Climate forced air-quality modeling at urban scale: sensitivity to model resolution, emissions and meteorology

K. Markakis\textsuperscript{1}, M. Valari\textsuperscript{1}, O. Perrussel\textsuperscript{2}, O. Sanchez\textsuperscript{2}, and C. Honore\textsuperscript{2}

\textsuperscript{1}Laboratoire de Meteorologie Dynamique, IPSL Laboratoire CEA/CNRS/UVSQ, Ecole Polytechnique, 91128 Palaiseau CEDEX, France
\textsuperscript{2}AIRPARIF, Association de surveillance de qualite de l’air en Île-de-France, 7 rue Crillon, 75004, Paris, France

Received: 12 November 2014 – Accepted: 3 February 2015 – Published: 20 February 2015

Correspondence to: K. Markakis (konstantinos.markakis@lmd.polytechnique.fr)

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

While previous research helped to identify and prioritize the sources of error in air-quality modeling due to anthropogenic emissions and spatial scale effects our knowledge is limited on how these uncertainties affect climate forced air-quality assessments. Using as reference a 10 yr model simulation over the greater Paris (France) area at 4 km resolution and anthropogenic emissions from a 1 km resolution bottom-up inventory, through several tests we estimate the sensitivity of modeled ozone and PM$_{2.5}$ concentrations to different potentially influential factors with a particular interest over the urban areas. These factors include the model horizontal and vertical resolution, the meteorological input from a climate model and its resolution, the use of a top-down emission inventory, the resolution of the emissions input and the post-processing coefficients used to derive the temporal, vertical and chemical split of emissions. We show that urban ozone displays moderate sensitivity to the resolution of emissions ($\sim 8\%$), the post-processing method ($6.5\%$) and model resolution ($\sim 5$) while annual PM$_{2.5}$ levels are particularly sensitive to changes in their primary emissions ($\sim 32\%$) and the resolution of the emission inventory ($\sim 24\%$) while model horizontal and vertical resolution are of little effect. In addition we use the results of these sensitivities to explain and quantify the discrepancy between a coarse ($\sim 50$ km) and a fine (4 km) resolution simulation over the urban area. We show that the ozone bias of the coarse run (+9 ppb) is reduced by $\sim 40\%$ by adopting a higher resolution emission inventory, by $25\%$ by using a post-processing technique based on the local inventory (same improvement is obtained by increasing model horizontal resolution) and by $10\%$ by adopting the annual emission totals of the local inventory. The bias on PM$_{2.5}$ follows a more complex pattern with the positive bias associated to the coarse run (+3.6 µg m$^{-3}$) increasing or decreasing depending on the type of the refinement. We conclude that in the case of fine particles the coarse simulation cannot selectively incorporate local scale features in order to reduce model error.
1 Introduction

Recent epidemiological findings stress the need to resolve the variability of pollutant concentrations at urban scale. The International Agency for Research on Cancer recently classified outdoor air pollution as a “leading environmental cause of cancer deaths” (Loomis et al., 2013) while new findings reveal that living near busy roads substantially increases the total burden of disease attributable to air pollution (Pascal et al., 2013). Research on future projections of air-quality should be addressed primarily at such scale especially given the fact that the efforts to mitigate air-pollution are more intense in areas where the largest health benefits are observed (Riahi et al., 2011).

Climate and atmospheric composition are related through a series of physical and chemical mechanisms and atmospheric feedbacks. A significant portion of the published literature on this issue uses global scale models to focus on the impact of climate on tropospheric ozone at global or regional scales (Brasseur et al., 1998; Liao et al., 2006; Prather et al., 2003; Szopa et al., 2006; Szopa and Hauglustaine, 2007). More recent studies have integrated advanced chemistry schemes capable of resolving the variability of pollutant concentrations at regional scale, which spans from several hours up to a few days, with chemistry transport models (CTMs) (Colette et al., 2012, 2013; Forkel and Knoche, 2006, 2007; Hogrefe et al., 2004; Katragkou et al., 2011; Knowlton et al., 2004; Lam et al., 2011; Langner et al., 2005, 2012; Nolte et al., 2008; Szopa and Hauglustaine, 2007; Tagaris et al., 2009; Zanis et al., 2011). Global models with a typical resolution of a few hundreds of kilometers and regional CTMs used at resolutions of a few tens of kilometers – and their parameterization of physical and chemical processes make them inadequate for modeling air-quality at urban scale (Cohan et al., 2006; Forkel and Knoche, 2007; Markakis et al., 2014; Sillman et al., 1990; Tie et al., 2010; Valari and Menut, 2008; Valin et al., 2011; Vautard et al., 2007).

The challenge here is how to model climate forced atmospheric composition with CTMs at fine resolution over urban areas, where emission gradients are particularly
sharp, without introducing large errors due to emissions’ and meteorology uncertainties as well as to CTMs numerical resolution. By principle, in the absence of plume in grid parameterization, emissions in CTMs are instantly mixed within the volume of model grid-cells before chemical reaction transport and mixing take place. When the volume of these cells is large compared to the characteristic time scale of these processes, sub-grid scale errors occur such as over-dilution of emissions leading to unrealistic representation of urban scale chemistry such as ozone titration. The resolution of meteorological modeling is another issue: LeRoyer et al. (2014) argue that only high-resolution meteorological modeling can correctly capture the urban heat island, also Flagg and Taylor (2011) showed that high-resolution modeling is very much dependent on the resolution of the surface layer input data.

Another key issue is the representativeness of top-down emission inventories over cities. The starting point of these inventories is emission annual totals for families of pollutants at continental, regional or national scale that are temporally and spatially downscaled based on proxies such as land-use and population data, activity-dependent time profiles and chemical speciation to provide gridded hourly emission fields suitable for modeling with CTMs. It has been shown that these inventories cannot adequately portray the plethora and complexity of the anthropogenic emissions over large cities (Markakis et al., 2010, 2012). In Markakis et al. (2014) we showed that the implementation of bottom-up emissions in a decade simulation over Paris correctly predicted ozone production in the city under a VOC-limited regime as previous studies confirm (Beekman and Derognat, 2003; Beekman and Vautard, 2010; Deguillaume et al., 2008) whereas the use of emission fluxes originating from a regional top-down inventory predicted ozone under NO$_x$-limited conditions. Such a discrepancy is critical when mitigation scenarios are investigated because they may lead to controversy when studying the ozone response in the future. As shown in Markakis et al. (2014) regional scale modeling and the use of top-down emissions can result to higher future reductions than the urban scale modeling using bottom-up emissions. Other challenges stem from the fact that emission projections are mostly based on scenarios developed to represent
changes at global scale and are rarely suited for assessment at regional let alone urban scales. This is because long-term projections are very constrained by the evolution of energy supply and demand, which is a large scale issue. There is no scope in modeling the evolution of emissions over a given urban area over the long term without considering the evolution of more global scenarios.

The major caveat of simulating regional scales at high resolution is the enormous computational demands and that is particularly relevant to climate studies where the simulated periods extend over several decades. To fill the gap between regional and city-scale assessments we need to combine in a single application the advantages of each scale; on one hand the high spatial coverage (but with low resolution) and on the other a good representation of emissions over cities. To achieve this goal we need to understand the major sources of error and their respective impact on climate forced atmospheric composition simulations at urban scale.

This study builds on the previous work of Markakis et al. (2014) where a qualitative comparison was accomplished between an urban (local) and a regional scale simulation over Paris. The aim of the present study is to disentangle modeling errors of climate forced atmospheric composition studies over finer scales due to different factors such as emission and meteorological input as well as model horizontal and vertical resolution. We use as reference run a 10 yr long simulation (1996–2005) over the Ile-de-France region in France (IdF) at 4 km resolution, using the high-resolution (1 km) bottom-up emission inventory of the region’s environmental agency (AIRPARIF, 2012). Boundary conditions for this run are taken from a regional scale simulation at 0.5° over Europe, where ECLIPSE top-down emissions were used (Klimont et al., 2013, 2015).

We carry out several sensitivity tests to quantify the impact of an envelope of effects such as (a) meteorology from a climate model vs. reanalysis data, (b) the spatial resolution of the meteorological input, (c) the air-quality model vertical resolution, especially close to the surface, (d) bottom-up vs. top-down emissions, (e) AIRPARIF vs. EMEP post-processing information (temporal, vertical and chemical split) of emissions to provide appropriate fluxes on the air-quality modeling mesh grid f) the resolution of the
emission input g) the CTM's horizontal resolution. We aim to point out the most influential parameters of model configuration to help improving regional scale climate change assessments.

