next decades in the context of changing climate requires the use of climate models to produce results that can be averaged in a climatologically and statistically sound manner. This is a very different approach from the one that is generally used for air quality hindcasts for the present period: analysed meteorological fields are used to represent specifically each date and hour. Differences arise both from the fact that a climate model run is a pure model output, with no influence from observations (which are useful to correct for a range of errors), and that in a "climate" set-up, simulations on a given day, month or even season cannot be related to any specific period of time (but can just be interpreted in a climatological sense). Hence, although an air quality model can be thoroughly validated in a "realistic" set-up using analysed meteorological fields, the question remains of how far its outputs can be interpreted in a "climate" set-up. For this purpose, we focus on Europe and on the current decade using three 6-yr simulations performed with the multiscale chemistry-transport model MOCAGE and use meteorological forcings either from operational meteorological analyses or from climate simulations. We investigate how statistical skill indicators compare in the different simulations, discriminating also the effects of meteorology on atmospheric fields (winds, temperature, humidity, pressure ...) and on the dependent emissions and deposition processes (volatile organic compound emissions, deposition velocities ...). Our results show in particular how differing boundary layer heights and deposition velocities affect horizontal and vertical distributions of species. When the model is driven by operational analyses, the simulation accurately reproduces the observed values of O_3, NO_x, SO_2 and, with some bias that can be explained by the set-up, PM_{10}. We study how the simulations driven by climate forcings differ, both due to the realism of the forcings (lack of data assimilated and lower resolution) and due to the lack of representation of the actual chronology of events. We conclude that the indicators such as mean bias, mean normalized bias, RMSE and deviation standards can be used to interpret the results with some

GMDD

5, 2083–2138, 2012

confidence as well as the health-related indicators such as SOMO35 and the number of days of exceedance of regulatory thresholds. These metrics are thus considered to be suitable for the interpretation of simulations of the future evolution of European air quality.

1 Introduction

The issues of climate change and air quality are intertwined: anthropogenic emissions contribute to climate change, and the evolution of the climate through changes in meteorological parameters (temperature, precipitation) impacts concentrations and distributions of pollutants in the atmosphere. In the lower troposphere, ozone (O_3) is a pollutant that affects human health (WHO, 2004; Schlink et al., 2006) and causes damages to crops (Fuhrer and Booker, 2003) and ecosystems. O_3 is a secondary pollutant; its principal precursors are carbon monoxide (CO), volatile organic compounds (VOCs) and nitrogen oxides (NO_x) emitted by both natural (biogenic) and anthropogenic (transport, industries) sources. Several studies have shown how O_3 photochemistry depends upon meteorological conditions such as temperature and precipitation. During summer, conditions of high temperature and low precipitation favour oxidant accumulation and surface concentrations of O_3 reach high values (Guicherit and van Dop, 1977; Sillman, 2000) and have the potential to exceed air quality standards. These conditions also favor the production of secondary pollutants such as sulphate and nitrate aerosols, and organic aerosols which can contribute to the high levels of particulate matter (PM) during summertime. Nevertheless, the frequency and intensity of pollution episodes vary considerably from year to year depending on weather: as an example the summer 2003 heat wave in Europe has been associated with exceptionally high O_3 concentrations (Langner et al., 2005; Vautard et al., 2005, 2007; Guerova and Jones, 2007; Solberg et al., 2008). In fall and winter, stagnant conditions also enhance levels of primary pollutants (SO_2, NO_x) in the atmosphere and thus concentrations of PM_{10} (particles with

GMDD
5, 2083–2138, 2012

an aerodynamic diameter smaller than 10 µm): another pollutant of concern connected to air quality and health problems.

The interactions between climate change and surface air quality have been already extensively studied. On a global scale, studies (Prather et al., 2003; Dentener, 2006) have for instance evaluated surface O_3 concentrations under an A2 scenario (IPCC AR4). Dentener (2006) showed that global mean surface O_3 may increase by about 4.3 ± 2.2 ppbv by the year 2030 and the area of global natural ecosystems exposed to critical nitrogen deposition may increase up to 25 % by 2030. Regional models centered over the continental United States have been used to examine US air quality in the future (Hogrefe et al., 2004; Knowlton and al., 2004; Dawson et al., 2009). Hogrefe et al. (2004) concluded that the average daily summertime maximum 8-h O_3 concentrations will increase by 2.7 ppbv and 4.2 ppbv for summers in the 2020 s and 2050 s, respectively. In the literature, a set of regional models have focused on the European region (Langner et al., 2005; Szopa et al., 2006; Meleux et al., 2007; Giorgi and Meleux, 2007). Zlatev (2007) and Langner et al. (2005) presented the impacts of climate change on air quality over Europe with a constant emission rate and showed an increase in photochemical production in future climate scenarios. In Meleux et al. (2007), the authors isolated the impacts of summer European climate change on the increase in O_3 levels by using the same emissions and global chemical boundary conditions for the present day and future periods. Carvalho et al. (2010) concluded that PM_{10} levels will be impacted by climate change depending on the month and region, with a maximum increase reaching 30 µg m^{-3} in September over Portugal. Szopa et al. (2006) also estimated that the O_3 concentration in July may increase up to 5 ppb across Europe by 2030. According to all these findings, the expectation of a warmer climate in the next decade may well affect air quality directly despite the regulations to reduce the emissions of pollutants (European Commission, 2008).

In most "climate" studies, air quality modeling systems are chemistry-transport models (CTM) that rely on global or regional climate models to provide the meteorological forcings for future periods. The purpose of this paper is to assess how realistic

GMDD
5, 2083–2138, 2012

air quality hindcasts are when driven by forcings from climate models for the current period over Europe, in comparison to a reference obtained using instead analysed meteorological forcings (analyses), which include meteorological observations and are thus very realistic and specific for each single date. The results will be evaluated using a range of statistical tools and air quality indicators. In our study, we have used the chemistry and transport model MOCAGE of Météo-France (Peuch et al., 1999) for three multi-year simulations covering the present time (2003–2008) over a European domain. Comparisons between the simulations and the AirBase observations allow us to infer how a range of statistical indicators are affected when using different types of forcings. This work will provide guidance on how far to interpret air quality hindcasts relying on climate model outputs, which is essential for the study of future air quality.

In this paper, Sect. 2 describes the modeling approaches and the numerical experiment design. We also discuss the statistical indicators and the representativeness of the measurement stations used for this study. Section 3 compares simulations in order to evaluate separately how emissions and meteorological changes affect the distributions of pollutants over Europe. Finally, the experiments run with the analyses and the climate forcings are compared against observations in Sect. 4. The statistical indicators that produce similar results for the two experiments will be the most useful ones to consider when examining future trends.

2 Methodology

2.1 Model set-up and experimental design

The model used in this study is the three-dimensional multi-scale chemistry-transport MOCAGE (Modèle de Chimie Atmosphérique à Grande Echelle), which simulates the interactions between the dynamical, physical and chemical processes in the troposphere and stratosphere (Peuch et al., 1999). The configuration allows for the representation of both long range transport of pollutants and regional impacts of pollutants

GMDD
5, 2083–2138, 2012

on air quality. This model is used for operational air quality forecasting in France (http://www.prevair.org, Honoré et al., 2008) and in the context of the GMES atmospheric monitoring service (Hollingsworth et al., 2008) and has been evaluated during several campaigns, see for instance: Dufour et al. (2004); Bousserez et al. (2007).

MOCAGE uses a semi lagrangian advection scheme (Williamson and Rasch, 1989) to transport chemical species. On the vertical, the configuration has 47 hybrid levels from the surface up to 5 hPa with a resolution of about 150 m in the lower troposphere increasing to 800 m in the higher troposphere. Turbulent diffusion is parameterized with the scheme of Louis (1979) and convective processes with the scheme of Bechtold et al. (2001). The chemical scheme used in this study is RACMOBUS: it is a combination of the stratospheric scheme REPROBUS (Lefèvre et al., 1994) and the tropospheric scheme RACM (Stockwell et al., 1997). Overall, this chemical scheme includes 119 individual species with 94 prognostic variables and 377 chemical reactions. In our study, a sulfur cycle has been implemented; the oxidation reactions in the gaseous and aqueous phases lead to the formation of sulphate aerosols as in Ménégoz et al. (2012). MOCAGE simulates the evolution of five types of aerosols: black carbon, sea salts, desert dusts, anthropogenic primary particulate matter and sulphates. They are compartmented in size bins (Martet et al., 2009) and divided into 6 bins for each aerosol compound: between 0.1 µm and 100 µm for dust aerosols, 0.001 µm and 10 µm for black carbon, 0.03 µm and 20 µm for sea salt, 0.005 µm and 10 µm for anthropogenic particulate matter and 0.01 µm and 20 µm for sulphates. Nitrate and organic aerosols are not taken into account in this study. A negative bias is thus expected by design on total PM.

The model uses two-way nested domains, on a $2^\circ \times 2^\circ$ horizontal grid over the globe and a $0.2^\circ \times 0.2^\circ$ horizontal grid over Europe (30° N/70° N; 15° W/35° E). Three-hourly forcings are used for meteorology in this study, from either operational analyses from Météo-France (ARPEGE, Courtier et al., 1991) or from climate simulations obtained with ARPEGE-Climate, (version 5.1, Déqué et al., 1994), for the present decade. The resolution of ARPEGE is in a T798 spectral stretched grid (resolution of around 15 km

GMDD
5, 2083–2138, 2012

over Europe and 60 km in the Pacific) while ARPEGE-Climate operates on a T63 triangular truncation, equivalent to a resolution of about 2.8°. Anthropogenic forcings of ARPEGE-Climate (GHG, aerosols) refer to the climatology of the present time while the SSTs (Sea Surface Temperatures) evolve along the simulation. Meteorological forcings are interpolated horizontaly on the two MOCAGE domains.

For the anthropogenic emissions, the inventory is the one (Visschedijk and Denier van der Gon, 2005; Pouliot et al., 2012) developed for the Global and regional Earth-system Monitoring using Satellite and in-situ data (GMES) project (Hollingsworth et al., 2008). This inventory has a high spatial resolution compatible with our model and a temporal resolution of 1 h. It is representative of the year 2003. We chose not to modify emissions depending on each specific year as this would be meaningless for runs driven by climate forcings. Biogenic emissions of isoprene and monoterpene are calculated offline with MEGANv2.04 model (Guenther et al., 2006). Two types of input files were required: the landcover variables (leaf area index, plant functional type and emissions factors) and the weather data. The landcover variables are available for spatial resolution of ∼ 5 km (150 s longitude × 150 s latitude). The meteorological fields (temperature, solar radiation) are provided either by ARPEGE analyses or ARPEGE-Climate simulations.

The numerical experiment design is as follows (Table 1). Three six-year periods of the current climate were simulated using different meteorological forcings and surface processes. ANALY simulation acts as the reference. INT is relied upon meteorological forcings from a climate model and surface exchanges (weather-dependent emissions and deposition velocities) are the same as ANALY. CLIM is driven by climate forcings and surface processes are computed with meteorological conditions of ARPEGE-Climate. The summer heat wave of 2003 was uncharacteristic of the current climate and studies have shown a similar pattern of heatwaves with future climate conditions (Meleux et al., 2007; Solberg et al., 2008). The climatological forcings from ARPEGE-Climate are representative of the current decade and do not reproduce the extremely hot and cold events. For this reason, year 2003 is not considered in the statistical

GMDD
5, 2083–2138, 2012

comparisons. We chose the 2004–2008 period for our simulations; 5 yr is a short time on the one hand to represent the meteorological variability over Europe. On the other hand, we need that emissions do not evolve too much in time during the period and 5 yr is certainly at the limit for such an assumption. The choice of 5 yr is thus clearly a compromise.

2.2 Statistical indicators

In Europe, air quality thresholds of acceptable levels of O_3 (European Commission, 2002), NO_2 and SO_2 (European Commission, 1999, 2001) and $PM_{2.5}$ and PM_{10} (European Commission, 1999, 2001, 2008) have been established in order to protect and inform populations, as described in Table 2. The impact of these air quality policies, in Europe and in the world as well, can be evaluated using numerical air quality modeling. As an example, in Europe, the CAFE (clean air for europe) program has been set up to assess the impacts of these policies on the pollutants' levels (Cuvelier et al., 2007).

In order to forecast air quality, to understand the dynamics of air pollution and to develop regulations to reduce emissions, air quality modeling systems are needed. A variety of metrics has been used over the years to evaluate the performance of air quality models (US-EPA, 1984, 1991; Chang and Hanna, 2004; Boylan and Russell, 2006). Mean bias (MB), mean normalized bias (MNB), root mean square error (RMSE) and correlation coefficient (CORR) are common statistical parameters used by the modeling community. Furthermore, the mean normalized bias error (MNBE) and the mean normalized gross error (MNGE), normalizing the bias and error for each model-observed pair by the observations, are also useful parameters. For the evaluation of particulate matter concentration, Boylan and Russell (2006) suggested the consideration of the mean fractional bias (MFB) and the mean fractional error (MFE) parameters instead of MNBE and MNGE. They proposed that the model performance criteria would be met when both MFE $\leq 75\,\%$ and MFB $\leq \pm 60\,\%$, respectively. The model performance goal would be met when both MFE $\leq 50\,\%$ and MFB $\leq \pm 30\,\%$. The US-EPA suggested several performance criteria for simulated O_3 such as MNBE $\leq \pm 15\,\%$ and

GMDD
5, 2083–2138, 2012

MNGE $\leq \pm 35\,\%$ (US-EPA, 1991) while the EC proposes a modeling quality objective given as a relative uncertainty (%): 50 % and 30 % for $PM_{10}/PM_{2.5}/O_3$ and NO_2/SO_2 annual average, respectively (European Commission, 2008).