2 Materials and methods

2.1 Meteorological and air-quality models’ setup

The IdF region is located at 1.25–3.58° E and 47.89–49.45° N with a population of approximately 11.7 million, more than two million of which live in the city of Paris (Fig. 1). The area is situated away from the coast and is characterized by uniform and low topography, not exceeding 200 m above sea level.

In order to simulate air-quality in the study region we employ a dynamical downscaling approach: at first the IPSL-CM5A-MR global circulation model (Dufresne et al., 2013) is used to derive projections of the main climate drivers (temperature, solar radiation etc.) using the RCP-4.5 dataset of greenhouse gas emissions (van Vuuren et al., 2011). Global climate output is downscaled with the Weather Research and Forecasting (WRF) mesoscale climate model (Skamarock and Klemp, 2008) over Europe at a 0.44° horizontal resolution grid (details on these simulations can be found in Kotlarski et al., 2014). For the purpose of the sensitivities presented in the paper we also employ meteorology driven by ERA reanalysis data at two resolutions; 0.11 and 0.44° (Vautard et al., 2013).

Pollutant concentrations at global scale are modeled with the LMDz-INCA chemistry model (Hauglustaine et al., 2004, 2013) forced with RCP-4.5 emissions. These concentration fields are downscaled at regional scale with the CHIMERE (2013a version) off-line chemistry-transport model (http://www.lmd.polytechnique.fr/chimere) in two steps: initially at 0.44° resolution grid (∼ 50 km) over Europe (EEA, 2104) and subsequently at 4 km resolution over the IdF region. The nesting scheme is presented in Fig. 1. CHIMERE is a cartesian mesh-grid model including gas-phase, solid-phase and
aqueous chemistry, biogenic emissions modeling depending on meteorology with the MEGAN model (Guenther et al., 2006), dust emissions (Menut et al., 2005) and resuspension (Vautard et al., 2005). Gas-phase chemistry is based on the MELCHIOR mechanism (Lattuati, 1997) which includes more than 300 reactions of 80 gaseous species. The aerosols model species are sulfates, nitrates, ammonium, organic and black carbon and sea-salt (Bessagnet et al., 2010) and the gas-particle partitioning of the ensemble Sulfate/Nitrate/Ammonium is treated by the ISORROPIA code (Nenes et al., 1998) implemented on-line in CHIMERE. CHIMERE has been benchmarked in the past in a number of model inter-comparison experiments (see Menut et al. (2013) and references therein).

For the reference run at urban scale (hereafter REF), we use the same model setup as in Markakis et al. (2014): the modeling domain has a horizontal resolution of 4 km and consists of 39 grid cells in the west–east direction, 32 grid cells in the north–south direction and 8 $\sigma$-p hybrid vertical layers from the surface (999 hPa) up to approximately 5.5 km (500 hPa) with the surface layer being 25 m thick. The configuration of the reference run represents the best compromise between local scale emission data and the high computational demand of a long-term simulation at fine resolution.

2.2 Climate and emissions

The RCP-4.5 long-term scenario of greenhouse gases used as global scale predictor of present-time climate displays a 20% GHG emission reduction for Europe, constant population at about 575 million inhabitants and mid-21st century change in global radiative forcing by 4 W m$^{-2}$, increasing to 4.5 W m$^{-2}$ by 2065 and stabilizing thereafter. The RCP-4.5 also includes century-long estimates of air pollutant emissions, including aerosols and was used to drive the global scale LMDz-INCA simulations.

The regional scale simulations for the present-time (2010) employ an emission database developed in the framework of the ECLIPSE (Evaluating the Climate and Air Quality Impacts of Short-Lived Pollutants) project (Klimont et al., 2013, 2015) implementing emission factors from GAINS (Amann et al., 2011). Present-time emissions
are compiled by the International Institute for Applied Systems Analysis (IIASA) and as regards Europe they include the results of the work undergone in the UNECE Convention on Long-Range Transboundary Air Pollution (CLRTAP). The emission estimates are available at a 0.5° × 0.5° resolution grid.

Present-time (2008) emission estimates for the IdF region are available in hourly basis over a 1 km resolution grid. This emission inventory is compiled by the Ile-de-France environmental agency and combines a large quantity of city-specific information (AIRPARIF, 2012) based on a bottom-up approach. The spatial allocation of emissions is either source specific (e.g. locations of point sources) or completed with proxies such as high-resolution population maps and a detailed road network. The inventory includes emissions of CO, NO\textsubscript{x}, Non-methane Volatile Organic Compounds (NMVOCs), SO\textsubscript{2}, PM\textsubscript{10} and PM\textsubscript{2.5} with a monthly, weekly and diurnal -source specific- temporal resolution. Emissions from point sources are included as area emissions in the model and the grid cells containing those sources adopt a vertical distribution across model layers which varies in time dependent on several meteorological variables such as temperature and wind inputted in a plume-rise algorithm (Scire et al., 1990). Consequently the distribution of emissions among different activity sectors reveals that in the IdF region the principal emitter of NO\textsubscript{x}, on annual basis, is the road transport sector (50 %), for NMVOCs the use of solvents (50 %) and for fine particles the residential sector (37 %).

The raw data of the 1 km resolution emissions were aggregated to the 4 km resolution grid.

2.3 Data and metrics for model evaluation

Model results from the different sensitivity runs are compared against observational data for O\textsubscript{3}, NO, NO\textsubscript{2} and PM\textsubscript{2.5}. Pollutant concentrations measured at 29 sites of the air-quality network of AIRPARIF (17 urban, 4 suburban and 8 rural) are compared to first-layer modeled concentrations on the grid-cells containing the corresponding monitor site. To benchmark model performance we use the skill score $S$ which is based...
on the equations of Mao et al. (2006):

\[ S = \frac{1}{2} \left( 1 - \left| \frac{\text{BIAS}}{\text{MGE}} \right| + \left| \frac{\text{MGE}}{\text{RMSE}} \right| \right) \]  

(1)

where MGE represents the absolute mean gross error and RMSE the root mean square error. A skill score close to 1 is indicative of an unbiased model with no significant errors present but in the case of biased results this rating masks the information on the magnitude of the bias and the corresponding error. For this reason alongside \( S \) we employ the mean normalized bias (MNB) and mean normalized gross error (MNGE) as regards ozone evaluation and the Mean Fractional Bias (MFB) and Mean Fractional Error (MFE) as regards PM\(_{2.5}\) (EPA, 2007).

We extract these metrics from the daily concentration values and not the decade average bearing in mind that this is not typical for runs forced by climate simulations but for operational forecast evaluation. We should note here, that it is reasonable to expect lower scores than those achieved in operational forecast analysis due to the presence of climate biases (Colette et al., 2013; Menut et al., 2013a). As in Markakis et al. (2014) we aim to evaluate our simulations by utilizing metrics that are time averaged on a scale finer than a climatological one.

2.4 Description of the sensitivity simulations

Through a number of test cases we study the ability of the model in predicting present-time decadal air-quality with respect to emission and meteorological input as well as model horizontal and vertical resolution. For that purpose we conduct five sets of 10 yr long simulations (1996–2005) over a 4 km resolution grid covering the IdF region (see Table 1). In all our comparisons we use as a measure of sensitivity of modeled ozone and PM\(_{2.5}\) the absolute difference between the mean of daily averaged concentrations (|\( \Delta c \)|) as well as the absolute change in the skill score \( S \). For ozone we also compare the MNB, MNGE and for PM\(_{2.5}\) the MFB and MFE. All scores are calculated to represent an average of all urban, suburban or rural stations. For PM\(_{2.5}\) for which only
observations from urban stations are available we represent the results for summer, winter and in annual basis for the urban stations.

The first sensitivity case focuses on the climate bias due to the meteorological forcing. It is well established that ozone and certain particulate matter species are sensitive to temperature changes (Fiore et al., 2012; Im et al., 2011, 2012; Jacob and Winner, 2009; Megaritis et al., 2014). Menut et al. (2003) using an adjoint model studied the sensitivity of ozone concentrations at the afternoon peak to numerous model processes and inputs for a typical summer episode in Paris and found that temperature and wind speed was the most influential parameters to the observed changes. For our test we utilize meteorological input that stems from a WRF run employing ERA40 reanalysis data over a 0.44° resolution regional scale grid (ERA05) and compare with the REF simulation utilizing climate model meteorology. Both configurations share identical emission inventories (AIRPARIF) and vertical resolution (8σ-p layers). Modeled meteorological fields are further interpolated over the 4 km-resolution IdF grid for the air-quality simulation. We note here, that interpolating the 0.44° resolution meteorology over the 4 km resolution CHIMERE grid adds a source of uncertainty in modeled pollutant concentrations, but due to the flat topography of the area and as shown in previous research studies in the same region, increasing the resolution of the meteorological input does not improve model performance (Menut et al., 2005; Valari and Menut, 2008). To study the impact of the resolution of the input meteorology here, we conduct a second sensitivity run where meteorological input stems from a WRF simulation using ERA40 reanalysis data over a finer resolution mesh with grid spacing of 0.11° (ERA01) and compare with the ERA05 run.