The model to data statistics MB, MNB, RMSE, correlation coefficient and sigma ratio are selected for the present study. The definitions of these metrics are indicated in Table 3. We also considered the mean diurnal cycle and the temporal series. Mean diurnal cycles are averaged over all available days of concentrations for each 24 h period, while time series are based on the daily mean. Seasonal mean statistics are computed, with seasons corresponding to summer (June, July, August and September) and winter (December, January, February and March). We chose to study these two seasons of interest in air pollution while autumn and spring are rather transitional seasons. As summarized in Table 4 metrics are calculated for hourly values and daily averages for NO_x, SO_2 and PM_{10} while the hourly value and daily maximum 8-h average concentrations (M×8h) statistics are computed for O_3 as the M×8h is one of the most important parameters to be considered for this species.

2.3 Observations and representativeness

In order to evaluate the performance of MOCAGE and to be in position to investigate and put into context the differences between the simulations, we used AirBase (Version 5) measurement data. The AirBase meta-data describes both the site area (urban, suburban or rural) and the site type (traffic, industrial or background). Giving the spatial resolution of our model ($0.2^\circ \times 0.2^\circ$), not all the reporting sites are representative enough. In Joly and Peuch (2012), an objective classification of the AirBase sites based on past measurements has been proposed in order to overcome issues of lack of homogeneity and erroneous information in the metadata. This classification allows for selection of the monitoring sites that are representative of the spatial resolution of our model. Through 10 classes, the less polluted stations (class 1) are distinguished from the very polluted sites (class 10). The robustness of this approach is obtained with

GMDD
5, 2083–2138, 2012

a pollutant-specific classification, taking into account that transport, chemistry and life time are specific to each pollutant.

In order to highlight the effect of site representativeness, we have compared the summertime (JJAS) average diurnal cycles for classes 1–2 (1 and 2), 1–5, 1–10 and 6–10 over France for O_3 and NO_x (Fig. 1). The median diurnal cycles observed (red) and simulated with MOCAGE (black) are presented for the year 2007. As shown in Fig. 1, the modelled variability of ozone and NO_x is not as pronounced as in observations, whatever the type of sites. During nighttime hours, ozone levels decrease when polluted stations are added to the sample (statistics for classes 1–2 versus for classes 1–10) due to titration by NO. Nevertheless, simulated O_3 values do not decrease as much as the observations for classes 6–10 and 1–10. As expected, a large variation in observed NO_x concentrations is seen when the stations considered are from different class types. Except when considering sites of classes 1–2, the amplitude of NO_x concentrations is generally under-estimated by the model. The amplitude of the median diurnal cycles changed significantly between categories 1–2 and 1–5 both for NO_x and O_3. For classes 6–10 and 1–10, i.e. when most polluted sites are considered, the model does not reproduce the high observed values because, at least in part, of the model's horizontal resolution.

To conclude, the choice of the sites used for verification according to an objective classification allows for an improvement in accounting for representativeness. For NO_x which are short lived species, it is necessary to reduce the sample of sites to classes 1–2 only to focus on sites that are representative enough for the model grid-size. Due to transport effects, 5 classes (1–5) can be used to evaluate simulations for longer-lived species such as O_3 in order to have a larger geographical basis. The spatial distribution of PM_{10} (not shown) has the same behavior as O_3 and the same conclusion can be applied. In conclusion, the performances of MOCAGE will be assessed by comparing simulations against observations at sites of classes 1–5 for O_3 and PM_{10} and of classes 1–2 for NO_x and SO_2 (not shown but same behavior as NO_x). The number of sites finally taken into account for each pollutant and country are summarized in Table 5.

GMDD
5, 2083–2138, 2012

3 Results

In the following section, we discuss MOCAGE's capability to simulate realistic air quality hindcasts when driven by forcings from climate modeling for the current period. We will evaluate statistical tools and air quality indices and compare how they evolve with different set of forcings. Two main parts can be distinguished. First, in Sect. 3.1, comparisons between ANALY and INT, as well as between INT and CLIM, allow us to detect the effects due to the meteorological forcings and to changes in surface exchange fluxes, respectively. As described previously, ARPEGE analyses and ARPEGE-Climate fields differ regarding horizontal resolution. Surface exchanges (emissions that depend upon meteorology, as well as surface deposition) have been computed with two sets of meteorological conditions (ARPEGE and ARPEGE-Climate). The differences are for the biogenic volatile organic compounds, desert dust and sea salt emissions as well as for deposition velocities, which depend on meteorology. Section 3.2 presents statistical skill scores of ANALY, CLIM against AirBase data (for "representative sites" only).

3.1 Comparisons between ANALY, INT and CLIM

3.1.1 Impact of changes in meteorological fields on European air pollution levels

Figure 2 shows the mean differences in surface temperature, precipitation, humidity and planetary boundary layer height (PBL) for the JJAS period between ANALY and INT run. The purpose is to evaluate briefly how climate meteorological forcings differ from the analyses. For the comparisons, we have thus averaged spatially the analyses fields to a horizontal resolution similar to the climate run. Focusing on the temperature first, similar structures are found over Europe and an increasing gradient from the northern to the Mediterranean areas. However, over the northeastern part of the domain, the temperatures simulated by ANALY are locally significantly higher up to 4–5 °C. In contrast, over the Italian and Greek areas, INT displayed higher temperatures.

GMDD
5, 2083–2138, 2012

The lower resolution of INT described previously explains partly the differences observed because ANALY display more structured fields. Orography is smoothed in INT and the fields are representative of the current decade. Concerning the humidity, the spatial pattern is correctly reproduced by ARPEGE-Climate, with higher humidity over the Mediterranean Sea. The PBL is another important field to be considered as it affects the dilution of pollutants in the atmosphere. As depicted in Fig. 2, higher average PBL in ANALY than in INT are seen over Spain and Greece.

In Fig. 3, we represent the average surface concentrations of O_3 (a), isoprene (b), NO_x (c), SO_2 (d) and sulphate aerosols (e) for the summertime 2004–2008 in ANALY and INT. Here, the observed changes in pollutant distributions are only due to differences in the meteorological conditions, as emissions and deposition velocities are identical. The spatial pattern of mean O_3 concentrations is similar for the two simulations over the European domain. The highest concentrations are found in Southern Europe, over the Mediterranean Sea (50–60 ppbv), caused by intense photochemical production of O_3 (EEA, 2005; Vautard et al., 2005). The meteorological fields such as temperature influence the production of O_3 (Meleux et al., 2007; Hedegaard et al., 2008). Fields of change in O_3 present some similarities to the changes in temperature (Fig. 2). As noticed previously, over Spain, Africa and Northern Europe, ANALY outputs higher temperatures up to 4–5 °C and higher O_3 concentrations (+6–8 ppbv). The highest positive temperature differences seen over Europe relate to the highest positive O_3 differences. Other studies have shown that among all the meteorological parameters, the one that causes the greatest impact on ozone was temperature (Dawson et al., 2007). Nevertheless, as explained in Katragkou et al. (2010), other variables such as differences in solar radiation and zonal and meridional winds impact the ozone concentrations. Similar temperatures and O_3 spatial patterns with the opposite sign are observed over the Mediterranean basin (Italy, Greece).

High concentrations of isoprene, in the range of 2.5–3 ppbv, are simulated over North Africa and Greece with the simulation INT. The biogenic emissions calculated to drive ANALY and INT are the same, the differences in simulated isoprene can not

GMDD
5, 2083–2138, 2012

be explained by the isoprene emissions or the temperature fields. A longitudinal cross-section (not shown) at a latitude of 36° N (across Africa) displays higher isoprene concentrations at the surface from 0 to 10° with INT. The accumulation of isoprene near the surface can mainly be explained by the boundary layer, which is less well-mixed in INT than in ANALY.

The simulated distributions of summertime average NO_x concentrations ($NO + NO_2$) show levels around 8–12 ppbv in the Netherlands, Belgium, central and Eastern England and the industrial Po Valley (Fig. 3). In ANALY and INT, higher concentrations of NO_x are also found over major shipping routes (North Sea, Gibraltar) and near emissions sources. Tropospheric columns of NO_x (not shown) are identical in ANALY and INT: the differences seen over Europe at the surface are mainly explained by differing boundary layer mixing in the two experiments.

Simulated SO_2 concentrations over Europe display their highest levels over Spain, Eastern Europe (Poland, Romania, Greece) as well as Belgium and UK. Over Northern Spain (Fig. 3), the concentrations reach up to more than 8 ppbv (due to large plant sources in coastal Spain in the emissions dataset). Similar geographical distributions and ground levels of SO_2 are observed in ANALY and INT. Concentrations of SO_2 lead to the formation of sulphate aerosols. Over the northeastern part of the domain, the levels of sulphate are higher in INT than in ANALY. Figure 2 shows higher precipitation and humidity in ANALY than in INT over this area. These differences imply enhanced transformation of SO_2 into sulphate aerosols but also increased wet deposition: those two contrasting effects can explain the differences seen in sulphate concentrations for the two simulations. Figure 4 represents a latitudinal cross-section of sulphate at the longitude of 30° E, averaged for the summertime period of 2004–2008 (JJAS). From 60° N to 70° N, the vertical extent of the sulphate distribution is lower in INT than in ANALY. The difference in sulphate near the surface is due to differing PBL mixing properties in the two simulations. Tropospheric columns of sulphate (Fig. 4) indeed indicate very similar quantities in the two simulations.

GMDD
5, 2083–2138, 2012

To sum up, meteorological forcings (temperature, humidity, horizontal and vertical winds) differ in ANALY and INT. These differences lead to change in the vertical and horizontal simulated distributions of pollutants. For all pollutants, differences are primarily due to differing PBL mixing heights in the two simulations.

3.1.2 Impact of changes in surface exchanges on European air pollution levels

Comparisons between INT and CLIM indicate the contribution of surface processes on the pollutant level changes. The differences between these simulations are related to the biogenic emissions of isoprene and terpene, the desert dust and sea salt emissions as well as the deposition velocities. In Fig. 5, the spatial distributions of summer isoprene emissions in INT and CLIM are shown. Temperature and solar radiation are key driving variables that regulate emissions of isoprene and other BVOCs (Guenther et al., 1993, 1995, 2006). In accordance with the temperature field differences (Fig. 2), higher levels of isoprene are emitted in INT over central Spain and Central Europe, Scandinavia and the northeastern part of Africa than in CLIM. The changes in isoprene emissions induce corresponding changes in the geographical pattern of isoprene concentrations. Elevated concentrations of isoprene (> 1 ppbv) are observed over Central Europe, Greece and North Africa in INT compared to CLIM (Fig. 7b).

Over Central Europe, O_3 deposition velocities are higher by up to 0.2–0.3 $cm\,s^{-1}$ in INT than in CLIM (Fig. 6). Average nighttime and daytime velocities have been calculated for both INT and CLIM; daytime is considered to be from 8:00 to 16:00 UTC and night-time is considered to be from 20:00 to 04:00 UTC. Over land, similar deposition velocities of about 0.54 and 0.57 $cm\,s^{-1}$ are calculated during daytime in INT and CLIM, respectively. Higher velocities are found during the day as it is known that O_3 deposition velocity has a strong diurnal cycle due to increase in surface resistance at night. The mean deposition fluxes ($\mu g\,m^{-3}\,s^{-1}$) of O_3, NO_x and SO_2 have been computed for the summertime period (Fig. 6). In Fig. 7, the changes in concentrations between INT and CLIM follow the changes in deposition fluxes and velocities. Where higher deposition fluxes are seen in INT than in CLIM (parts of Spain, England, Italy and the North and

GMDD
5, 2083–2138, 2012

West of France), higher concentrations of ozone are simulated in CLIM. On the contrary, over other areas (mainly in Central Europe), the mean deposition flux is higher in CLIM compared to INT, leading to higher concentrations in INT than in CLIM.

In contrast with O_3, smaller differences in flux deposition are observed for NO_x and SO_2; nevertheless, these differences lead to differing concentrations between INT and CLIM. In the case of NO_x, higher concentrations observed in CLIM over the northern area (as in England, Belgium) are related to lower flux deposition at the surface. In addition, SO_2 concentrations rise by 1–1.5 ppbv in CLIM over Northern Spain and Belgium (Fig. 7c); the simulated changes are related to the SO_2 deposition fluxes (Fig. 6c). Very similar distribution and levels of sulphate aerosol are observed in INT and CLIM (Fig. 7e) across Europe.

In summary, the comparisons between ANALY and CLIM (Fig. 8) have revealed the contribution of both meteorological and flux changes on simulated air pollutants. The differences linked to the meteorological parameters or surface processes are pollutant dependent. Depending on the species that are considered, the differences can be driven mainly by the meteorological fields or the emission inventories. The meteorological and surface process effects can also compensate each other. Over the whole domain, the changes in sulphate concentrations between ANALY and CLIM are mostly determined by chemical, physical and dynamical processes due to the meteorological fields (humidity, precipitation). The major changes in isoprene concentrations (Spain, North Africa, Greece) are attributed to dynamical processes (boundary layer height). For the short lived species NO_x and SO_2, we see that the larger changes are localized near the high emission spots. In case of SO_2, the differences between ANALY and CLIM over Europe are explained by both the changes in deposition fluxes and by the meteorological fields. The O_3 concentration differences between the two simulations are partly related to the changes in meteorological fields (such as temperature) but are principally due to the changes in deposition velocities.