The third sensitivity test addresses the issue of model’s vertical resolution (VERT). A previous sensitivity analysis conducted with the same air-quality model showed only small changes in modeled ozone and PM$_{10}$ concentrations over the IdF region due to increase in model vertical resolution (Menut et al., 2013b). On the other hand Menut et al. (2003) showed that vertical diffusivity was one of the most influential parameters to the observed daily peak concentrations of ozone for a typical summertime episode in
IdF. Here, we undertake a similar analysis but in a climate modeling framework, where enhanced meteorological bias is expected. VERT implements a 12 vertical $\sigma$-p layers instead of 8. The major difference between the two configurations (REF vs. VERT) is not the number of layers but the depth of the first model layer, which is reduced from 20 to 8 m in VERT.

The fourth sensitivity case estimates the discrepancy in modeled ozone and PM$_{2.5}$ concentrations between two runs where emission totals stem from different inventories, namely the local AIRPARIF inventory and the ECLIPSE regional-scale dataset. In Menut et al. (2003) it was shown that due to the surface emissions, ozone concentrations in the afternoon peak hour had the second largest sensitivity after meteorology. In Markakis et al. (2014) we compared the two approaches as for their ability to correctly represent ozone photo-chemical production under typical anticyclonic summer conditions and also found important differences. In the present work we push the analysis a step further and quantify model response to the emission input over longer timescales. For this purpose we compile a new 4 km resolution emission dataset over the IdF domain (ANN) in which annual emission fluxes match the ECLIPSE emissions ($0.5^\circ$ resolution) but are downscaled spatially and temporally to obtain 4 km-resolution and hourly emissions based on the local scale information implemented in the bottom-up approach of the AIRPARIF emission inventory. The same approach is applied on the chemical speciation of the inventory’s pollutants to obtain emissions for all the species required for the air-quality simulation chemical mechanism. Therefore the only difference amongst the two runs stem from the use of different annual quantified emission fluxes for the region (Table 1). To give a sense of the discrepancies between the two inventories over the IdF region we compare the annual domain-wide fluxes of NO$_x$, NMVOCs and PM$_{2.5}$ (Fig. 2). NMVOCs emissions are considerably higher in the ECLIPSE inventory while NO$_x$ emissions are lower. In terms of photochemical ozone production, this makes the ECLIPSE inventory more favourable of NO$_x$-limited conditions than the bottom-up AIRPARIF inventory, which is consistent with the findings of Markakis et al. (2014). Fine particles emissions are 2.4 times more in ECLIPSE,
which probably stems from the use of a population proxy to spatially allocate winter-time emissions from wood-burning. We note here, that the interest of comparing the two emission inventories is strictly to quantify the added value of implementing local scale information in city-scale climate studies and not by any means to compare qualitatively the two datasets. It is clear that ECLIPSE dataset is not meant to accurately represent emissions at such fine scale.

In the fifth sensitivity case we study the impact of the post-processing methodology e.g. the process followed in order to split the annual emission totals into hourly emission fluxes for all the species and vertical layers required by the air-quality model. Menut et al. (2012a) showed that model performance improves when time-variation profiles developed on the basis of observations are applied for the temporal allocation of emissions instead of the EMEP coefficients. Mailler et al. (2013) found that model results are highly sensitive to the coefficients used for the vertical distribution of emissions. For this test emission totals must match between the two emission datasets. We compile a new emission dataset (POST) where the ECLIPSE annual totals are spatially (both horizontally and vertically) and temporally downscaled on the 4 km-resolution IdF grid. This procedure is based on coefficients extracted from the ECLIPSE post-processed inventory which in turn derive from the EMEP model. Comparing between the POST and ANN runs (Table 1) we can model the impact of integrating a bottom-up approach in regional emission modeling on pollutant concentrations.

Finally the impact of model horizontal resolution is a crucial issue for air-quality modeling. As regards urban ozone there are plentiful studies on the effect of model resolution refinement with an overall tendency to show improvement of the model’s quality when increasing resolution from about 30–50 to 4–12 km (Arunachalam et al., 2006; Cohan et al., 2006; Tie et al., 2010; Valari and Menut, 2008). On the other hand reports are scarce for fine particles: Punger and West (2013) show that increasing the resolution from 36 to 12 km improved the 1 h daily maximum concentrations but not the daily average, Stroud et al. (2011) reported better agreement of fine particles of organic origin with measurements from a modeling exercise at a 2.5 km resolution domain over
a 15 km resolution domain while Queen and Zhang (2008) also show improvement but their results include the effect of increasing the resolution of the meteorological input as well. Valari and Menut (2008) showed that the impact of the resolution of emissions on modeled concentrations of ozone may be higher than the model resolution itself. These question has not yet been raised in the framework of climate driven atmospheric composition modeling at the local scale. In our study we disentangle the impact of the resolution of the emission dataset used as input for the air-quality simulation from the effect of model resolution itself by conducting two more tests. In the first test we employ the 0.5° resolution simulation (REG hereafter) from which all aforementioned simulations take their boundary conditions. We also compile the AVER database which uses as a starting point the modeled concentrations at 4 km resolution from the POST run spatially averaged over the 0.5° grid-cells of the REG resolution mesh. REG vs. AVER (see Table 1) can provide information on the influence of model resolution while comparing AVER against POST provides the sensitivity to the resolution of the emission inventory only.

3 Model evaluation

3.1 Evaluation of present-time meteorology

There are three WRF simulations involved in the study: (i) climate model driven meteorology downscaled from a global scale climate model (MET_CLIM), (ii) meteorology from reanalysis datasets at 0.5° resolution (MET_ERA05) and (iii) meteorology downscaled from reanalysis data at 0.11° (MET_ERA01). In this section we present a short evaluation of these datasets comparing model results against surface observations from seven meteorological monitoring sites existing in the domain. We note here, that from these monitors only one is located inside the highly urbanized city of Paris. A thorough evaluation of the reanalysis dataset in Europe may be found in Menut et al. (2012b).
The mean wintertime (DJF) and summertime (JJA) modeled and observed daily average values are compared for four different meteorological variables relevant for air-quality, namely 2 m temperature, 10 m wind speed, relative humidity and total precipitation (Table 2). A strong positive bias is observed in modeled wind speed for both MET_CLIM and MET_ERA05 meteorology especially during the winter period. Such a bias, consistent with previous studies (see e.g. Jimenez et al. (2012) for WRF or Vautard et al. (2012) for other models), is expected to enhance pollutants' dispersion and lead to less frequent stagnation episodes. The bias is stronger for the MET_CLIM dataset than for the MET_ERA05. A systematic wet bias in both summertime and wintertime precipitation is observed for the two datasets. This can significantly reduce PM$_{10}$ concentrations through rain scavenging (Fiore et al., 2012; Jacob and Winner, 2009). MET_ERA05 fields provide a better representation of precipitation especially in wintertime where the bias is reduced by a factor of more than 2 compared to MET_CLIM. Summertime temperature is adequately represented in the climate dataset whereas a wintertime weak cold bias ($-0.3^\circ$C) is observed. A strong hot bias during the winter is found for the reanalysis meteorology. A warmer climate can increase ozone formation through thermal decomposition of PAN releasing NO$_x$ (Sillman and Samson, 1995). RH is generally well represented in both cases.

Finally we notice that the finer resolution reanalysis dataset (MET_ERA01) is not able to reduce the observed domain-wide biases of the coarse meteorological run with the exception of specific locations such as the Montsouris station in Paris where the bias in wintertime precipitation and wind speed bias is reduced by 22 and 40% respectively.

3.2 Evaluation of the reference simulation (REF)

Mean modeled daily surface ozone and the daily maximum of 8h running means (MD8 h) are compared against surface measurements in urban, suburban and rural stations (Fig. 3a). The results presented are averaged over the ozone period (April–August). We also use odd oxygen $O_x = O_3 + NO_2 - 0.1 \times NO_x$ (Sadanaga et al., 2008)
as an indicator of the efficiency of the model to represent photochemical ozone build-up. Contrary to O$_3$, the concentration of O$_x$ is conserved during the fast reaction of ozone titration by NO and is therefore, a useful metrics for the evaluation of the photochemical ozone build-up by ruling out titration near high NO$_x$ sources (Vautard et al., 2007).