GMDD
5, 2083–2138, 2012

3.2 Statistical results: ANALY and CLIM against AirBase

3.2.1 Model interannual variability

Figure 9 represents the temporal series of the model (ANALY in black line; CLIM in gray line) and monthly measured AirBase data (red line) as an average of the daily mean O_3 (a), NO_x (b), SO_2 (c) and particulate matter PM_{10} (d) from 2004 to 2008 across the European domain. If we subtract the observed and simulated annual cycle averaged for the period 2004–2008 from these time series, positive and negative anomalies remain. The meteorology of ANALY is expected to follow the day-to-day variability in a more realistic way than CLIM, which reproduces the climate of the decade 2000–2010. The inter-annual variability of O_3 simulated by ANALY follows the measured variability of meteorological events in terms of correlation and amplitude. The positive and negative anomalies estimated for ANALY are indeed mostly correlated with the observations ($CORRO_3H = 0.61$). This is obvious for the case of summer 2006 heat wave. A positive anomaly, slightly underestimated, is calculated with ANALY while this summer was particularly extreme with an intense production of O_3 (Struzewska and Kaminski, 2008) compared to the mean level of the period 2004–2008. In Table 6 are summarized the statistics of the hourly mean and daily M×8h O_3 levels, averaged for the annual and seasonal period of 2004–2008 over all the European stations considered (Table 5). If the annual trend and the inter-season variability are reproduced, a systematic negative bias is detected with CLIM throughout the summers ($MBO_3H = 4.6\,\mu g\,m^{-3}$, $MBO_3MAX = -5.2\,\mu g\,m^{-3}$).

The model-observation comparisons of NO_x presented in Fig. 9b highlight an accurate annual amplitude of the concentrations for ANALY while the NO_x values simulated by CLIM are overestimated during wintertime ($MBNO_xDM = 8\,\mu g\,m^{-3}$). The dynamics and the intensities of the anomalies are often well-captured by the simulation ANALY; the positive anomalies estimated at the beginning of 2005 and 2006 are for instance well correlated with the observations. Winter pollution characterized by high levels of

GMDD
5, 2083–2138, 2012

NO_x (30–40 µg m^{-3}) are depicted on the time series. For ANALY, the statistics compiled in Table 7 show better performance in term of correlation in winter ($CORRNO_xH = 0.42$; $CORRNO_xDM = 0.55$) than in summer ($CORRNO_xH = 0.29$; $CORRNO_xDM = 0.43$). During winter, chemical processes that lead to O_3 production are less dominant compared to the transport and could explain such differences (Bessagnet et al., 2004).

Time series of monthly mean concentrations of SO_2 from the AirBase stations and the model simulations are represented in Fig. 9c. Results show that the SO_2 concentrations are overestimated for both ANALY and CLIM. Overall, a good agreement is observed for the anomalies between ANALY and the observations in term of amplitude. From the year 2006, we notice a decrease in the observed SO_2 concentrations due to the regulations reducing anthropogenic emissions. Indeed, emissions have been reduced in the sector of power and heat generation with the emission abatement strategies in some European countries during the period of 2000–2010 (EEA, 2007). In our simulations, we kept the same emissions inventory representative of the year 2003. According to the time series of anomalies, and as described in Table 8, the biases calculated in ANALY and CLIM are of the same order of values ($MBSO_2DM = 0.37$ for ANALY and $MBSO_2DM = 0.45$ for CLIM).

Although the simulation ANALY presents a persistent negative bias, it has the capability to reproduce the dynamics of PM_{10} for each year. The underestimation of PM_{10} can be explained principally by the lack of secondary particulate and nitrate aerosols in our representation of PM_{10}. During summer, when photochemistry favors the formation of these particulates, the biases between simulated and observed concentrations become greater ($MBPM_{10} = -11.9$ µg m^{-3}, Fig. 9). In the case of CLIM, the time series of anomalies displays the capability of the model to reproduce the particulate matter events ($CORRPM_{10} = 0.39$), although the model hardly reproduces their amplitude.

GMDD
5, 2083–2138, 2012

3.2.2 Statistical results

The statistics of the model are spatially displayed in Fig. 10: we illustrate the mean biases for O_3 daily M×8h values, as well as for NO_x and SO_2 daily mean concentrations across the representative European stations. The scores are averaged for the summer season (JJAS). Regarding the results of simulation ANALY in the case of O_3, two distinct spatial regimes can be distinguished from the figures: positive and low biases are shown over Germany while negative biases are noticed in Spain and Italy. The correlations (not shown) are also more elevated in the northern part of Europe, notably in Germany (0.6–0.8) while in Southern Europe (Italy and Spain), the performance of the model is rather low (Pay et al., 2010). Comparisons between the statistical metrics of ANALY and CLIM indicate comparable biases over Europe for the daily M×8h O_3. The spatial distribution and the amplitude of the negative and positive biases are mostly similar, except for the stations in Germany and France, which display higher positive (30–40 $\mu g\,m^{-3}$) and negative biases in CLIM than in ANALY, respectively. As shown in Table 6, for ANALY, annual correlations of hourly values and daily M×8h O_3 concentrations reach 0.61 and 0.7, respectively. Also, annual and seasonal MNB values for hourly and daily M×8h O_3 show good performance in accordance with the recommendation of US-EPA (MNBE $\leq \pm 15\,\%$). In summer, the model tends to slightly underestimate the hourly levels of O_3 ($MBO_3H = -2.9\,\mu g\,m^{-3}$) and daily M×8h concentrations ($MBO_3MAX = -8.1\,\mu g\,m^{-3}$), while it overestimates concentrations in winter months ($MBO_3H = 6.4\,\mu g\,m^{-3}$; $MBO_3MAX = 5.7\,\mu g\,m^{-3}$). Correlation values are lowest for both hourly values and daily M×8h over winter: 0.58 and 0.65, respectively. During winter, O_3 is controlled by processes other than photochemistry (such as boundary conditions, deposition, titration), O_3 is thus sensitive to physical and dynamic characteristics (vertical resolution, emissions, deposition velocities). Higher negative bias are thus exhibited by CLIM in wintertime ($MBO_3H = -8.2\,\mu g\,m^{-3}$; $MBO_3MAX = -9.1$) as seen in Fig. 9 due to the differences in deposition velocities and vertical resolution. Daily M×8h are also best reproduced by the model (correlations between 0.63 and 0.7);

GMDD
5, 2083–2138, 2012

the variability of daily M×8h is mainly driven by photochemistry as well as the boundary layer height. To examine if the model is able to simulate the variability of O_3 concentrations, we used the sigma ratio, which is the standard deviation of the modeled time series divided by the standard deviation of the observed time series (Table 3). In Table 6 are summarized the sigma ratio values, averaged for the period of 2004–2008. For both simulations, the model underestimates the observed variability of hourly and daily M×8h values, except for the winter season. In addition, two health related parameters are considered: SOMO35 and the number of exceedance days. SOMO35 corresponds to the yearly sum of the differences between daily maximum 8 h running average concentrations that are greater than 35 ppb (Amann et al., 2005). It is used as an indicator for O_3 health impact and is recommended by the World Health Organization (WHO). The values of SOMO35 are summarized in Table 6. When averaged over all of the European stations considered, the observed seasonal levels reach 2221 µg m^{-3} d and 496 µg m^{-3} d in summer and winter, respectively (4118 µg m^{-3} d for the all year). Both simulations catch the levels of SOMO35 and the seasonal variation (van Loon et al., 2007). According to ANALY, about 40 % of the SOMO35 is produced during summer and 20 % during winter. Over the summer period, the number of days with O_3 exceeding the 120 µg m^{-3} threshold for the daily maximal 8-h is underestimated in ANALY (n = 5.7 days) and fairly well estimated in CLIM (n = 12 days), in comparison to the observations (n = 15.4 days) from the European stations. The mean ozone concentrations above the threshold of 120 µg m^{-3} for the M×8h simulated by ANALY are mostly in line with the observations or within the inter-annual variability. More elevated values are reached in CLIM than ANALY, as shown for the French and Italian stations (Fig. 11a). Figure 11c shows the percentile of daily O_3 maximum simulated by ANALY and CLIM simulations. The interval between the 20th and 70th percentiles display similar values for both simulations. The occurrence of extreme values (maxima) is underestimated by the model for both simulations. As seen previously, above the threshold of 180 µg m^{-3}, CLIM simulates a higher number of occurrences than the observations. These figures

GMDD
5, 2083–2138, 2012

depict however overall that MOCAGE driven by climate model outputs as forcings is able to simulate realistic ozone concentrations over Europe.

As exposed in Table 5, 354 stations were used to provide NO_x measurements throughout Europe. Considering the spatial distribution of mean biases for daily mean NO_x, statistic results show satisfactory seasonal mean bias for ANALY ($MBNO_xDM = -2.4\,\mu g\,m^{-3}$ during summer) without spatial pattern between north and south. Similar geographical distributions and values of mean bias are displayed for CLIM (Fig. 10) while the summer mean bias reaches $-2.1\,\mu g\,m^{-3}$ (Table 7). The spatial distribution of the correlation coefficients shows a large variability per station: while northern stations display high correlations ($0.6 < r < 0.8$), low correlations are observed in Southern Europe ($r < 0.4$). The performances of the model are reduced with CLIM for all seasons, notably during winter (Fig. 7) when the concentrations of NO_x are overestimated as shown in Fig. 9. Thus, in winter, $MNBNO_xH$ and $MNBNO_xDM$ reach 55.1 % and 54.3 %, respectively; also $RMSENO_xH$ and $RMSENO_xDM$ reach $29.2\,\mu g\,m^{-3}$ and $25.7\,\mu g\,m^{-3}$, respectively. In summer, the MNB values for hourly and daily mean are near the uncertainty proposed by EC and US-EPA for the ANALY simulation. Globally, the annual and seasonal daily mean statistics present better performances in comparison with the hourly values.

For SO_2, low correlations are mainly concentrated in Southern Europe (coefficients under 0.2), as in Spain, while some northern stations display high correlations ($r > 0.7$) in regard to statistic results of ANALY. Averaged over all European stations, the summertime mean correlation of daily mean SO_2 reaches 0.3. Considering the mean bias for summer (Fig. 10), low biases are depicted across all the stations ($MBSO_2DM = -0.07$ for ANALY and CLIM). Nevertheless, the stations located in Poland display high positive biases ($> 2\,\mu g\,m^{-3}$). The uncertainties of the emissions inventory in Eastern Europe may contribute to the higher bias observed. In some stations in Spain, higher bias are also observed, due to the emission inventory in some part. The regulation of SO_2 emissions in Spain have lead to an emission reduction of 50 % (Spain Environment Ministry, 2011). As shown in Table 8, highest correlations are obtained in winter

GMDD
5, 2083–2138, 2012

($CORRSO_2H$ = 0.27; $CORRSO_2DM$ = 0.40) but the concentrations are overestimated leading to a high value of MNB close to the EC criteria.

As seen in Fig. 9, the model presents a systematic negative bias for the simulated concentrations of PM_{10}. For the ANALY simulation, the correlation coefficient for the annual daily mean is 0.39, while it reaches 0.2 and 0.48 for the summer and winter season, respectively (Table 9). The spatial distribution of mean bias and correlations (not shown here) present a homogeneous pattern over Europe. The annual MFE ($MFEPM_{10}$ = 75.1 %) and MFB ($MFBPM_{10}$ = −59.5 %) calculated for ANALY does not meet the performance criteria or the performance goal proposed by Boylan and Russell (2006). The performance of the model is better during winter when the $MBPM_{10}$ is about $-4.6\,\mu g\,m^{-3}$ (against $-11.9\,\mu g\,m^{-3}$ in summer) and the mean correlation reaches 0.48 (against 0.2 for the summer). The underestimation of PM_{10} can be explained by the lack of secondary particulate and nitrate aerosols in our representation of PM_{10}. Differences between the seasons are linked to chemical processes, dominant in summer, which favor the formation of these particulates and increase the bias between simulated concentrations (lacking these chemical processes) and observations.

Several air quality models operated in Europe have been evaluated either individually or in comparison to other models in the literature. In the following discussion, a quick comparison with other regional air quality models (Hass et al., 2003; van Loon et al., 2004, 2007) and MOCAGE will be carried out in order to situate our model among the community. For this reason, we used the studies similar with our simulation ANALY, which had a long time scale of 1 yr over the Europe domain on a regional scale with horizontal resolutions similar to MOCAGE. Also, these models were evaluated against ground observations at rural sites from AirBase and EMEP. Concerning O_3 daily M×8h, satisfactory performances are displayed with MOCAGE, in terms of annual MNB values: 0.04 % versus −1 to 10 %; correlations: 0.7 versus 0.69–0.84 (van Loon et al., 2007) and RMSE: $21.3\,\mu g\,m^{-3}$ versus $18.1–25.5\,\mu g\,m^{-3}$ (van Loon et al., 2004). Values for summer and winter daily M×8h O_3 are also in the range of other models, as for the correlations (0.63 versus 0.61–0.77 for summer; 0.65 versus 0.45–0.62

GMDD
5, 2083–2138, 2012

for winter) and the MNB (−9.44 % versus −5 to 8 % for summer; 9.8 % versus −20 to 15 % for winter) according to the study of van Loon et al. (2007). The MOCAGE performances for NO_x can be compared with the performance of NO_2 in other models. The annual correlation of daily mean NO_x obtained in this study reaches 0.61, compared to 0.03–0.52 (Hass et al., 2003; van Loon et al., 2004) and the RMSE value is around 9.8 µg m^{-3} versus 8.5–13.9 µg m^{-3} (Hass et al., 2003; van Loon et al., 2004). As for O_3, the MOCAGE results for SO_2 show good performances in comparison with the other studies. The annual daily mean correlation is among the higher value (0.36 versus 0.24–0.49). The calculated RMSE reaches 3.2 µg m^{-3} against 2.7–10.9 µg m^{-3} for the other models (Hass et al., 2003; van Loon et al., 2004). For PM_{10}, statistical results are in the same range as for other studies: nevertheless, the annual daily mean correlation is rather low and reaches 0.39 compared to 0.38–0.55; the annual RMSE is 15.1 µg m^{-3} (versus 12.4–16.6 µg m^{-3}).

To summarize, MOCAGE performs well according to the comparisons between ANALY and AirBase observations as discussed previously. The statistical scores of O_3, NO_x and SO_2 display satisfactory performances compared to other studies while the accuracy in our representation of PM_{10} exhibit poorer results which are expected by design. Comparisons between the simulations ANALY and CLIM have shown that the geographical distribution of mean biases are quite similar for each pollutant considered. In this section, the model to observation comparisons were based on a common approach, which consists in comparing each year of the simulation with the matching measured values from the AirBase database. In CLIM, the meteorological forcings are representative of the current decade: there is no particular match in the sequence of years and the representativeness of the skill scores can be assessed by permutations of all years. By doing the same for ANALY, the comparisons will allow us to determine which statistical tools are useful to consider for future studies.

GMDD
5, 2083–2138, 2012

3.2.3 Impacts of chronology of pollution events

We evaluated each year of the simulations with 5 yr of measurements, giving 25 model-to-date pairs of statistics for both ANALY and CLIM. On the time basis of the 2004–2008 period, we calculated every possible permutation, but on this period of 5 yr, we did not consider the same year of measurements more than once. We also filtered out the cases when one or more simulated years correspond with the same years of data. Finally, these conditions let us consider 44 realizations which provide a large statistical basis.

In order to give a concise statistical summary, we used the Taylor diagrams, which indicate how well observed and simulated patterns match each other in terms of correlation and normalized standard deviation NSD (Taylor, 2001). The correlation coefficient R gives a measure of the co-variance of simulated and observed values. The NSD gives a measure of the amplitude of the variance in modeled values versus observed values. When NSD reaches a value lower than 1, it means that the temporal standard deviation in simulated values is lower than observed. Figure 12 shows the normalized Taylor plots that summarize the ANALY and CLIM permutations. The statistics are computed for the summertime daily M×8h O_3 concentrations (a), daily averages of NO_x (b) and SO_2 (c). For each plot, "ANALY" refers to the reference model–observation analyses (black cross) whereas "ANALY-p" (blue symbols) and "CLIM-p" (red symbols) refer to the permutation cases. Considering the results of ozone first, comparisons between ANALY and ANALY-p confirm there is no day-to-day variability with the permutations of ANALY-p as shown by the correlations. As summarized in Table 10, for ANALY-p, the median value of the correlations reach 0.04 (against 0.63 for ANALY). The median RMSE of ANALY-p are not as good as the median RMSE for ANALY for both the hourly ($RMSEO_3H = 31\,\mu g\,m^{-3}$ for ANALY-p and $21.4\,\mu g\,m^{-3}$ for ANALY) and daily M×8h ($RMSEO_3MAX = 30.9\,\mu g\,m^{-3}$ for ANALY-p and $24.9\,\mu g\,m^{-3}$ for ANALY) of O_3. Nevertheless, very similar values of $MNBO_3MAX$ and $MNBO_3H$ in the ranges of −9 % and −4.5 %, respectively are obtained. The standard deviation values (σ) are similar in

GMDD
5, 2083–2138, 2012

ANALY and ANALY-p and range around 0.66 meaning that the model underestimates the daily M×8h ozone variability. The same conclusions can be extended to the daily averages of NO_x, SO_2 (Table 11) and PM_{10} (Table 12) concentrations. ANALY and ANALY-p only differ by the correlations while the MB, RMSE and variances are quite similar. For the daily mean NO_x (Fig. 12), the NSD are close to the reference meaning that the amplitude of the simulated NO_x agrees with the observations. To sum up, these results point out that, for all the species, the correlations calculated for ANALY-p are weaker than ANALY due to the permutations. The RMSE values increase by 19 %, 11 % and 8 % for the hourly values of O_3, NO_x and SO_2, respectively from ANALY to ANALY-p. However, for the metrics MB, MNB and σ, similar values are computed.

For O_3 (Fig. 12), low and similar correlations of daily M×8h levels are calculated for both ANALY-p and CLIM-p. As shown in Table 10, the correlations between the observed and simulated hourly values (0.33 for ANALY-p and 0.34 for CLIM-p) are higher than daily M×8h. The daily variability of ozone, characterized by higher levels during afternoon and lower values during nighttime hours is still captured with the permutations. The RMSE values show lower performances for CLIM-p compared to ANALY-p for both daily M×8h (+10 %) and hourly O_3 (+13 %) levels. The σ values indicate the tendency of ANALY-p to underestimate ozone variance in summer. The CLIM-p simulations show better σ values (0.9 and 0.98 versus 0.67 and 0.75 for ANALY-p) which is unexpected and can not be interpreted as greater performance. The MNB of M×8h ozone show a tendency of model underestimation as the median reaches −9.1 % in ANALY-p and −6.4 % in CLIM-p. For the hourly and daily average NO_x and SO_2 concentrations, the simulations ANALY-p and CLIM-p are now well correlated (Table 11). During summer, ANALY-p and CLIM-p underestimate the daily mean and hourly NO_x concentrations. From ANALY-p to CLIM-p, the median $MNBNO_xDM$ and $MNBNO_xH$ change by about 14 % while the $RMSENO_xDM$ and $RMSENO_xH$ change by about 7 %. Concerning the amplitude of the NO_x variances (Fig. 12), a satisfactory agreement is observed, σ reaches 0.98 in ANALY-p and 1.07 in CLIM-p for daily mean NO_x (Table 11). For the SO_2 results, the variance is overestimated by the model for both ANALY-p and CLIM-p.

GMDD
5, 2083–2138, 2012

For the daily mean of PM_{10} concentrations, the amplitude of the variances are in line for ANALY-p and CLIM-p simulations. The simulations underestimate the variability of the PM_{10} (σPM_{10} = 0.49 for ANALY-p and σPM_{10} = 0.59 for CLIM-p). $MBPM_{10}$ reaches similar values for ANALY-p and CLIM-p. MFB and MFE metrics do not reach the performance criterion.

To summarize, the use of permutations have made the comparisons between the simulations suitable for discussion of the statistical results. Comparisons between the reference case "ANALY" with "ANALY-p" corroborate the incorrect phasing between the measurements and simulations when the day-to-day variability is not reproduced. Finally, the results allow us to conclude that statistical metrics such as variances, MB, MNB and RMSE give robust and sound information when climate forcings are used to drive the model.

4 Conclusions

This paper has investigated how different the hindcasts of an air quality modelling system are when using two different types of meteorological forcings: meteorological analyses or fields from a climate model. This is ground work needed to qualify and properly interpret statistical conclusions that can be drawn from simulations of air quality in a future climate.

The comparisons between three 5-yr experiments allow us to quantify the relative importance of changes in surface fields and upper air meteorology. We find that both elements contribute to changes in O_3 concentrations. Differences in sulphate aerosols and in isoprene (as a proxy for biogenic volatile organic compounds) are mainly related to changes in meteorology and mixing while it is the contrary for SO_2 and NO_x, which essentially depend upon changes in surface fluxes.

The skill of the reference simulation (analysed forcings) to reproduce European surface observations is in the range of previous reference evaluation studies (Hass et al., 2003; van Loon et al., 2004, 2007). As expected, the simulation based upon forcings

GMDD
5, 2083–2138, 2012

from a climate model is not as skillful at reproducing observations, since it cannot follow day-to-day variations by design. Nevertheless, the geographical distributions of the mean biases are similar in the two simulations. The comparisons of SOMO35 and the distribution of O_3 maximum percentiles support the capability of MOCAGE, driven by meteorological fields from a climate model, to simulate realistic European ozone levels.

The objective of this work was to determine useful statistical metrics that can be used for models driven by climate model meteorological parameters. For O_3, we show that simulations using either analyses or climate model forcings follow the same tendency: hourly and M×8h concentrations are slightly underestimated and the biases and RMSE are in the same range of values. Similar conclusions are observed for the daily averaged and hourly values of NO_x and SO_2, as well as daily PM_{10}. The amplitude of variance is accurately reproduced when the model is driven by climate fields. Finally, as for the standard deviation, statistical results of MB, MNB and RMSE can be interpreted with some degree of confidence.

Acknowledgements. This work was partly founded by ADEME (Agence De l'Environnement et de la Maitrise de l'Energie, http://www.ademe.fr). We would like to thank N. Poisson (ADEME), L. Rouil (INERIS) and M. Beekman (CNRS/LISA) for useful discussions.

The publication of this article is financed by CNRS-INSU.

GMDD

5, 2083–2138, 2012

References

Amann, M., Bertok, I., Cofala, J., Gyarfas, F., Heyes, C., Klimont, Z., Schopp, W., and Winiwater, W.: Clean air for Europe (CAFE) programme final report, Tech. rep., 2005. 2101

Bechtold, P., Bazile, E., Guichard, F., Mascart, P., and Richard, E.: A mass-flux convection scheme for regional and global models, Q. J. Roy. Meteor. Soc., 127, 869–886, 2001. 2088

Bessagnet, B., Hodzic, A., Vautard, R., Beekmann, M., Cheinet, S., Honoré, C., Liousse, C., and Rouïl, L.: Aerosol modeling with CHIMERE-preliminary evaluation at the continental scale, Atmos. Environ., 38, 2803–2817, 2004. 2099

Bousserez, N., Attié, J.-L., Peuch, V.-H., Michou, M., Pfister, G., Edwards, D., Emmons, L., Mari, C., Barret, B., Arnold, S. R., Heckel, A., Richter, A., Schlager, H., Lewis, A., Avery, M., Sachse, G., Browell, E. V., and Hair, J. W.: Evaluation of the MOCAGE chemistry transport model during the ICARTT/ITOP experiment, J. Geophys. Res., 112, D10S42, doi:10.1029/2006JD007595, 2007. 2088

Boylan, J. and Russell, A.: PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models, Atmos. Environ., 40, 4946–4959, 2006. 2090, 2103

Carvalho, A., Monteiro, A., Solman, S., Miranda, A., and Borrego, C.: Climate-driven changes in air quality over Europe by the end of the 21st century, with special reference to Portugal, in: Environmental Science and Policy, Elsevier, Oxford, England, 445–558, 2010. 2086

Chang, J. and Hanna, S.: Air quality model performance evaluation, Meteor. Atmos. Phys., 87, 167–196, 2004. 2090

Courtier, P., Freydier, C., Geleyn, J. F., Rabier, F., and Rochas, M.: The ARPEGE project at Météo France, in: Atmospheric Models, vol. 2, Workshop on Numerical Methods, Reading, UK, 193–231, 1991. 2088

Cuvelier, C., Thunis, P., Vautard, R., Amann, M., Bessagnet, B., Bedogni, M., Berkowicz, R., Brandt, J., Brocheton, F., Builtjes, P., Coppalle, A., Denby, B., Douros, G., Graf, A., Hellmuth, O., Honoré, C., Hodzic, A., Jonson, J., Kerschbaumer, A., de Leeuw, F., Minguzzi, E., Moussiopoulos, N., Pertot, C., Pirovano, G., Rouïl, L., Schaap, M., Stern, R., Tarrason, L., Vignati, E., Volta, M., White, L., Wind, P., and Zuber, A.: CityDelta: a model intercomparison study to explore the impact of emission reductions in European cities in 2010, Atmos. Environ., 41, 189–207, 2007. 2090

GMDD
5, 2083–2138, 2012

Dawson, J. P., Adams, P. J., and Pandis, S. N.: Sensitivity of ozone to summertime climate in the Eastern USA: a modeling case study, Atmos. Environ., 41, 1494–1511, 2007. 2094
Dawson, J. P., Racherla, P. N., Lynn, B. H., Adams, P. J., and Pandis, S. N.: Impacts of climate change on regional and urban air quality in the Eastern United States: role of meteorology, J. Geophys. Res., 114, D05308, doi:10.1029/2008JD009849, 2009. 2086
Dentener, F. E. A.: The global atmospheric environment for the next generation, Environ. Sci. Technol., 40, 3586–3594, doi:10.1021/es0523845, 2006. 2086
Déqué, M., Dreveton, C., Braun, A., and Cariolle, D.: The ARPEGE/IFS atmosphere model: a contribution to the French community climate modelling, Clim. Dynam., 10, 249–266, 1994. 2088
Dufour, A., Amodei, M., Ancellet, G., and Peuch, V.-H.: Observed and modelled "chemical weather" during ESCOMPTE, Atmos. Res., 74, 161–189, doi:10.1016/atmosres.2004.04.013, 2004. 2088
EEA: Air pollution by ozone in Europe in summer 2004, Technical Report, 3/2005, Copenhagen, Denmark, available at: http://reports.eea.eu.int, 2005. 2094
EEA: Air pollution in Europe 1999–2004, Technical Report, 2/2007, Office for Official Publications of the European Communities, Luxembourg, 79 pp., 2007. 2099
European Commission: Council directive 1999/30/EC of 22 April 1999 relating to limit values for sulphur dioxide, nitrogen dioxide and oxides of nitrogen, particulate matter and lead in ambient air, Technical Report 1999/30/EC, L163, Off. J. Eur. Comm., 1999. 2090
European Commission: Commission decision of 17 October 2001 amending annex v to Council directive 1999/30/EC relating to limit values for sulphur dioxide, nitrogen dioxide and oxides of nitrogen, particulate matter and lead in ambient air (text with EEA relevance) (notified under document number C (2001) 3091, Technical Report 2001/744/EC, 278, Off. J. Eur. Comm., 2001. 2090
European Commission: Directive 2002/3/EC of the European Parliament and of the Council of 12 February 2002 relating to ozone in ambient air, Technical Report 2002/3/EC, l67, Off. J. Eur. Comm., 2002. 2090
European Commission: Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe, Technical Report 2008/50/EC, L152, Off. J. Eur. Comm., 2008. 2086, 2090, 2091
Fuhrer, J. and Booker, F.: Ecological issues related to ozone: agricultural issues, Environ. Int., 29, 141–154, 2003. 2085