The model performs well in the urban areas capturing the mean daytime ozone levels (bias +1.8 ppb) while O$_x$ is also accurately represented with an underestimation of only 4.1% illustrating the efficiency of the model to reproduce both daytime formation and titration of urban ozone. The bias in daytime average is smaller and less than 1 ppb. The O$_x$ bias in daily averages is similar to the daytime one, suggesting underestimation of nighttime titration. This is consistent with other studies using CHIMERE (Van Loon et al., 2007; Vautard et al., 2007; Szopa et al., 2009). Model benchmark ratings show a high skill score (0.78) while MNB and MNGE are +20.6 and 38.9 respectively.

We observe an overestimation of mean daytime suburban ozone (+5 ppb). The small bias in O$_x$ (+0.6 ppb) suggests that the problem stems from the representation of local titration and more specifically daytime titration; the daily average ozone bias drops to +3.9 ppb while O$_x$ is accurately represented in this case (−0.2 ppb). Suburban stations present the lowest skill score (0.63) compared to urban and rural. Model performance over rural stations is adequate, with an overestimation in mean daily ozone of 8.2% (bias = +2.8 ppb) and a good skill score (0.73). We identified two major downwind locations in the IdF domain and found that they represent the lowest biases (less than 0.1 and 1.1 ppb for the south-west and north-east directions respectively). The bias of the daytime average reaches +2.1 ppb.

Ozone daily maxima in the urban and rural stations are underestimated by 10% (−4.2 ppb) and 7% (−3.2 ppb) respectively but we consider the magnitude of the underestimation small given the climate framework of the simulation. Daily average ozone is better represented than daily maxima, highlighting model sensitivity to accumulated errors (Valari and Menut, 2008). Modeled peak concentrations are particularly sensi-
tive to temperature compared to the daily averages as shown in Menut at al. (2003). This could also be due to the fact that 4 km is still an insufficient model resolution.

The evaluation of PM$_{2.5}$ (Fig. 3b) shows a good representation of daily average levels during wintertime where the highest annual concentrations are presented (bias less than 1 µg m$^{-3}$). In annual basis the bias is also small while a larger underestimation is predicted for the summertime season (bias = 2.8 µg m$^{-3}$). The latter can be due to underestimation of summertime emission fluxes (resuspension emissions are not considered in our simulations) and underestimation of secondary organic aerosols formation (Hodzic et al., 2010; Markakis et al., 2014; Solazzo et al., 2012). The overestimation in wind and precipitation also contributes to the observed PM underestimation. Wintertime and annual statistics show a high skill score. Interestingly in wintertime and in annual basis the site located in downtown Paris presents the lowest bias (< 0.3 µg m$^{-3}$). Overall the results indicate that the fine scale setup is able to predict the main patterns of ozone and fine particle pollution in the area.

4 Sensitivity cases

4.1 Sensitivity to climate model driven meteorology (REF vs. ERA05)

The goal of this case study is to estimate the discrepancy between an air-quality model run where regional meteorology is downscaled with WRF from reanalysis data (ERA05) and a simulation where meteorology is downscaled from a global scale climate model (REF). The wet bias in MET_CLIM meteorology is significantly reduced with meteorology from reanalysis data (Sect. 3.1). This is expected to have a significant role in the modeled PM concentrations. Another influential factor is the colder bias found in summertime temperature in the MET_ERA05 dataset. This may lead to decrease in the reaction rates, less biogenic emissions and consequently to less ozone. The lower bias in 10 m wind speed under MET_ERA05 is bound to lead to less dispersion and higher surface concentrations. We also compare the average modeled boundary layer
height (PBL) for the summer and winter periods between the two datasets: PBL is reduced by 5 and 12% in summer and winter respectively when reanalysis data are used instead of climate model output. This may result in less dilution of emissions and therefore higher surface concentrations for primary emitted species, such as PM and NO$_x$.

Comparing the results of the two air-quality model runs for ozone (Fig. 4a and Table 3) we find only a small sensitivity of ozone to using meteorology from a climate model or reanalysis data over all three types of monitor sites, urban, suburban and rural (|Δc| ~ 1 ppb or 3.4%) suggesting a small improvement of model performance with the reanalysis dataset which stems from the fact that titration is more realistically represented in ERA05 (the difference is O$_x$ between the two runs is negligible). The response of urban daily maximum values to the meteorological dataset is also negligible (|Δc| = 0.1 ppb or 0.3%).

Wintertime PM$_{2.5}$ concentrations, on the contrary show a large sensitivity to the meteorological dataset. The change in the daily average is 3.1 µg m$^{-3}$ (17.6%) while summertime levels remain unchanged (Table 3). Focusing on the annual averages, the small underestimation observed in the REF run turns into small overestimation in the ERA05 run (|Δc| = 1.4 µg m$^{-3}$ or 9.4%). The use of reanalysis data leads to a strong overestimation of wintertime concentrations (Fig. 4b), which stems directly from the reduction (and improvement) of precipitation by a factor of 2 in the meteorology from reanalysis. This leads to the conclusion that the small bias observed in the REF simulation during wintertime (Fig. 4b) could be due model error compensation such as unrealistically high precipitation and possible inhibition of vertical mixing or overestimation of wintertime emissions. The scores suggest a slight deterioration in model performance when passing from meteorology from a climate model to reanalysis meteorology in both winter and summer but improvement when focusing on the annual statistics.
We conclude that using climate model driven meteorology has a small impact on modeled ozone whereas larger sensitivity is observed for wintertime PM$_{2.5}$ levels due to modeled precipitation.

4.2 Sensitivity to the resolution of the meteorological input (ERA01 vs. ERA05)

Here we model the sensitivity of modeled ozone and PM$_{2.5}$ concentrations to the resolution of the meteorological input (Fig. 5 and Table 3). Daily average ozone shows a very weak response over urban and rural sites ($|\Delta c| < 0.4$ ppb or $< 0.8\%$) and daily urban maxima improve slightly with the ERA01 run ($|\Delta c| = 0.4$ ppb or 1%). At the suburban area the impact, though small ($|\Delta c| = 1.4$ ppb or 4.3%), is definitely higher than over urban or rural sites. O$_x$ change at the suburban area (not shown) is much weaker compared to ozone ($|\Delta c| \sim 0.5$ ppb or 1.2%) showing that the increase in the resolution of meteorology has an impact on the representation of ozone titration leading to improved model performance. Skill score over suburban sites increases by 9% while NMB improves by 22% from 26.1 in ERA05 to 20.3 in ERA01. Interestingly, the response of suburban ozone to the resolution of the meteorological input is the strongest modeled sensitivity for this variable amongst all the studied cases.

Weak sensitivities are modeled for PM$_{2.5}$ (Table 3) during summertime ($|\Delta c| = 0.3\mu$gm$^{-3}$ or 3.4%) and on annual basis ($|\Delta c| = 0.6\mu$gm$^{-3}$ or 4%), but stronger during the winter season ($|\Delta c| = 1.3\mu$gm$^{-3}$ or 6.8%). In fact, wintertime statistics suggest that model bias actually increases with the refinement of the meteorological grid as a consequence of the reduced modeled precipitation (less scavenging) and wind speed (weaker dispersion) in MET_ERA01 compared to the climate model driven meteorology (Sect. 3.1). Again this points to the same error compensation scheme described in the REF vs. ERA05 comparison (Sect. 4.1).

We conclude that the resolution of the meteorological input has a small impact on modeled ozone while moderate sensitivity is observed for suburban ozone and wintertime PM$_{2.5}$.
4.3 Sensitivity to the resolution of the vertical grid (REF vs. VERT)

This study addresses the impact of the resolution of the vertical mesh and more specifically of the thickness of the first model layer, on modeled ozone and PM$_{2.5}$ concentrations (Fig. 6). Mean daily ozone is practically insensitive to the refinement of the vertical mesh at the urban, suburban and rural areas (Table 3). Similarly maximum ozone at the urban area changes by only 0.5 ppb (1.4 %) with increased bias in the VERT run. Changes in summertime and annual modeled PM$_{2.5}$ concentrations are also small, while the wintertime daily average shows some weak sensitivity ($|\Delta c| = 0.5 \mu g m^{-3}$ or 2.2 %). All scores are hardly affected.

Interestingly, the impact of the refinement of the vertical grid on daily averaged O$_x$ is much stronger that on ozone: O$_x$ changes by 0.9 ppb in the urban and suburban areas. The change in O$_x$ is reasonable since in VERT, NO$_x$ emissions are released within a surface layer thinner by 60 % compared to REF (from 20 to 8 m) leading to higher NO$_x$ concentrations. That should normally affect titration which is the driver of urban ozone concentrations. The fact that ozone remains insensitive to the change in NO$_x$ concentrations suggests that some other modeled processes counteracts titration. To further investigate this issue we study the change in dynamical processes such as vertical mixing and dry deposition. We extract the vertical diffusion coefficient $K_z$ ($m^2 s^{-1}$) and dry deposition rates ($g m^{-3}$) for ozone, NO$_2$ and PM$_{2.5}$ for all grid cells that include an urban monitor site and looked how modeled sensitivities change as a function of these parameters (Fig. 7).