GMDD
5, 2083–2138, 2012

Giorgi, F. and Meleux, F.: Modelling the regional effects of climate change on air quality, C. R. Geosci., 339, 721–733, 2007. 2086
Guenther, A., Zimmerman, P., Harley, P., Monson, R., and Fall, R.: Isoprene and monoterpene emission rate variability: model evaluation and sensitivity analysis, J. Geophys. Res., 98, 12609–12617, doi:10.1029/93JD00527, 1993. 2096
Guenther, A., Hewitt, C. N., Erickson, D., Fall, R., Geron, C., Graedel, T., Harley, P., Klinger, L., Lerdau, M., Mckay, W., Pierce, T., Scholes, B., Steinbrecher, R., Tallamraju, R., Taylor, J., and Zimmerman, P. A.: A global model of natural volatile organic compound emissions, J. Geophys. Res., 100, 8873–8892, doi:10.1029/94JD02950, 1995. 2096
Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmos. Chem. Phys., 6, 3181–3210, doi:10.5194/acp-6-3181-2006, 2006. 2089, 2096
Guerova, G. and Jones, N.: A global model study of ozone distributions during the August 2003 heat wave in Europe, Environ. Chem., 4, 85–292, 2007. 2085
Guicherit, R. and van Dop, H.: Photochemical production of ozone in Western Europe (1971–1975) and its relation to meteorology, Atmos. Environ., 11, 145–155, doi:10.1071/EN07027, 1977. 2085
Hass, H., van Loon, M., Kessler, C., Stern, R., Matthijsen, J., Sauter, F., Zlatev, Z., Langner, J., Foltescu, V., and Schaap, M.: Aerosol modeling: results and intercomparison from European regional scale modeling systems, Technical Report, EUROTRAC 2 Report, EUREKA Environmental Project, GLOREAM, 2003. 2103, 2104, 2107
Hedegaard, G. B., Brandt, J., Christensen, J. H., Frohn, L. M., Geels, C., Hansen, K. M., and Stendel, M.: Impacts of climate change on air pollution levels in the Northern Hemisphere with special focus on Europe and the Arctic, Atmos. Chem. Phys., 8, 3337–3367, doi:10.5194/acp-8-3337-2008, 2008. 2094
Hogrefe, C., Lynn, B., Civerolo, K., Ku, J.-Y., Rosenzweig, J. R. C., Goldberg, R., Gaffin, S., Knowlton, K., and Kinney, P. L.: Simulating changes in regional air pollution over the Eastern United States due to changes in global and regional climate and emissions, J. Geophys. Res., 109, D22301, doi:10.1029/2004JD004690, 2004. 2086
Hollingsworth, A., Engelen, R., Textor, C., Boucher, O., Chevallier, F., Dethof, A., Elbern, H., Eskes, H., Flemming, J., Granier, C., Kaiser, J. W., Morcrette, J. J., Rayner, P., Peuch, V.-H., Rouïl, L., and the GEMS consortium: Toward a monitoring and forecasting system for

GMDD
5, 2083–2138, 2012

atmospheric composition: the GEMS project, Bull. Amer. Meteor. Soc., 89, 1147–1164, doi:1S230.1175/2008BAM55.1, 2008. 2088, 2089

Honoré, C., Rouïl, L., Vautard, R., Beekmann, M., Bessagne, B., Dufour, A., Elichegaray, C., Flaud, J.-M., Malherbe, L., Meleux, F., Menut, L., Martin, D., Peuch, A., Peuch, V.-H., and Poisson, N.: Predictability of European air quality: assessment of 3 years of operational forecasts and analyses by the PREV'AIR system, J. Geophys. Res., 113, D04301, doi:10.1029/2007JD008761, 2008. 2088

Joly, M. and Peuch, V.-H.: Objective classification of air quality monitoring sites over Europe, Atmos. Environ., 47, 111–123, 2012. 2091

Katragkou, E., Zanis, P., Tegoulias, I., Melas, D., Kioutsioukis, I., Krüger, B. C., Huszar, P., Halenka, T., and Rauscher, S.: Decadal regional air quality simulations over Europe in present climate: near surface ozone sensitivity to external meteorological forcing, Atmos. Chem. Phys., 10, 11805–11821, doi:10.5194/acp-10-11805-2010, 2010. 2094

Knowlton, K., Rosenthal, J. E., Hogrefe, C., Lynn, B., Gaffin, S., Goldberg, R., Rosenzweig, C., Civerolo, K., Ku, J. Y., Kinney, P. L., and al.: Assessing ozone-related health impacts under changing climate, Environ. Health Persp., 112, 1557–1563, 2004. 2086

Langner, L., Bergstrom, R., and Foltescu, V.: Impact of climate change on surface ozone and deposition of sulphur and nitrogen in Europe, Atmos. Environ., 39, 1129–1141, doi:10.1016/j.atmosenv.2004.09.082, 2005. 2085, 2086

Lefèvre, F., Brasseur, G. P., Folkins, I., Smith, A. K., and Simon, P.: Chemistry of the 1991–1992 stratospheric winter: three dimensional model simulations, J. Geophys. Res., 99, 8183–8195, 1994. 2088

Louis, J.-F.: Parametric model of vertical eddy fluxes in the atmosphere, Bound.-Lay. Meteorol., 17, 187–202, 1979. 2088

Martet, M., Peuch, V.-H., Laurent, B., Marticorena, B., and Bergametti, G.: Evaluation of long-range transport and deposition of desert dust with the CTM MOCAGE, Tellus, 61B, 449–463, 2009. 2088

Meleux, F., Solmon, F., and Giorgi, F.: Increase in summer European ozone amounts due to climate change, Atmos. Environ., 41, 7577–7587, 2007. 2086, 2089, 2094

Ménégoz, M., Voldoire, A., Teyssèdre, H., Salas Y Mélia, D., Peuch, V.-H., and Gouttevin, I.: How does the atmospheric variability drive the aerosol residence time in the Arctic region?, Tellus, 64, doi:10.3402/tellusb.v64i0.11596, 2012. 2088

GMDD

5, 2083–2138, 2012

Pay, M., Piot, M., Jorba, O., Gassó, S., Gonçalves, M., Basart, S., Dabdub, D., Jiménez-Guerrero, P., and Baldasano, J. M.: A full year evaluation of the CALIOPE-EU air quality modeling system over Europe for 2004, Atmos. Environ., 44, 3322–3342, 2010. 2100

Peuch, V.-H., Amodei, M., Barthet, T., Cathala, M.-L., Michou, M., and Simon, P.: MOCAGE, MOdéle de Chimie Atmosphérique à Grande Echelle, in: Proceedings of Météo France: Workshop on atmospheric modelling, Toulouse, France, 33–36, 1999. 2087

Pouliot, G., Pierce, T., Denier van der Gon, H., Schapp, M., Moran, M., and Nopmongcol, U.: Comparing emission inventories and model-ready emission datasets between Europe and North America for the AQMEII project, Atmos. Environ., 53, 4–14, 2012. 2089

Prather, M., Gauss, M., Berntsen, T., Isaksen, I., Sundet, J., Bey, I., Brasseur, G., Dentener, F., Derwent, R., Stevenson, D., Grenfell, L., Hauglustaine, D., Horowitz, L., Jacob, D., Mickley, L., Lawrence, M., von Kuhlmann, R., Muller, J.-F., Pitari, G., Rogers, H., Johnson, M., Pyle, J., Law, K., van Weele, M., and Wild, O.: Fresh air in the 21st century?, Geophys. Res. Lett., 30, 1100, doi:10.1029/2002GL016285, 2003. 2086

Schlink, U., Herbarth, O., Richter, M., Dorling, S., Nunnari, G., Cawley, G., and Pelikan, E.: Statistical models to assess the health effects and to forecast ground-level ozone, Environ. Modell Softw., 21, 547–558, doi:10.1016/j.envsoft.2004.12.002, 2006. 2085

Sillman, S.: Ozone production efficiency and loss of NO_x in power plant plumes: photochemical model and interpretation of measurements in Tennessee, J. Geophys. Res., 105, 9189–9202, doi:10.1029/1999JD901014, 2000. 2085

Solberg, S., Hov, O., Sovde, A., Isaksen, I., Coddeville, P., De Backer, H., Forster, C., Orsolini, Y., and Uhse, K.: European surface ozone in the extreme summer 2003, 113, D07307, doi:10.1029/2007JD009098, 2008. 2085, 2089

Stockwell, W. R., Kirchner, F., Khun, M., and Seefeld, S.: A new mechanism for regional atmospheric chemistry modelling, J. Geophys. Res., 102, 25847–25879, 1997. 2088

Struzewska, J. and Kaminski, J.: Formation and transport of photooxidants over Europe during the July 2006 heat wave – observations and GEM-AQ model simulations, Atmos. Chem. Phys., 8, 721–736, doi:10.5194/acp-8-721-2008, 2008. 2098

Szopa, S., Hauglustaine, D., Vautard, R., and Menut, L.: Future global tropospheric ozone changes and impact on European air quality, Geophys. Res. Lett., 115, L14805, doi:10.1029/2006GL025860, 2006. 2086

Taylor, K.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res., 106, 7183–7192, 2001. 2105

GMDD
5, 2083–2138, 2012

US-EPA: Interim procedures for evaluating air quality models, Technical Report. EPA-450/4-91-013, Research Triangle Park, Environmental Protection Agency, Office of Air Quality Planning and Standards, NC, US, 1984. 2090

US-EPA: Guideline for regulatory application of the urban airshed model, Technical Report, EPA-450/4-91-013, Research Triangle Park, Environmental Protection Agency, Office of Air Quality Planning and Standards, NC, US, 1991. 2090, 2091

van Loon, M., Roemer, M., Builtjes, P., Bessagnet, B., Rouïl, L., Christensen, J., Brandt, J., Fagerli, H., Aasón, L., and Rodgers, I.: Model Inter-comparison in the Framework of the Review of the Unified EMEP Model, TNO Report, Technical Report R2004/282, 53 pp., 2004. 2103, 2104, 2107

van Loon, M., Vautard, R., Schaap, M., Bergstrom, R., Bessagnet, B., Brandt, J., Builtjes, P., Christensen, J., Cuvelier, C., Graff, A., Jonson, J., Krol, M., Langner, J., Roberts, P., Rouïl, L., Stern, R., Tarrasòn, L., Thunis, P., Vignati, E., White, L., and Wind, P.: Evaluation of long-term ozone simulations from seven regional air quality models and their ensemble, Atmos. Environ., 41, 2083–2097, doi:10.1016/j.atmosenv.2006.10.073, 2007. 2101, 2103, 2104, 2107

Vautard, R., Honore, C., Beekmann, M., and Rouïl, L.: Simulation of ozone during the August 2003 heat wave and emission control scenarios, Atmos. Environ., 39, 2957–2967, doi:10.1016/j.atmosenv.2005.01.039, 2005. 2085, 2094

Vautard, R., Builtjes, P., Thunis, P., Cuvelier, K., Bedogni, M., Bessagnet, B., Honoré, C., Moussiopoulos, N., Schaap, M., Stern, R., Tarrason, L., and van Loon, M.: Evaluation and intercomparison of ozone and PM_{10} simulations by several chemistry-transport models over 4 European cities within the city-delta project, Atmos. Environ., 41, 173–188, 2007. 2085

Visschedijk, A. H. J. and Denier van der Gon, H. A.: Gridded European anthropogenic emission data for NO_x, SO_2, NMVOC, NH_3, CO, PM_{10}, $PM_{2.5}$ and CH_4 for the year 2000, TNO B O-A Rapport 2005/106, 2005. 2089

WHO: Health aspects of air pollution results from the WHO project "Systematic review of health aspects of air pollution in Europe", Tech. rep., 2004. 2085

Williamson, D. L. and Rasch, P. J.: Two-dimensional semi-lagrangian transport with shape-preserving interpolation, Mon. Weather Rev., 117, 102–129, 1989. 2088

Zlatev, Z.: Comprehensive air pollution studies by the Danish Eulerian Model, in: Air, Water and Soil Quality Modelling for Risk and Impact Assessment, edited by: Ebel, A., Davitashvili, T., Nato Science for Peace and Security, Environmental Security, 293–302, 2007. 2086

GMDD

5, 2083–2138, 2012

Table 1. Simulations considered for the present study.

Simulations	Periods	Meteorological forcings	Emissions			Deposition velocities
			anthropogenic	biogenic	desert dust, sea salt	
ANALY	2003–2008	ARPEGE	GEMS	MEGANv2.4 /ARPEGE	ARPEGE	ARPEGE
INT	6 yr of 2000–2010 climate	ARPEGE-Climate	GEMS	MEGANv2.4 /ARPEGE	ARPEGE	ARPEGE
CLIM	6 yr of 2000–2010 climate	ARPEGE-Climate	GEMS	MEGANv2.4 /ARPEGE-Climate	ARPEGE-Climate	ARPEGE-Climate

GMDD
5, 2083–2138, 2012

Table 2. Some regulatory European air quality thresholds of O_3, NO_2, SO_2 and PM_{10} levels.

Pollutant	Parameter	Threshold values
O_3	hourly average	180 µg m^{-3}
	daily maxima of 8 h running	120 µg m^{-3}
NO_2	hourly average	200 µg m^{-3}
	annual average	40 µg m^{-3}
SO_2	hourly average	350 µg m^{-3}
	daily average	125 µg m^{-3}
PM_{10}	daily average	50 µg m^{-3}

GMDD
5, 2083–2138, 2012

Table 3. Definition of the metrics used in the evaluation of the MOCAGE model performance. M refers to the model, O refers to the observations. $\bar{M} = \frac{1}{N} \sum(\mathrm{M}_i)$ et $\bar{O} = \frac{1}{N} \sum(O_i)$.