NO$_2$ concentrations increase with the refinement of the first model vertical layer for all vertical mixing conditions (Fig. 7a). However it is only under low vertical mixing ($1 < K_z < 5 m^2 s^{-1}$) that ozone sensitivity becomes positive (Fig. 7b). Under stronger turbulence ($K_z > 5 m^2 s^{-1}$), the 12-layer setup leads to higher first-layer NO$_2$ concentrations stronger titration leading to negative values for ozone sensitivity (such conditions account for the 93 % of the simulated period). On the other hand the refinement of the vertical mesh primarily affects NO$_2$ deposition rates which accelerate by 14.3 %
but leaving ozone deposition rates unaffected. We may assume that under low mixing conditions, the increased deposition rate of NO$_2$ slows down the increase in NO$_2$ concentration due to the emission effect and dynamical processes become more influential than titration. As a result the surface layer is enriched in ozone by getting mixed with air from higher atmospheric layers (Menut et al., 2013b).

For almost the entire $K_z$ range, PM$_{2.5}$ concentrations increase with VERT (Fig. 7c). This is due to the fact that emissions are released in smaller volumes as discussed above. On the other hand, here too, the refinement of the vertical resolution enhances deposition rate. These two conflicting effects explain the small impact of model vertical resolution on PM$_{2.5}$ concentrations.

Both ozone and PM$_{2.5}$ sensitivities to the refinement of the vertical mesh are small. Our analysis suggests that in both cases this is the result of two competing processes, either titration against vertical mixing (ozone) or emission vs. deposition (PM$_{2.5}$). Although in the Ile-de-France area (low topography) the overall effect is insignificant, it may not be the case in other regions with more complex topography.

4.4 Sensitivity to the annual emission totals (REF vs. ANN)

This case study compares modeled concentrations between two runs where annual emission totals stem from either the AIRPARIF inventory (REF) or the ECPLISE dataset (ANN). Changes in modeled urban daily average ozone concentrations are small ($|\Delta c| = 0.8$ ppb or 2.5%) with the regional inventory (ECLIPSE annual totals) to tend to increase the bias of the REF run (Fig. 8a and Table 3). Sensitivity is weaker for O$_x$ (0.4 ppb or 1%) suggesting that the main reason for the improvement brought about by the use of the local inventory (REF run) is due to a better representation of the ozone titration process. At the suburban area, the sensitivity is larger ($|\Delta c| = 1.1$ ppb or 3.2%) and of the same order of magnitude as sensitivities to climate model driven meteorology and to the resolution of the meteorological input. The weaker change in suburban O$_x$ ($|\Delta c| = 0.1$ ppb or 0.3%) suggests that this area benefits more than the urban area from the improvement in the titration process. The skill score associated
Changes in daytime averages at both urban and suburban areas are similar to those in the daily averages suggesting that modeled sensitivity stems mainly from daytime titration. Rural ozone is practically unaffected (|Δc| = 0.3 ppb or 1 %). It is noteworthy that the absolute change in modeled ozone concentrations is in the order of 1 ppb or less despite the large differences in ozone precursors’ emissions between the local and the regional inventory.

Changes in fine particles concentrations in summertime, wintertime and in the annual daily average are much stronger than ozone (|Δc| = 4.1 µg m⁻³ or 33 %, 6.6 µg m⁻³ or 33.8 % and 5.5 µg m⁻³ or 31.9 % respectively). PM\textsubscript{2.5} concentrations modeled with the ANN run are significantly higher than those modeled with the REF run (Fig. 8b). Wintertime bias in the ANN run reaches 5.8 µg m⁻³ showing that fine particle emissions from the ECLIPSE inventory are overestimated (see also Fig. 2). The main source of primary wintertime PM\textsubscript{2.5} emissions over the IdF region as well as in Paris in the ANN run is wood burning (see discussion in Sect. 2.4), which is unrealistic for a city like Paris and stems directly from the use of the population proxy to spatially allocate national totals over finer scale. This is consistent to the fact that the summertime bias in the ANN run is much lower (+1.4 µg m⁻³). In fact, in this case the ANN bias is even smaller than the REF bias (−2.8 µg m⁻³) enhancing our hypothesis that summertime fine particle emissions in the AIRPARIF inventory are underestimated (see also Sect. 2.1). The REF skill score is higher than the ANN score in wintertime and lower in summertime.

We conclude that ozone sensitivity to the annual emission totals is low but strong for fine particles.

### 4.5 Sensitivity to emission post-processing (ANN vs. POST)

Here we use identical annual totals but two different methods for their vertical and temporal allocation to obtain hourly fluxes over the 4 km resolution domain and different matrices for their chemical speciation. The ANN dataset uses the AIRPARIF bottom-up approach whereas the EMEP methodology is applied on the POST dataset. To compile the ANN inventory we had to extract the post-processing coefficients of the
bottom-up inventory and apply them on the ECLIPSE annual totals. This procedure though was not emission source-sector oriented and this inconsistency definitely affects model results. On the other hand the post-treatment of sectoral raw emissions in large-scale applications are typically based on sectoral coefficients that don’t link back to the same quantified emissions either. For example in the regional application used this study (REG) the per SNAP sector ECLIPSE raw emissions are treated with SNAP level EMEP information that stems from the EMEP inventory having different synthesis of sub-SNAP sources from that of ECLIPSE. Therefore when we compare ANN with POST we consider that what we observe is the bias of this inconsistency in regional modeling. The question raised is: what is the benefit of adopting a bottom-up post-processing for regional scale air-quality modeling?

The effect on ozone over the urban area is moderate (|Δc| = 1.9 ppb or 6.4 %) (Fig. 9a and Table 3). Model bias is reduced from +4.5 ppb in POST to +2.6 ppb in ANN. Ozone sensitivity in this case, is twice as high as the sensitivity to climate model driven meteorology and even higher compared to the impact of annual totals. The ANN simulation is able to increase the skill score by 14 % and reduce MNB by 26 %. The low O\textsubscript{x} sensitivity suggests that discrepancies are mainly due to a better representation of ozone titration. Suburban and rural ozone is practically insensitive to the post-processing technique. Even if emission totals are the same between the two configurations ozone concentrations over the urban area are lower in the ANN run than in the POST run because the ANN has more ground-layer NO\textsubscript{x} emissions than POST enhancing ozone titration. This stems from the fact that the annual emission totals are allocated in the model’s vertical layers very differently. Following the AIRPARIF post-processing (ANN) all urban emissions are released in the surface layer because no major point sources can be found within the urban area. On the contrary, the regional scale post-processing (POST) does not resolve the urban from the suburban and rural areas, where industrial zones are located and assigns only 70 % of the total NO\textsubscript{x} emission in Paris in the first model layer.

Another important piece of information of the post-processing of emissions regards their diurnal variation. Although the time scale of a climate forced run largely exceeds
the hourly basis we aim to illustrate how important can the choice of the diurnal patterns be to the final modeled concentrations. Figure 10a shows the average diurnal variation of modeled and observed urban ozone for ANN and POST (for the modeled fields we use the grid cells of the monitoring sites). The two downscaling approaches compared here, apply different diurnal profiles on emissions to provide hourly fluxes. Between 10:00 LT and 15:00 LT, ANN underestimates ozone concentrations due to too much NO emissions enhancing titration and this is maximized in the local peak (15:00 LT) where NO concentrations are overestimated by a factor of 2. The daily maximum concentration shows the highest sensitivity in the emission post-treatment among all the presented cases ($|\Delta c| = 2.2$ ppb). This is consistent with Menut et al. (2003) who also found that the afternoon peak concentrations at a typical summertime episode in Paris is very sensitive to the NO emissions change. In the evening (after 15:00 LT) ANN deviates faster than POST from the observations because the afternoon peak in traffic emissions is more pronounced in the AIRPARIF diurnal profile compared to that used in the ECLIPSE processing which represents an average situation of anthropogenic sources hence a smoother variation. These results indicate that the diurnal variability of modeled ozone over the urban area is very sensitive to the choice of the diurnal profile. But in the climate concept where hourly values are timely too short to take into account the sensitivity is considered moderate as seen in Table 3.

Modeled PM$_{2.5}$ sensitivity is significant for both summer and wintertime ($|\Delta c| = 3.4$ µg m$^{-3}$ or 24.8% and 4.6 µg m$^{-3}$ or 18.3% respectively) (Table 3). POST wintertime bias is almost two times higher than ANN (Fig. 9b). A late afternoon peak is modeled with ANN accounting for the traffic emissions, whereas PM$_{2.5}$ evening levels modeled with the POST run (after 20:00 LT) are related to the residential heating activity.

What we can conclude is that in a climate forced–air quality framework the model response for daily average ozone by 6.2% is rather small considering the significant differences that the two post-processing approaches prescribe for the vertical distribution of emissions and their diurnal variation. Fine particle concentrations are much more sensitive to the applied emission post-processing technique.
4.6 Sensitivity to the emission inventory resolution (POST vs. AVER)

Here, we quantify the effect of the resolution of the emission input from the impact of model resolution. Results show that in the urban areas this sensitivity is the most influential amongst all tests presented in this paper with ozone changes reaching $|\Delta c| = 2.8$ ppb or $8.3\%$ (Fig. 11a). The change in daily average $O_x$ is smaller but comparable ($|\Delta c| = 1.2$ ppb or $2.9\%$) suggesting that ozone titration is not the only model processes that is affected by the increase in the resolution of the emission dataset. The skill score and MNB improve significantly in the POST run (Table 3). The increase in the resolution of the emission input leads to a reduced positive bias from $+7.3$ ppb (AVER) to $+4.5$ ppb (POST). Ozone precursors’ emissions from urban sources are mixed with the lower emissions from the surrounding suburban and rural areas inside the large cells of the coarse mesh-grid. This leads to lower titration rates and therefore, higher ozone levels. AVER overestimates ozone peaks by $0.8$ ppb while POST underestimates them by $−1.2$ ppb. The sensitivity of ozone concentration at the hour of the afternoon peak is linked to NO$_x$ concentration at the same hour, which reaches a local maximum due to the evening rush hour (see also Sect. 4.5). Suburban and rural ozone is less sensitive than urban ($|\Delta c| = 0.7$ ppb), with scores practically unchanged (Table 3).

Fine particles concentrations are also very sensitive to the resolution of the emission input, especially in wintertime ($|\Delta c| = 7.1$ µg m$^{-3}$ or $30\%$), with higher concentrations modeled with the refined emission inventory in POST (Table 3). This is because in the coarser inventory represented here by AVER, emissions in the high emitting areas in the city are smoothed down and diluted when averaged with emissions of the less polluted outer areas.

We conclude that the resolution of the emission input is the most influential factor from all the studied cases, even more than model resolution itself. PM$_{2.5}$ showed higher sensitivity than ozone concentrations. The non-linear nature of ozone chemistry suggests that it is important for the ozone precursor emissions to be concentrated correctly to the high emitting areas such as the urban centres.
4.7 Sensitivity to model horizontal resolution (AVER vs. REG)

Here, we study the sensitivity of ozone and PM$_{2.5}$ to model horizontal resolution. We compare simulations at two different spatial resolutions, the AVER run (averaged over the grid-cells of the coarser grid) and the REG simulation on a grid of 0.5$^\circ$ resolution (Fig. 12). REG models higher ozone concentrations than AVER over the urban area ($|\Delta c| = 1.7$ ppb or 4.7%). As discussed above, NO$_x$ emissions in the REG simulation are lower than in REF due to dilution in the coarser grid cells leading to lower ozone titration rates. Suburban and rural ozone is very little sensitive to model resolution ($|\Delta c| = 0.5$ ppb or 1.4% and 0.2 ppb or 0.5% respectively) because photochemical build-up occurs at larger time and space scales compared to titration and the refinement of the model grid does not provide much new information to the modeling. This confirms the results in Markakis et al. (2014). The effect on modeled PM$_{2.5}$ is very small with concentrations slightly higher over the finer mesh grid.

We may conclude that the benefit of increasing model resolution is insignificant for both ozone and PM$_{2.5}$ especially taking into account the large increase of model resolution from 50 to 4 km.

5 Sources of error in regional climate forced atmospheric composition modeling

In this paper we utilize simulations at two spatial scales: at urban scale over a grid of 4 km resolution using the AIRPARIF bottom-up inventory of anthropogenic emissions (REF) and a regional scale run at 0.5$^\circ$ resolution where emissions stem from the ECLIPSE top-down inventory (REG). Both realizations implement identical climate driven meteorology (at 0.44$^\circ$ resolution) and an 8-layer vertical mesh therefore are susceptible to the same source of error due to climate model driven meteorology, the resolution of the meteorological input and the resolution of the vertical grid. However the remaining biases presented in Table 3 over urban areas e.g. the emissions resolution,
the model horizontal resolution, the annual quantified fluxes and the post-processing method concern mainly the REG run. As regards ozone REG has a positive bias of 9 ppb over the city of Paris while the bias of REF is only +1.8 ppb (Fig. 13a). The question we raise is what are the main sources of uncertainty in regional scale climate driven air-quality simulations and how these could be eliminated or at least reduced.

With this study we are able to identify the source of the excess of $|\Delta c| = 7.2$ ppb of ozone modeled with the REG run compared to REF (Table 4); 26.4 % ($|\Delta c| = 1.9$ ppb) is related to the post-processing of the annual emissions totals based on the EMEP factors, 11.1 % ($|\Delta c| = 0.8$ ppb) to the annual emission totals in the ECLIPSE inventory, 23.6 % ($|\Delta c| = 1.7$ ppb) to coarse model resolution and 38.9 % ($|\Delta c| = 2.8$ ppb) to the coarse resolution of the ECLIPSE emission inventory.

Considering the discrepancies in the inventorying methodologies used to compile the ECLIPSE and the AIRPARIF datasets (top-down vs. bottom-up), it is very interesting that the least influential factor for ozone would be the annual emissions totals. It seems that the regional simulation would not benefit much from the integration of the local annual totals alone but a more important gain would stem from the application of the AIRPARIF post-processing methodology. The added value from both these factors would reduce the positive bias of REG by 2.7 ppb. Even largest improvement comes through the better spatial representation of ozone precursors emissions in the local emission inventory ($|\Delta c| = 2.8$ ppb) leading to more faithful titration process; $O_x$ levels are very close in REF and REG (Fig. 13a). It could therefore argued that without increasing model resolution of which the gain would reach only 1.7 ppb, the REG simulation would benefit significantly by simply integrating the aforementioned local scale information.

The difference in modeled ozone between REF and REG is much smaller over the suburban area ($|\Delta c| = 2.4$ ppb) and the most influential factor to this difference is the annual emission totals covering 45.8 % of this difference. Finally as regards ozone one important result of this study is that in the climate–air quality framework modeled concentrations from a coarse resolution run well agree with the much more intensive,
in terms of computational time, fine resolution run and the bias is considered of small magnitude (Fig. 13a). This is because the formation of rural ozone is a slower process than in urban areas and comparable to the characteristic transport time of precursor’s pollutants to the coarse grid cell.

Focusing on the wintertime PM$_{2.5}$ concentrations where the largest annual levels are observed, are better simulated with the REF run with a bias of $-0.8 \mu$gm$^{-3}$ and a high skill score of 0.78 compared to a strong positive bias of $+3.6 \mu$gm$^{-3}$ and a skill score of 0.68 with the REG run (Fig. 13b). We should remind here that both runs suffer from a strong wet bias reducing significantly PM$_{2.5}$ concentrations (see also Sect. 3.1). Contrary to ozone, where information from the local scale improves in all cases model performance, the resolution of the emission inventory seems to deteriorate the modeling of PM$_{2.5}$ with increase in the bias by 7.1 µgm$^{-3}$. All other features have a positive effect; model resolution reduces the bias by 0.4 µgm$^{-3}$, annual emission totals by 6.6 µgm$^{-3}$ and post-processing of the annual totals by 4.5 µgm$^{-3}$. This essentially means that the regional realization cannot selectively incorporate any combination of local-scale features in order to improve performance as in the case of ozone. But the results indicate that by simply integrating bottom-up post-processing technique would result in an overall bias of the regional application of $-0.9 \mu$gm$^{-3}$.

6 Conclusions

In the present paper we assess the sensitivity of ozone and fine particles concentrations with respect to emission and meteorological input with a 10 yr long climate forced atmospheric composition simulation at fine resolution over the city of Paris.