Mean bias	$\mathrm{MB} = \frac{1}{N} \sum_{i=1}^{n} (M_i - O_i) = \bar{M} - \bar{O}$
Mean normalized bias	$\mathrm{MNB} = \frac{1}{N} \sum_{i=1}^{n} \frac{M_i - O_i}{O_i}$
Mean fractional error	$\mathrm{MFE} = \frac{2}{N} \sum_{i=1}^{n} \frac{\lvert M_i - O_i \rvert}{(M_i + O_i)}$
Mean fractional bias	$\mathrm{MFB} = \frac{2}{N} \sum_{i=1}^{n} \frac{(M_i - O_i)}{(M_i + O_i)}$
Correlation coefficient	$r = \frac{\sum_{i=1}^{n} (M_i - \bar{M})(O_i - \bar{O})}{\left\{ \sum_{i=1}^{n} (M_i - \bar{M})^2 \sum_{i=1}^{n} (O_i - \bar{O})^2 \right\}^{1/2}}$
Root mean square error	$\mathrm{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{n} (M_i - O_i)^2}$
Sigma ratio	$\sigma = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^{n} (M_i - \bar{M})^2}}{\sqrt{\frac{1}{N} \sum_{i=1}^{n} (O_i - \bar{O})^2}}$

GMDD
5, 2083–2138, 2012

Table 4. Metrics considered in the evaluation of O_3, NO_x, SO_2 and PM_{10} concentrations.

Indicators	Parameters	Codes			
		O_3	NO_x	SO_2	PM_{10}
Mean bias	hourly value	MBO_3H	$MBNO_xH$	$MBSO_2H$	–
	daily mean	–	$MBNO_xDM$	$MBSO_2DM$	$MBPM_{10}$
	M×8h	MBO_3MAX	–	–	–
Mean normalized bias	hourly value	$MNBO_3H$	$MNBNO_xH$	$MNBSO_2H$	–
	daily mean	–	$MNBNO_xDM$	$MNBSO_2DM$	–
	M×8h	$MNBO_3MAX$	–	–	–
Mean fractional bias	daily mean	–	–	–	$MFBPM_{10}$
Mean fractional error	daily mean	–	–	–	$MFEPM_{10}$
Correlation coefficient	hourly value	$CORRO_3H$	$CORRNO_xH$	$CORRSO_2H$	–
	daily mean	–	$CORRNO_xDM$	$CORRSO_2DM$	$CORRPM_{10}$
	M×8h	$CORRO_3MAX$	–	–	–
Roor mean square error	hourly value	$RMSEO_3H$	$RMSENO_xH$	$RMSESO_2H$	–
	daily mean	–	$RMSENO_xDM$	$RMSESO_2DM$	–
	M×8h	$RMSEO_3MAX$	–	–	–
Sigma ratio	hourly value	σO_3H	σNO_xH	σSO_2H	–
	daily mean	–	σNO_xDM	σSO_2DM	σPM_{10}
	M×8h	σO_3MAX	–	–	–
SOMO35*	M×8h	SOMO35	–	–	–

* SOMO35 annual sum of excess of daily maximum 8-h means ozone over the cut-off of 35 ppb.

GMDD
5, 2083–2138, 2012

Table 5. Number of representative sites available by countries and species. Countries are, from left to right, France (FR), Spain (ES), England (GB), Germany (DE), Italy (IT) and Poland (PL).

Species	Classes	Countries						
		Europe	FR	ES	GB	DE	IT	PL
O_3	1–5	970	201	191	33	202	54	47
NO_x	1–2	354	68	115	–	97	9	–
SO_2	1–2	342	31	101	7	101	17	7
PM_{10}	1–2	824	168	151	47	259	–	28

GMDD
5, 2083–2138, 2012

Table 6. Seasonnal and annual statistics obtained with MOCAGE over Europe at the AirBase stations. Statistics are averaged for the 5 yr period. Summer: June, July, August and September; winter: December, January, February and March. The calculated statistics are: mean bias (MBO_3H, MBO_3MAX; $\mu g\,m^{-3}$), mean normalized bias ($MNBO_3H$, $MNBO_3MAX$, %), correlation coefficient ($CORRO_3H$, $CORRO_3MAX$), Root Mean Square Error ($RMSEO_3H$, $RMSEO_3MAX$; $\mu g\,m^{-3}$) and sigma ratio (σH, σMAX). Statistics are computed for the O_3 hourly value and O_3 M×8h. SOMO35 values are in $\mu g\,m^{-3}\,d$.

O_3 Metrics	ANALY			CLIM		
	Year	JJAS	DJFM	Year	JJAS	DJFM
MBO_3H	2.7	−2.9	6.4	−5.3	−4.6	−8.2
$MNBO_3H$	5.1	−4.5	13.8	−9.8	−7.7	−17.7
$CORRO_3H$	0.61	0.59	0.58	0.3	0.32	0.17
$RMSEO_3H$	24.7	24.9	24.3	34.3	35.4	34.1
σH	0.87	0.73	1.04	0.98	0.97	1.08
MBO_3MAX	0.01	−8.1	5.7	−6.3	−5.2	−9.1
$MNBO_3MAX$	0.04	−9.44	9.8	−8.8	−5.9	−15.5
$CORRO_3MAX$	0.7	0.63	0.65	0.4	0.04	0.25
$RMSEO_3MAX$	21.3	21.4	21.5	32.7	35.2	32.2
σMAX	0.85	0.66	1.09	0.99	0.90	1.22
SOMO35	3925	1552	915	3628	2043	531

GMDD
5, 2083–2138, 2012

Table 7. Same as Table 6 for hourly value and daily mean NO_x concentrations.

NO_x Metrics	ANALY			CLIM		
	Year	JJAS	DJFM	Year	JJAS	DJFM
$MBNO_xH$	−0.19	−2.5	2.2	2.1	−2.1	8.4
$MNBNO_xH$	−1.2	−32.5	13.9	17.0	−28.8	55.1
$CORRNO_xH$	0.46	0.29	0.42	0.14	0.08	0.05
$RMSENO_xH$	13.2	8.8	15.9	20.5	10.6	29.2
σNO_xH	1.20	0.91	1.19	1.64	1.00	1.78
$MBNO_xDM$	−0.12	−2.4	2.2	2.1	−2.1	8.0
$MNBNO_xDM$	−0.67	−32.4	14.6	17.5	−27.9	54.3
$CORRNO_xDM$	0.61	0.43	0.55	0.20	0.09	0.07
$RMSENO_xDM$	9.8	6.1	11.9	17.6	7.6	25.7
σNO_xDM	1.34	0.98	1.30	1.92	1.07	2.08

GMDD
5, 2083–2138, 2012

Table 8. Same as Table 6 for hourly value and daily mean SO_2 concentrations.

SO_2 Metrics	ANALY			CLIM		
	Year	JJAS	DJFM	Year	JJAS	DJFM
$MBSO_2H$	0.39	−0.06	0.79	0.46	−0.06	1.3
$MNBSO_2H$	18.6	−2.9	36.7	24.2	−3.1	50.8
$CORRSO_2H$	0.23	0.15	0.27	0.02	−0.01	−0.01
$RMSESO_2H$	4.3	3.3	4.9	5.5	3.8	7.4
σSO_2H	1.37	1.16	1.53	1.65	1.21	2.05
$MBSO_2DM$	0.37	−0.07	0.79	0.45	−0.07	1.25
$MBNSO_2DM$	17.8	−3.5	35.4	23.2	−4.0	49.7
$CORRSO_2DM$	0.36	0.3	0.4	0.03	−0.01	−0.01
$RMSESO_2DM$	3.2	2.3	3.7	4.5	2.8	6.29
σSO_2DM	1.53	1.35	1.70	1.92	1.47	2.42

GMDD
5, 2083–2138, 2012

Table 9. Same as Table 6 for the PM_{10} daily mean. The calculated statistics are: $MBPM_{10}$ ($\mu g\,m^{-3}$), $MFBPM_{10}$ (%), $MFEPM_{10}$ (%), $CORRPM_{10}$, $RMSEPM_{10}$ ($\mu g\,m^{-3}$) and σPM_{10}.

PM_{10} Metrics	ANALY			CLIM		
	Year	JJAS	DJFM	Year	JJAS	DJFM
$MBPM_{10}$	−8.2	−11.9	−4.6	−5.3	−11.8	4.6
$MFBPM_{10}$	−59.5	−94.6	−33.6	−48.5	97.9	3.5
$MFEPM_{10}$	75.1	96.9	58.0	85.4	101.7	77.4
$CORRPM_{10}$	0.39	0.2	0.48	0.04	−0.04	0.03
$RMSEPM_{10}$	15.1	14.4	15.1	21.6	15.3	28.3
σPM_{10}	0.96	0.49	1.02	1.39	0.60	1.59

GMDD
5, 2083–2138, 2012

Table 10. Seasonnal JJAS statistics obtained with the permutation ANALY-p and CLIM-p. Statistics are the median values of the permutations. The calculated statistics are: mean bias ($\mu g\,m^{-3}$), mean normalized bias (%), correlation coefficient, Root Mean Square Error ($\mu g\,m^{-3}$) and sigma ratio. Statistics are computed for the O_3 hourly value and daily M×8h concentrations.

O_3 Metrics	ANALY-p	CLIM-p	ANALY
MBO_3MAX	−7.9	−5.6	−8.1
$MNBO_3MAX$	−9.1	−6.4	−9.4
$CORRO_3MAX$	0.04	0.07	0.63
$RMSEO_3MAX$	31	34.6	21.4
σO_3MAX	0.67	0.9	0.66
MBO_3H	−2.8	−5.0	−2.9
$MNBO_3H$	−4.4	−7.9	−4.5
$CORRO_3H$	0.33	0.34	0.59
$RMSEO_3H$	30.9	35.1	24.9
σO_3H	0.75	0.98	0.73

GMDD
5, 2083–2138, 2012

Table 11. Same as Table 10 for hourly values and daily average of NO_x and SO_2.

NO_x Metrics	ANALY-p	CLIM-p	ANALY	SO_2 Metrics	ANALY-p	CLIM-p	ANALY
$MBNO_xDM$	−2.5	−2.0	−2.4	$MBSO_2DM$	−0.09	−0.04	−0.07
$MNBNO_xDM$	−32.6	−27.8	−32.4	$MNBSO_2DM$	−2.8	−1.0	−3.5
$CORRNO_xDM$	0.09	0.07	0.43	$CORRSO_2DM$	−0.004	0.005	0.3
$RMSENO_xDM$	7.1	7.6	6.1	$RMSESO_2DM$	2.6	2.7	2.3
σNO_xDM	0.98	1.07	0.98	σSO_2DM	1.39	1.47	1.35
$MBNO_xH$	−2.5	−2.1	−2.5	$MBSO_2H$	−0.08	−0.04	−0.06
$MNBNO_xH$	−32.7	−28.6	−32.5	$MNBSO_2H$	−2.6	−0.68	−2.9
$CORRNO_xH$	0.09	0.07	0.29	$CORRSO_2H$	−0.008	−0.01	0.15
$RMSENO_xH$	9.9	10.5	8.8	$RMSESO_2H$	3.6	3.8	3.3
σNO_xH	0.90	0.98	0.91	σSO_2H	1.18	1.21	1.16

GMDD
5, 2083–2138, 2012

Table 12. Seasonnal JJAS statistics obtained with the permutation ANALY-p and CLIM-p. Statistics represent the median values of the permutations. The calculated statistics are: mean bias ($\mu g\,m^{-3}$), mean fractional bias (%), mean fractional error (%), correlation coefficient, Root Mean Square Error ($\mu g\,m^{-3}$) and sigma ratio. Statistics are computed for the PM_{10} daily mean concentrations.

PM_{10} Metrics	**ANALY-p**	**CLIM-p**	**ANALY**
$MBPM_{10}$	−11.9	−11.8	−11.9
$MFBPM_{10}$	−93.8	−98.6	−94.6
$MFEPM_{10}$	97	102.2	96.9
$CORRPM_{10}$	0.01	0.01	0.2
$RMSEPM_{10}$	14.9	15.1	14.4
σPM_{10}	0.49	0.59	0.49

GMDD
5, 2083–2138, 2012

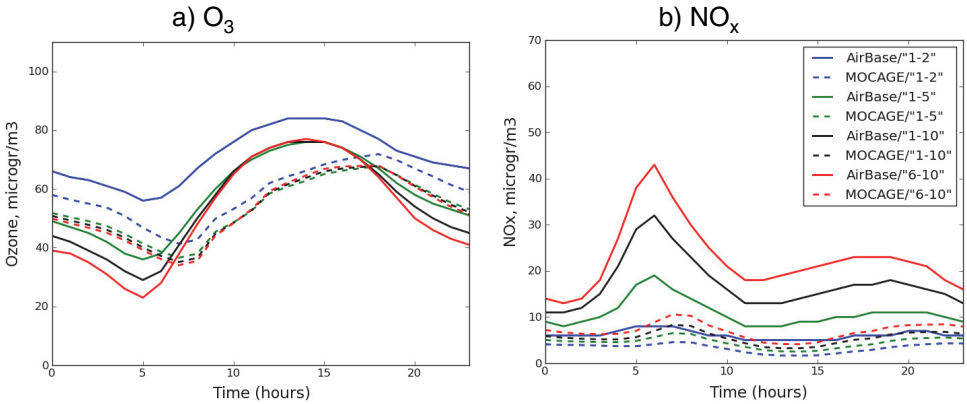

Fig. 1. Mean diurnal cycles of O_3 (left) and NO_x (right), as a function of hour, simulated with MOCAGE (dashed lines) and observed by AirBase averaged for the summer (JJAS) of 2007. Diurnal cycles are represented for the classes 1–2, 1–5, 1–10 and 6–10.