As a general observation our study shows that overall ozone response is considered low to moderate while PM$_{2.5}$ concentrations were generally very sensitive for the presented cases. The largest sensitivity in average daily ozone concentrations was observed in the urban areas primarily due to the resolution of the emission inventory ($|\Delta c| = 2.8$ ppb or 8.3 %) and secondly to the post-processing methodology applied on...
the annual emission totals ($|\Delta c| = 1.9$ ppb or 6.2\%). These sensitivities are attributed to changes in the titration process. When post-processing coefficients were derived from the bottom-up AIRPARIF inventory instead of EMEP, too much ozone titration takes place at the hour of the ozone peak and the sensitivity of daily maximum reached its highest value among all the studied cases ($|\Delta c| = 2.2$ ppb or 5.8\%). It is interesting that despite the fact that ozone precursor’s emissions are very different between the bottom-up and the top-down inventories, ozone sensitivity to the annual totals was shown to be very small ($|\Delta c| = 0.8$ ppb or 2.5\%). Also modeled ozone is fairly insensitive to the use of climate model or reanalysis meteorology as well as to the resolution of the meteorological input dataset. Finally all cases of suburban and rural ozone both for average and maximum concentrations showed a sensitivity of less than 5\%.

Regarding PM$_{2.5}$ concentrations amongst all the presented factors, the emissions related were those are shown to be the most influential. The corresponding sensitivity to the use of annual emission totals from a top-down and a bottom-up inventory reached 33\% in summer, 33.8\% in winter and 31.9\% for the daily average annual concentrations. This is connected to the downscaling methodology applied on the regional-scale totals in the ECLIPSE inventory. Using population as proxy for their spatial allocation leads to overestimation of particle emissions from wood-burning over the Paris area. Large sensitivity was also shown due to the resolution of the emission inventory (20.3\% in the summer, 30\% in the winter and 24.2\% in annual basis) because coarser inventory smoothen the sharp emission gradients over the urban area leading to less primary emissions. Fine particle concentrations were also sensitive to the applied emission post-processing technique (22.1\% in summer and 16.7\% in winter). Only wintertime PM$_{2.5}$ concentrations were significantly affected by the meteorological related sensitivities; by 17.6\% due to the use of meteorology from reanalysis instead of climate (mainly because the prescribed changes in modeled precipitation) and by 6.8\% due to refinement of the meteorological grid.

Both ozone and PM$_{2.5}$ are little sensitive to model’s vertical resolution (changes of less than 2.2\%). Nevertheless we provide evidence that this low sensitivity may be
the result of counteracting factors such as ozone titration, dry deposition and vertical mixing, too much dependent on local topography to be able to generalize for other regions. Also we note the weak sensitivity to modeled concentrations to the increase of model horizontal resolution from 50 to 4 km which in all case never exceeded 4.7 %.

To fill the gap between regional and city-scale assessments we have to combine in a single application the advantages of regional and local scale applications; the low resolution (but high spatial coverage) from one hand and the good representation of emissions (but limited area of coverage) on the other. The results of this study move towards that goal and can be used in order to identify the main sources of error in regional scale climate forced air-quality modeling over the urban area. These biases could be taken into account used in policy relevant assessments.

The difference in modeled daily average ozone between the local and regional application over the urban areas ($|Δc| = 7.2$ ppb) is attributed to several sources of error: 38.9 % is related to the resolution of the emission inventory, 26.4 % stems from the post-processing of national annual emission totals, 23.6 % is due to model resolution (4 km or 0.5°) and 11.1 % is associated to the annual emission totals used as starting point for the compilation of the anthropogenic emission dataset. Although the greatest benefit in the regional-scale modeling seems to come through the increase in the resolution of the emission inventory, simpler actions may be also meaningful, such as the integration of the locally developed annual totals and downscaling coefficients derived from the existing bottom-up modeling systems reducing the bias of the regional application by 37.5 %.

As regards PM$_{2.5}$ modeling our study shows that the regional realization cannot selectively incorporate any combination of local-scale features in order to improve performance as in the case of ozone. The simulation at regional scale (REG) predicts an excess of 3.6 $\mu$g m$^{-3}$ during wintertime compared to the fine scale simulation (REF) showing a bias of $-0.8$ $\mu$g m$^{-3}$ and this is attributed to the allocation of wood-burning emissions over the Paris area. Therefore, the most influential factor for PM$_{2.5}$ modeling is the resolution of the emission input ($\text{REG} - \text{REF} = +7.1$ $\mu$g m$^{-3}$).
But the implementation of the refined emission resolution of the local inventory alone would not benefit the regional simulation (which would increase the overall bias to $10.7 \, \mu g \, m^{-3}$), neither the implementation of the annual emissions of the bottom-up inventory alone ($\text{REG} - \text{REF} = -6.6 \, \mu g \, m^{-3}$) which would generate an overall negative bias of $3 \, \mu g \, m^{-3}$. A simpler action would be to integrate the post-processing bottom-up technique ($\text{REG} - \text{REF} = -4.5 \, \mu g \, m^{-3}$) giving an overall bias in REG of $-0.9 \, \mu g \, m^{-3}$.

Acknowledgements. This work is supported by the ERA-ENVHEALTH project (grand agreement n° 219337), co-funded by the European Commission under the 7th Framework Programme. The authors would also like to acknowledge Laurent Menut (LMD/Institute Pierre-Simon Laplace) and Augustin Colette (INERIS) for their insightful comments that helped improve this manuscript.

References


EPA: Guidance on the Use of Models and Other Analyses for Demonstrating Attainment of Air Quality Goals for Ozone, PM$_{2.5}$ and Regional Haze, EPA-454/B-07-002, April, 2007.


Table 1. Parameterization of the different sets of simulations presented in the paper. Changes with respect to the REF case are marked in italic. Changes with respect to a simulation other than REF are marked in bold.

<table>
<thead>
<tr>
<th></th>
<th>Annual emission totals&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Air-quality model resolution</th>
<th>Emission inventory resolution</th>
<th>Emission post-processing&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Climate/reanalysis meteorology and resolution</th>
<th>Number of layers in air-quality model</th>
</tr>
</thead>
<tbody>
<tr>
<td>REF AIRPARIF</td>
<td>4 km</td>
<td>4 km</td>
<td>Bottom-up</td>
<td>RCP-4.5 (0.44°)</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>REG&lt;sup&gt;c&lt;/sup&gt; ECLIPSE 0.5°</td>
<td>0.5°</td>
<td>0.5°</td>
<td>Top-down</td>
<td>RCP-4.5 (0.44°)</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Sensitivity simulation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERA05 AIRPARIF</td>
<td>4 km</td>
<td>4 km</td>
<td>Bottom-up</td>
<td>ERA (0.44°)</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>ERA01&lt;sup&gt;d&lt;/sup&gt; AIRPARIF</td>
<td>4 km</td>
<td>4 km</td>
<td>Bottom-up</td>
<td>ERA (0.11°)</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>VERT AIRPARIF</td>
<td>4 km</td>
<td>4 km</td>
<td>Bottom-up</td>
<td>RCP-4.5 (0.44°)</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>ANN ECLIPSE</td>
<td>4 km</td>
<td>4 km</td>
<td>Bottom-up</td>
<td>RCP-4.5 (0.44°)</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>POST&lt;sup&gt;e&lt;/sup&gt; ECLIPSE</td>
<td>4 km</td>
<td>4 km</td>
<td>Top-down</td>
<td>RCP-4.5 (0.44°)</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>AVER&lt;sup&gt;f&lt;/sup&gt; ECLIPSE</td>
<td>4 km</td>
<td>0.5°</td>
<td>Top-down</td>
<td>RCP-4.5 (0.44°)</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> The resolution of the emission inventory of AIRPARIF is 1 km (aggregated to 4 km for the purpose the local simulations) and the ECLIPSE inventory 50 km.

<sup>b</sup> Temporal, vertical allocation and chemical speciation.

<sup>c</sup> This simulation is used as boundary conditions for all local scale simulations.

<sup>d</sup> The ERA01 simulation is compared with the ERA05 not with the REF.

<sup>e</sup> The POST simulation is compared with the ANN not with the REF.