GMDD
5, 2083–2138, 2012

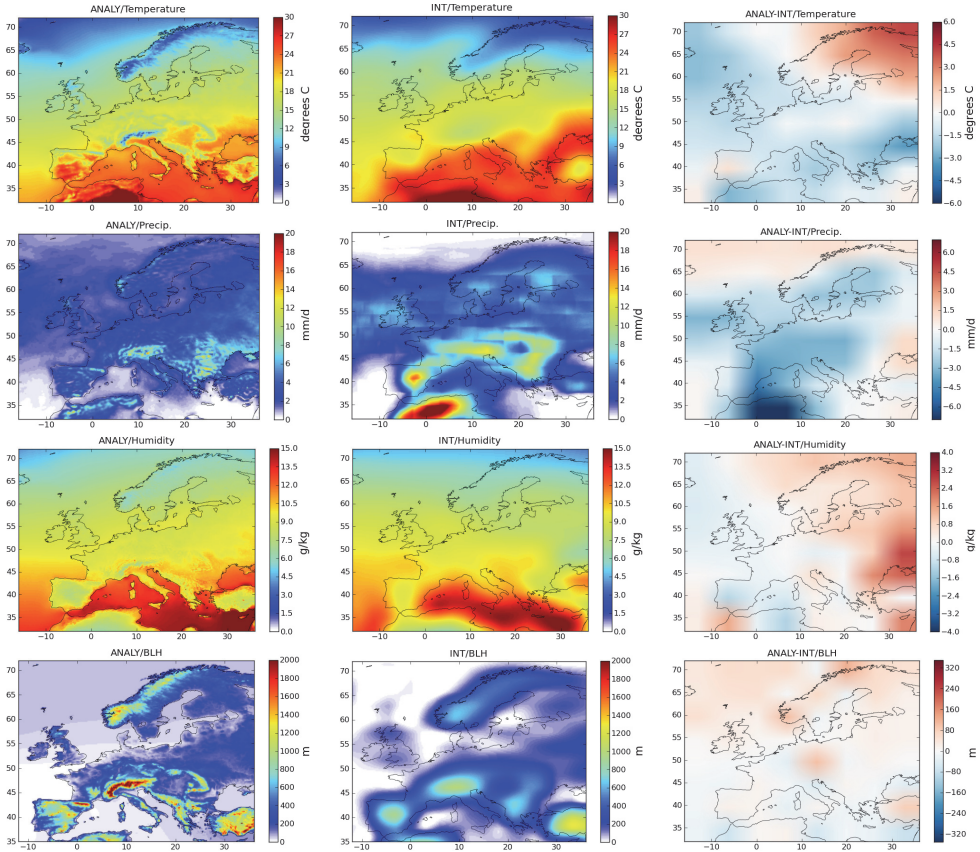

Fig. 2. From top to bottom: average summertime surface temperature (°C), precipitation ($mm\,d^{-1}$), humidity ($g\,kg^{-1}$) and planetary boundary layer height (m) for the summer period (JJAS) of ANALY and INT. Differences between ANALY and INT are shown on the right.

GMDD
5, 2083–2138, 2012

Air quality hindcasts driven by forcings from climate model

G. Lacressonnière et al.

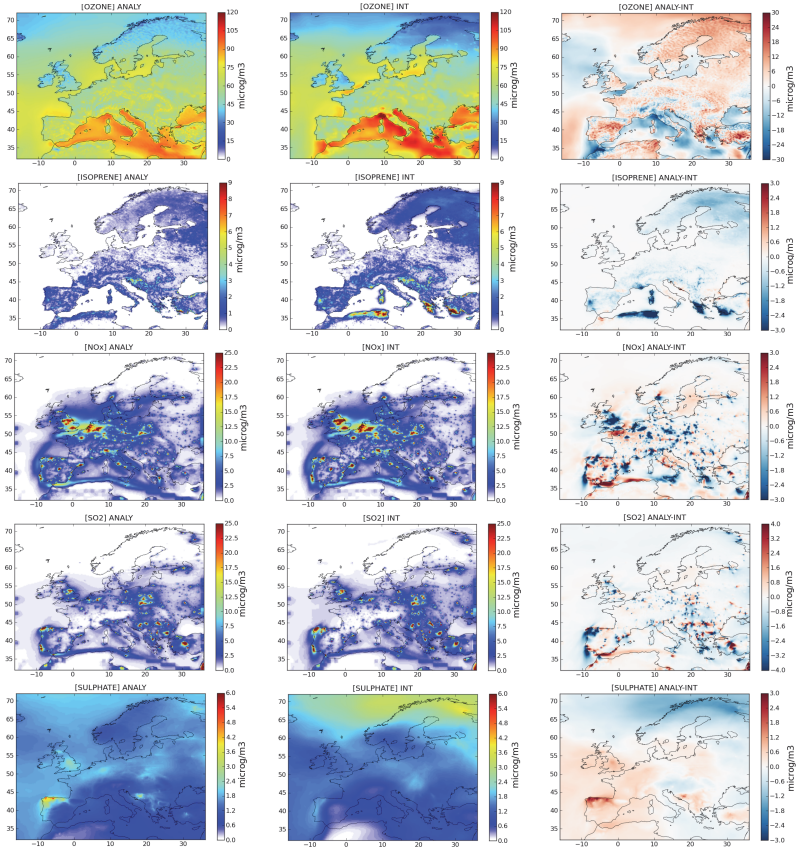

Fig. 3. From top to bottom: simulated average JJAS surface O_3, isoprene, NO_x, SO_2 and sulphate fields for ANALY and INT. Differences between ANALY and INT are shown on the right. Units are in $\mu g\, m^{-3}$.

GMDD
5, 2083–2138, 2012

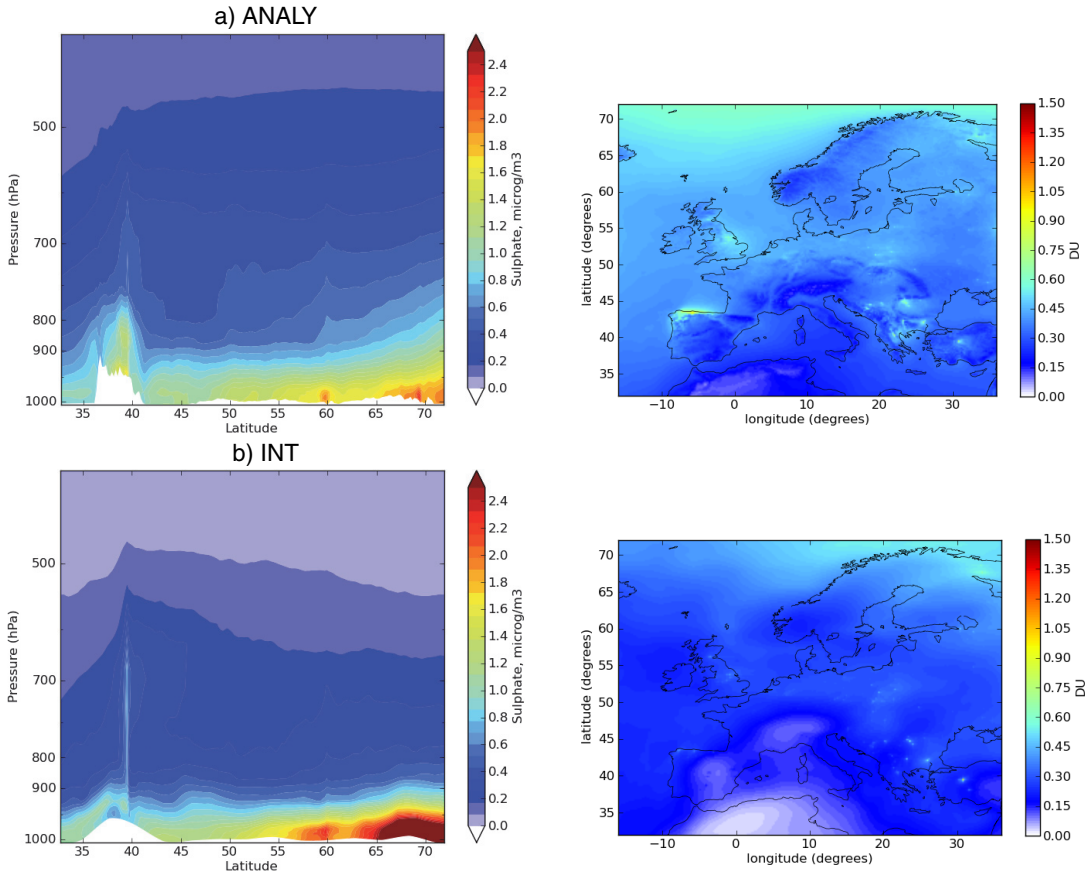

Fig. 4. On the left: latitudinal cross-section at 30° E of sulphate concentrations (μg m^{-3}) averaged for JJAS period. Levels are in hPa. On the right: sulphate column (DU) averaged for the summertime period of ANALY and INT.

GMDD
5, 2083–2138, 2012

Air quality hindcasts driven by forcings from climate model

G. Lacressonnière et al.

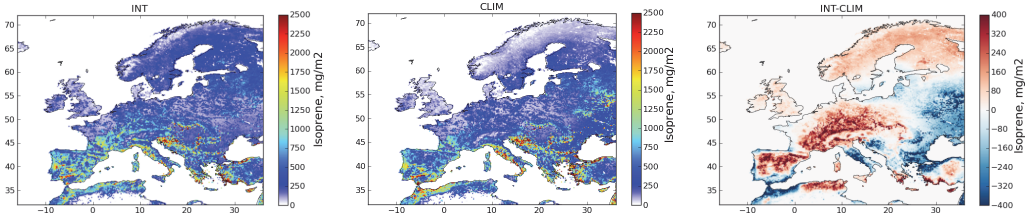

Fig. 5. Emissions of isoprene for the summertime period, averaged for 2004–2008 in the INT (left) and CLIM (middle) simulations. Differences between INT and CLIM are shown on the right figure.

GMDD
5, 2083–2138, 2012

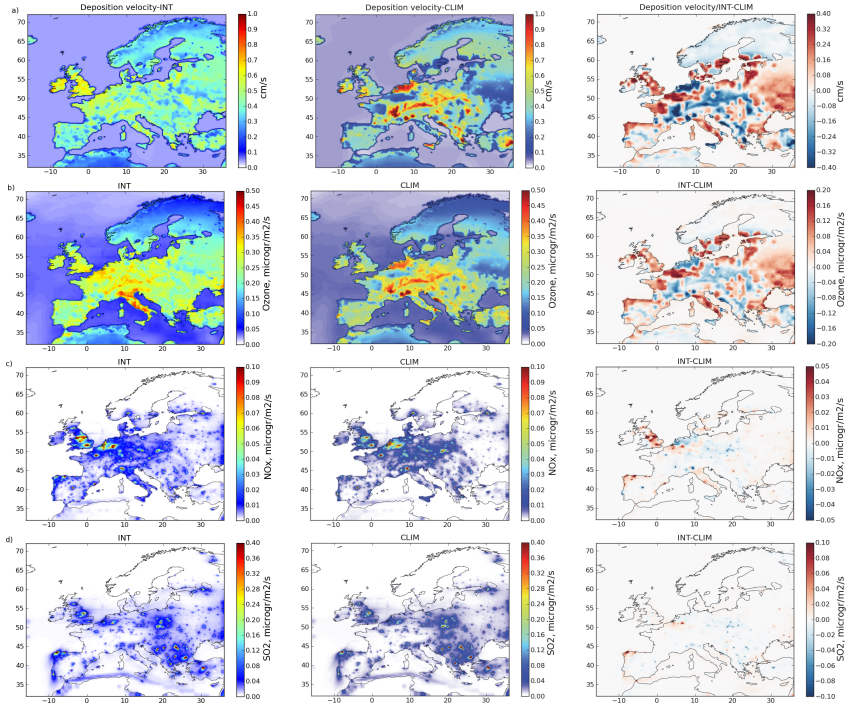

Fig. 6. From top to bottom: **(a)** O_3 deposition velocity averaged for the summertime period simulated by INT (left) and CLIM (middle). Differences between INT and CLIM are shown on the right. For INT, daytime and nightime mean deposition velocities reach 0.57 $cm\,s^{-1}$ and 0.24 $cm\,s^{-1}$, respectively over land (0.06 $cm\,s^{-1}$ and 0.05 $cm\,s^{-1}$ over sea). For CLIM, daytime and nightime mean deposition velocities reach 0.54 $cm\,s^{-1}$ and 0.24 $cm\,s^{-1}$ over land (0.05 $cm\,s^{-1}$ and 0.04 $cm\,s^{-1}$ over sea). **(b)** deposition flux of O_3, **(c)** deposition flux of NO_x and **(d)** deposition flux of SO_2. Deposition flux are in $\mu g\,m^{-3}\,s^{-1}$ and averaged for the summertime of INT and CLIM simulations.

GMDD
5, 2083–2138, 2012

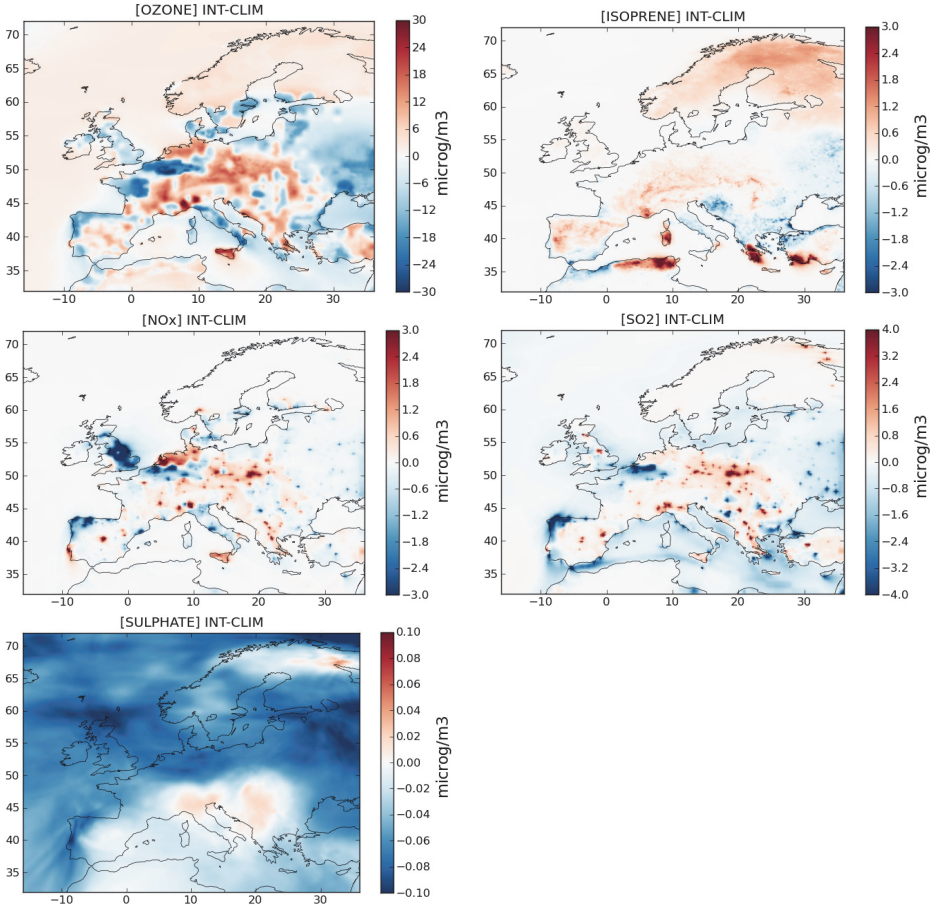

Fig. 7. Differences in simulated average surface O_3, isoprene, NO_x, SO_2 and sulphate fields between INT and CLIM for the summertime (JJAS). Species are in unit of $\mu g\, m^{-3}$.

GMDD
5, 2083–2138, 2012

Air quality hindcasts driven by forcings from climate model

G. Lacressonnière et al.

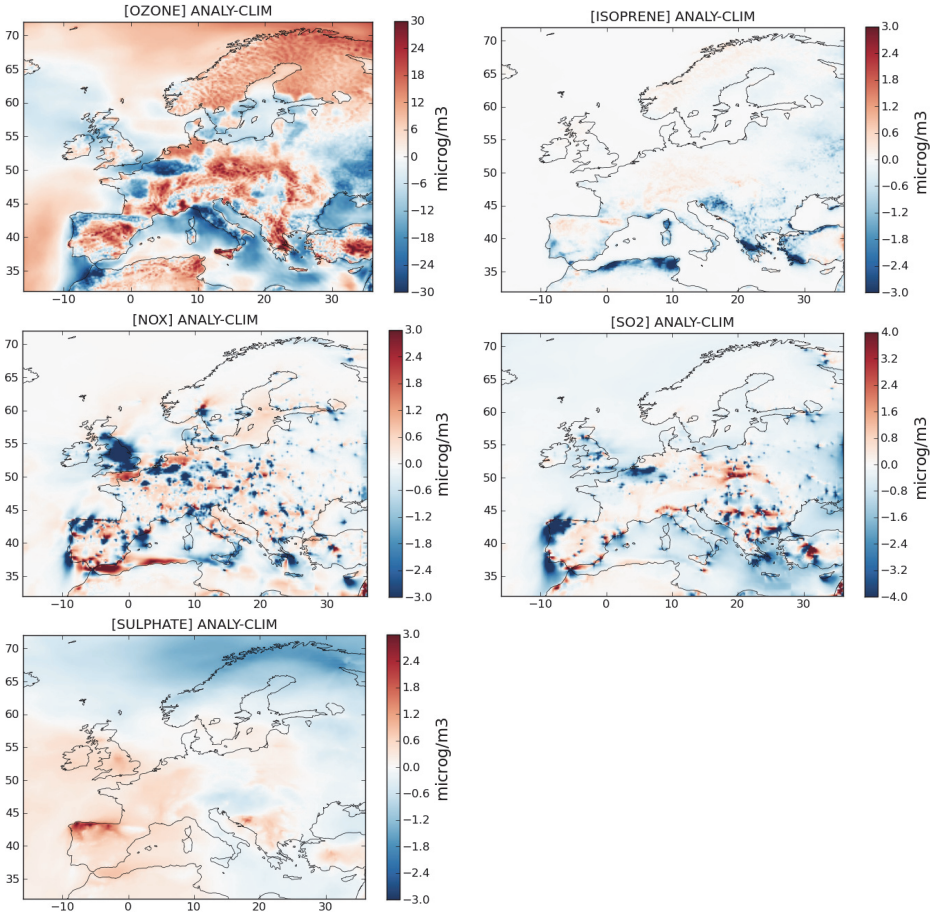

Fig. 8. Same as Fig. 7 but comparing ANALY and CLIM.

GMDD
5, 2083–2138, 2012

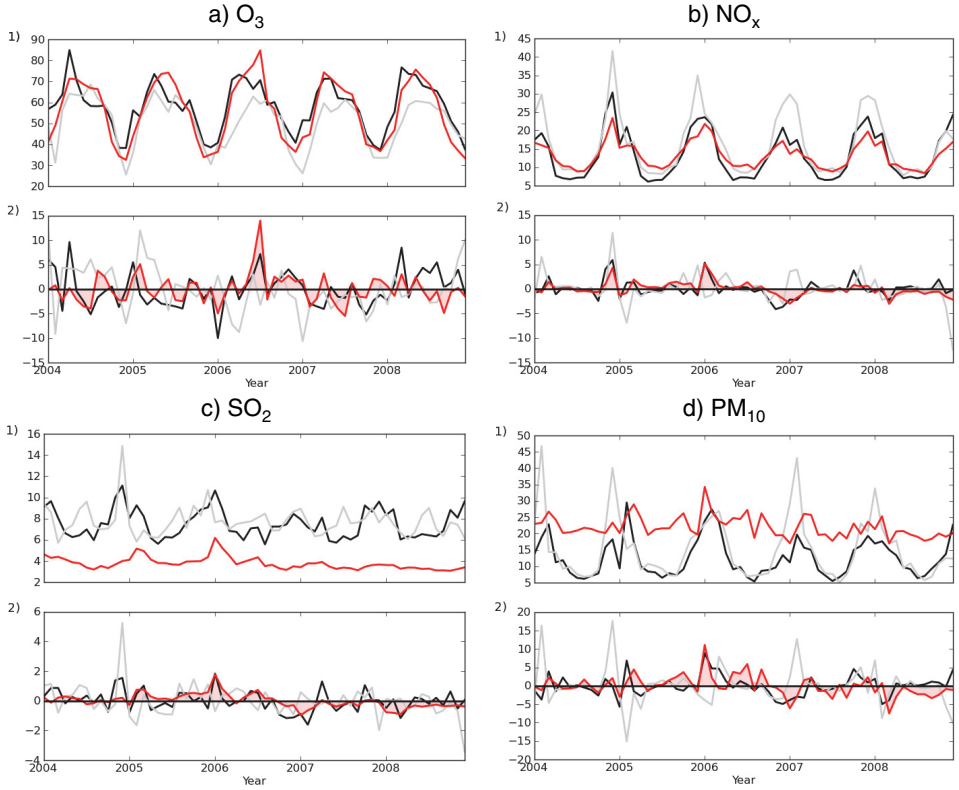

Fig. 9. 1. Simulated (ANALY: black lines; CLIM: gray lines) and measured at the AirBase stations (red lines) time series of monthly mean concentrations of O_3 **(a)**, NO_x **(b)**, SO_2 **(c)** and PM_{10} **(d)**. The time series are plotted from 1 January 2004 to 31 December 2008 and averaged over the European domain. 2. Anomalies calculated when substracting the average annual series from the the time series in 1. Concentrations are in $\mu g\, m^{-3}$.

GMDD
5, 2083–2138, 2012

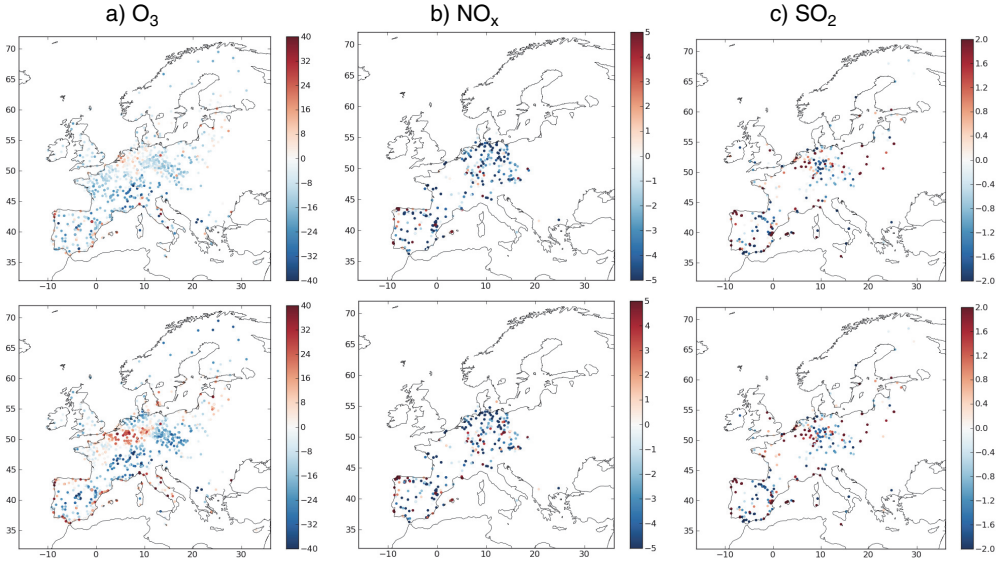

Fig. 10. Spatial distribution of mean bias ($\mu g\,m^{-3}$) for daily M×8h O_3 **(a)**, daily mean NO_x **(b)** and daily mean SO_2 **(c)** concentrations for the average summertime period of 2004–2008. The two rows represent ANALY (top) and CLIM (bottom).

GMDD
5, 2083–2138, 2012

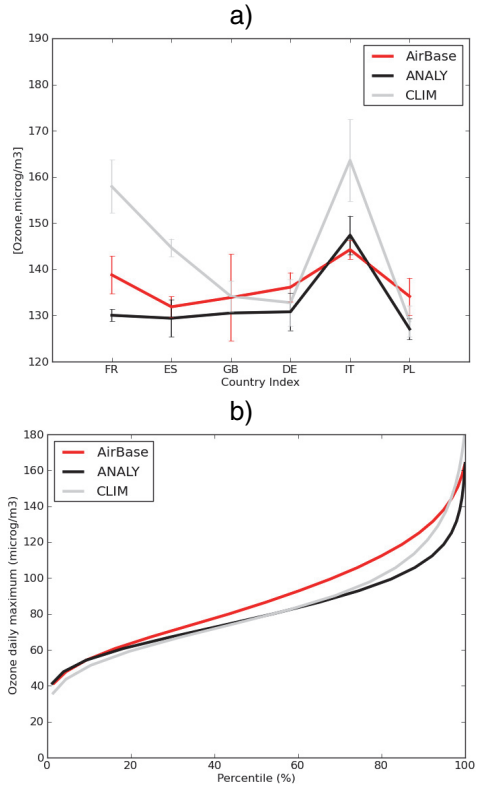

Fig. 11. (a) Summertime average ozone concentrations (over the stations available) above the threshold of 120 µg m^{-3} over 6 European countries. FR is France, ES is Spain, DE is germany, GB is England, IT is Italy, PL is Poland. The standard deviation measuring the inter-annual variability is represented by the vertical bars. **(b)** Distribution of daily O_3 maximum percentiles for AirBase measurements (red) and MOCAGE simulations (ANALY in black, CLIM in gray).

GMDD
5, 2083–2138, 2012

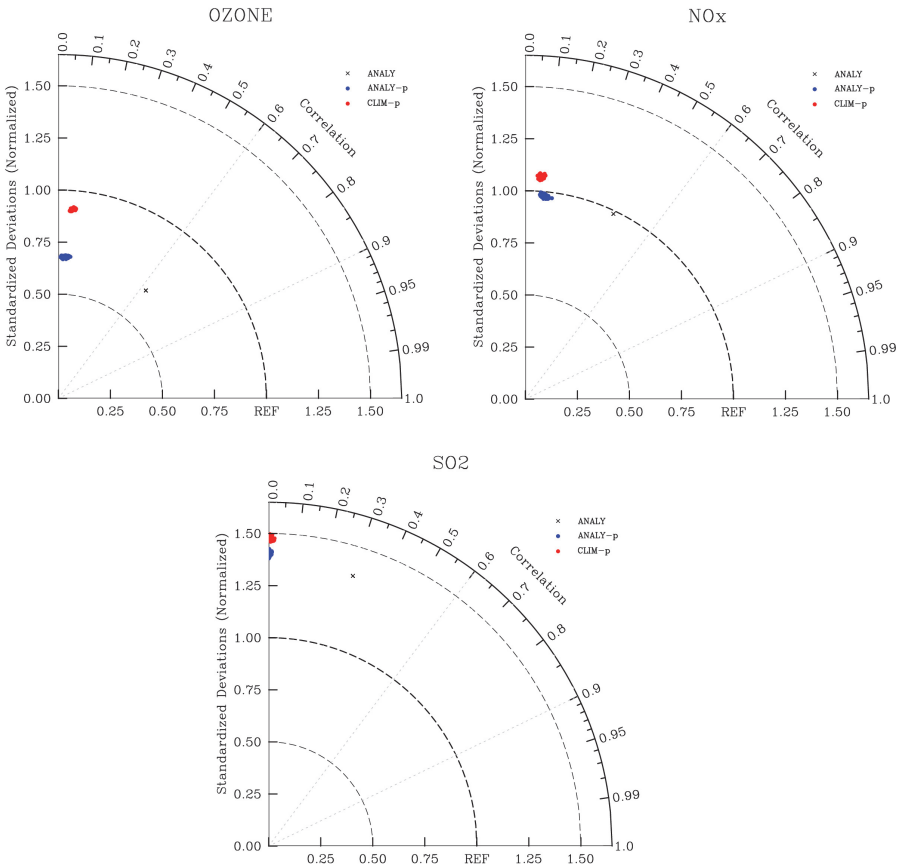

Fig. 12. Taylor plots of the comparison between modelled and observed M×8h O_3 concentrations, daily mean NO_x and daily mean SO_2. The radial distance from the origin corresponds to NSD and *R* corresponds to the azimuthal position.

GMDD
5, 2083–2138, 2012