<sup>f</sup> This is not a standalone simulation. Concentrations modeled at 4 km resolution with the POST run are averaged spatially to match the cells of REG (0.5° resolution simulation). AVER results are compared to REG to quantify the effect of model resolution and with POST to quantify the effect of the resolution of the emission inventory.
Table 2. Observed and modeled daily average meteorological variables over the Île-de-France region. MET_CLIM dataset stems from a climate model and MET_ERA05, MET_ERA01 from reanalysis data at 0.5° and 0.1° resolution respectively. Absolute model bias is given in parenthesis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>MET_CLIM</th>
<th>MET_ERA05</th>
<th>MET_ERA01</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Summer (JJA)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2 (°C)</td>
<td>19.19</td>
<td>19.14 (-0.05)</td>
<td>18.28 (-0.91)</td>
<td>18.19 (-1.0)</td>
</tr>
<tr>
<td>WS10 (m s⁻¹)</td>
<td>2.9</td>
<td>4.0 (+1.1)</td>
<td>3.8 (+0.9)</td>
<td>3.8 (+0.9)</td>
</tr>
<tr>
<td>RH (%)</td>
<td>69.1</td>
<td>68.1 (-1.0)</td>
<td>68.3 (-0.8)</td>
<td>67.3 (-1.8)</td>
</tr>
<tr>
<td>PRECIP (mm day⁻¹)</td>
<td>0.076</td>
<td>0.108 (+0.032)</td>
<td>0.097 (+0.021)</td>
<td>0.098 (+0.022)</td>
</tr>
<tr>
<td><strong>Winter (DJF)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2 (°C)</td>
<td>4.3</td>
<td>4.0 (-0.3)</td>
<td>6.0 (+1.7)</td>
<td>5.8 (+1.3)</td>
</tr>
<tr>
<td>WS10 (m s⁻¹)</td>
<td>3.6</td>
<td>6.2 (+2.6)</td>
<td>5.7 (+2.1)</td>
<td>5.5 (+1.9)</td>
</tr>
<tr>
<td>RH (%)</td>
<td>85.0</td>
<td>80.3 (-4.7)</td>
<td>79.7 (-5.3)</td>
<td>79.5 (-5.5)</td>
</tr>
<tr>
<td>PRECIP (mm day⁻¹)</td>
<td>0.069</td>
<td>0.112 (+0.043)</td>
<td>0.089 (+0.02)</td>
<td>0.087 (+0.018)</td>
</tr>
</tbody>
</table>
Table 3. Absolute difference (and percentage in parenthesis) between daily averaged ozone (ppb) and PM$_{2.5}$ (µgm$^{-3}$) from two climate forced air-quality runs. The most influential factor for each sensitivity test is marked in bold.

<table>
<thead>
<tr>
<th></th>
<th>Ozone</th>
<th>PM$_{2.5}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Urban</td>
<td>Suburban</td>
</tr>
<tr>
<td>Climate meteo (REF vs. ERA05)</td>
<td>1.0 (3.4 %)</td>
<td>1.1 (3.2 %)</td>
</tr>
<tr>
<td>Meteo. resolution (ERA05 vs. ERA01)</td>
<td>0.2 (0.6 %)</td>
<td><strong>1.4 (4.3 %)</strong></td>
</tr>
<tr>
<td>Vertical resolution (REF vs. VERT)</td>
<td>0.3 (1.2 %)</td>
<td>&lt; 0.1 (0.2 %)</td>
</tr>
<tr>
<td>Annual emis. totals (REF vs. ANN)</td>
<td>0.8 (2.5 %)</td>
<td>1.1 (3.2 %)</td>
</tr>
<tr>
<td>Emission post-proc. (ANN vs. POST)</td>
<td>1.9 (6.4 %)</td>
<td>0.1 (0.4 %)</td>
</tr>
<tr>
<td>Emission resolution (POST vs. AVER)</td>
<td><strong>2.8 (8.3 %)</strong></td>
<td>0.7 (1.9 %)</td>
</tr>
<tr>
<td>Model resolution (AVER vs. REG)</td>
<td>1.7 (4.7 %)</td>
<td>0.5 (1.4 %)</td>
</tr>
<tr>
<td></td>
<td>Summer</td>
<td>Winter</td>
</tr>
<tr>
<td>Climate meteo (REF vs. ERA05)</td>
<td>&lt; 0.1 (0.05 %)</td>
<td>3.1 (17.6 %)</td>
</tr>
<tr>
<td>Meteo. resolution (ERA05 vs. ERA01)</td>
<td>0.3 (3.4 %)</td>
<td>1.3 (6.8 %)</td>
</tr>
<tr>
<td>Vertical resolution (REF vs. VERT)</td>
<td>&lt; 0.1 (0.3 %)</td>
<td>0.5 (2.2 %)</td>
</tr>
<tr>
<td>Annual emis. totals (REF vs. ANN)</td>
<td><strong>4.1 (33.0 %)</strong></td>
<td>6.6 (33.8 %)</td>
</tr>
<tr>
<td>Emission post-proc. (ANN vs. POST)</td>
<td>3.4 (24.8 %)</td>
<td>4.5 (18.3 %)</td>
</tr>
<tr>
<td>Emission resolution (POST vs. AVER)</td>
<td>2.1 (20.3 %)</td>
<td><strong>7.1 (30.0 %)</strong></td>
</tr>
<tr>
<td>Model resolution (AVER vs. REG)</td>
<td>0.4 (4.1 %)</td>
<td>0.4 (1.9 %)</td>
</tr>
</tbody>
</table>
Table 4. Top row presents the coarse resolution application (REG) model bias of the April–August average urban ozone and wintertime urban PM$_{2.5}$. Subsequently, marked with italics the signals – measured as the absolute concentration change from REG – of several refinements such as increase of resolution (model or emissions) and adaptation of annual quantified fluxes and post-processing of a bottom-up inventory. The individual signals sum up to the absolute bias found under the fine resolution simulation (REF).

<table>
<thead>
<tr>
<th>Ozone</th>
<th>Ozone (ppb)</th>
<th>PM$_{2.5}$ (µg m$^{-3}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>REG (50 km)</td>
<td>+9.0</td>
<td>+3.6</td>
</tr>
<tr>
<td>Model resolution</td>
<td>−1.7</td>
<td>−0.4</td>
</tr>
<tr>
<td>Emissions resolution</td>
<td>−2.8</td>
<td>+7.1</td>
</tr>
<tr>
<td>Annual emission totals</td>
<td>−0.8</td>
<td>−6.6</td>
</tr>
<tr>
<td>Emissions post-processing</td>
<td>−1.9</td>
<td>−4.5</td>
</tr>
<tr>
<td>REF (4 km)</td>
<td>+1.8</td>
<td>−0.8</td>
</tr>
</tbody>
</table>
Figure 1. Overview of the coarse (D1 having 50 km resolution) and local scale (D2, illustrated by the red rectangle having 4 km resolution) simulation domains. In D2 the city of Paris is located in the area enclosed by the purple line. Circles correspond to sites of the local air-quality monitoring network (AIRPARIF) with red for urban, blue for suburban and black for rural.
Figure 2. Domain-wide annual emissions of NO\textsubscript{x}, NMVOC (left axis) and PM\textsubscript{2.5} (right axis) from the local (bottom-up) and the regional (top down) inventory (summed across the vertical column).
Figure 3. (a) Scatter plots and scores of daily average ozone concentrations at urban, suburban and rural stations from the REF simulation. Odd oxygen (O$_x$) and daily maximum values at urban locations are also shown. (b) Daily average PM$_{2.5}$ concentrations in wintertime (DJF), summertime (JJA) and on annual basis over urban stations.
Figure 4. Scatter plots and scores for the sensitivity test on climate model driven meteorology for ozone and PM$_{2.5}$. 
Figure 5. Scatter plots and scores for the sensitivity test on the resolution of meteorology for ozone and PM$_{2.5}$. 
Figure 6. Scatter plots and scores for the sensitivity test on the model vertical resolution for ozone and PM$_{2.5}$. 
Figure 7. Difference in average daily simulated NO$_2$ (a), ozone (b) and PM$_{2.5}$ (c) concentrations between VERT (12 vertical layers) and REF (8 vertical layers) at urban areas per range of $K_z$ (bins of 1 m$^2$ s$^{-1}$). Positive differences indicate that the refined vertical mesh leads to increased pollutant concentration and vice versa. The occurrence of sensitivity values within each $K_z$ range is also provided.
Figure 8. Scatter plots and scores for the sensitivity test on the annual emission totals for ozone and PM$_{2.5}$. 
Figure 9. Scatter plots and scores for the sensitivity of ozone and PM$_{2.5}$ on the post-processing (temporal analysis and chemical speciation) technique applied on the annual emission totals.
Figure 10. Average diurnal variation of ozone from April to August (a) and wintertime PM$_{2.5}$ (b) concentrations in the urban area.
Figure 11. Scatter plots and scores for the sensitivity test on the resolution of the emission inventory for ozone and PM$_{2.5}$. 
Figure 12. Scatter plots for the sensitivity test on model resolution for ozone and PM$_{2.5}$. 
Figure 13. (a) Scatter plots of daily average ozone concentrations at urban, suburban and rural stations from the REF and REG simulations. The odd oxygen (O$_x$) and daily maximum at urban locations is also shown. (b) Daily average PM$_{2.5}$ concentrations in wintertime (DJF), summertime (JJA) and on an annual basis over urban stations.