Reviewer 1

My overall rating for this paper is "minor revisions"; there are a few places where the text should be clarified. They are however important clarifications – the authors' title implies that a 4km resolution CTM and 0.11 degree/12.2km resolution driving meteorological model (or even a 4km resolution meteorological model) captures urban meteorology and chemistry. This may not be the case, from some recent work they reference. I wouldn't say that this invalidates any of the authors' results, but their conclusions and discussion need to acknowledge that some important meteorological circulation effects may only be captured at finer scales than employed in their study. There are three aspects to this concern, which I outline below, all of which can be addressed by adding appropriate caveats to their text. There are also a number of smaller concerns and spelling/grammar issues which follow, which would clarify and improve the readability of the paper.

Larger issues:

(1) I think it's important that the authors acknowledge that at least some of these results may be specific to the region they have studied (Ile de France). For example, one of their findings was that model vertical and horizontal resolution (both in the driving meteorology and the resolution of the CTM that made use of that meteorology) had a relatively minor effect on O3. Meanwhile, I'm aware of studies in coastal environments (where the city is located along a large lake or ocean coastline) where the resolution can have a profound effect on the ability of the meteorological model to capture the land-sea (or land-lake) breeze, and that in turn has a significant impact on ozone predictions. Similarly, though the authors note themselves (second sentence of section 2.1) that the region under study is far from the coast and has limited topography, these factors if present could significantly increase the relative impact of horizontal and vertical resolution. In that respect, some caveats should be placed within the text acknowledging the potential limitations of their study towards similar resolution studies in other domains (e.g. in the Abstract, page 4768, line 17, change "are of little effect." To "are of little effect for the regional/urban domain and models studied here."

Response: We acknowledge the fact that some of our results related to model resolution are representative to the specific characteristics of our study area. The abstract was revised according to the reviewer's suggestion: "*The air quality model horizontal and vertical resolution have little effect on model predictions for the specific study domain*." We also insert the following text on page. 4784, line 27: "*Never the less this result could reflect the local area's characteristics (flat terrain, away from the coast) confirming previous studies (Menut et al., 2005; Valari and Menut, 2008). In regions with more complex topography or close to the coast the resolution of the meteorological input could have a profound effect on the simulated meteorological conditions (Leroyer et al., 2014).*"

(2) The authors also need to discuss the accuracy of their meteorological model in the context of resolution – is the lack of a resolution impact of the meteorological model on chemical predictions reflective of resolution not being important (the authors' conclusion, I think) or reflective of a meteorological model whose performance has not improved in going to higher resolutions? They later (page 4776, lines 18-19) reference previous work that shows that the flat homogeneous topography of the region results in little benefit in going to higher resolution. But what was the highest resolution attempted in the referenced work by Menut et al, 2005, and Valari and Menut, 2008)? The highest resolution of the input meteorology used in the authors' current paper is 0.11 degrees or about 12.2km – this is inappropriately coarse to resolve much of the urban circulation (cf. also line 4782, line 2 regarding the CTM itself – I agree, this is very likely, so caveats in the abstract and conclusions need to be added to that effect). In the Leoyer et al paper quoted by the authors, substantial changes in vertical and horizontal transport in an urban environment for a meteorological model occurred mostly in the transition from 2.5km to 1km to even higher horizontal resolutions (e.g., 250m). The resolution impact in the IdF region could thus be substantial, but just not yet seen on the still relatively coarse resolutions used by the authors in this set of tests. 12.2 km meteorology inputted into a 4km CTM won't see much if anything of the urban circulation, and a 4km resolution met model might not see much of that circulation either. The authors' conclusion should be modified to include caveats to that effect, with the message going from, e.g., "horizontal and vertical resolution are not important" to "input horizontal and vertical resolution and CTM resolution were not important for the regional/urban domains and model resolutions studied here. The relative importance of resolution may however increase with further reductions in grid size in both the driving meteorological model (c.f. Leoyer et al, 2014), and the CTM making use of that meteorology."

Response: We agree. The increase of the meteorological model's resolution from 0.5° to 0.1° might not be enough for the chemical model to produce noticeable concentration responses. In Menut et al., 2005 the increase is from 0.5° to 3km and in Valari and Menut, 2008 from 0.5° to 5km. Based on the Leroyer et al., 2014 results this might not be sufficient. Never the less according to Leroyer et al. 2014: "*The necessity of using subkilometer meteorological numerical systems remains, however, questionable. The improvement in forecasting obtained by going to subkilometer modeling systems, with the use of a detailed surface description and of a more physically based (less parameterized) representation of atmospheric processes, still has to be rigorously demonstrated*". Despite that it is useful to include such discussion in the text e.g. page. 4784, line 27: "*We note here that the refinement in the resolution of the meteorological model from 0.5° to 0.1° may not be sufficient for the CTM to simulate noticeable concentration responses. For example Leroyer et al. (2014) (see also references therein) observed that substantial changes in vertical and horizontal transport in an urban environment occurred mostly in the transition from resolutions of 2.5km to 1km and even higher (250m).*" We also revise a sentence in the conclusions section (page 4795, line 4): "*We also note the weak sensitivity of modeled concentrations to the increase in the CTM's and the meteorological model's horizontal resolution at least for the area and the range of resolutions studied here.*"

(3) Page 4785, section 4.3. I'm assuming here that this sensitivity test looks at the vertical resolution of the CTM, but there is no corresponding increase in vertical resolution of the driving meteorological data (this should be clarified - mentioned in the first sentence)? This needs to be explicitly made clear in the text. This is another case where the impact of "vertical resolution" needs to be split conceptually in the text between "vertical resolution of the driving meteorology" and "vertical resolution of the CTM". My point here is that a meteorological model with a higher vertical resolution will almost certainly generate different vertical diffusion coefficients, than one with a lower vertical resolution. Increasing the resolution of the CTM may not capture this effect. A caveat to that effect should be in the conclusions and the abstract with regards to the vertical resolution issue. The authors should avoid the use of phrasing like "of the model's vertical resolution" and instead be using "of the CTM's vertical resolution" or "of the driving meteorological model's vertical resolution".

Response: WRF model runs on a 31 vertical layer grid, which we consider highly resolved, this is now added in the modeling setup description. Studying the effect of the vertical resolution of the meteorological model could have been a separate sensitivity test which we didn't chose to study. In any case the WRF meteorology is interpolated to the CTM's (i.e. CHIMERE) vertical grid therefore, technically, increasing the number of vertical layers in CHIMERE from 8 to 12 will result in a refinement of the meteorological input used for the chemical simulation as well. This was added in section 2.4 (description of the sensitivity simulations). With this clarification the phrase added in the abstract ("*The air quality model horizontal and vertical resolution have little effect on model predictions for the specific study domain*.") inevitable covers this -indirect- refinement of the vertical resolution of meteorology. Finally, we changed the ambiguous term 'model vertical resolution' into 'CTM's vertical resolution' throughout the manuscript as suggested by the reviewer.

Minor issues:

Abstract and conclusions: needs to be clarified with regards to the resolution of the input meteorology versus the resolution of the CTM used for the modelling.

Response: These additions were implemented, please refer to the responses in the "major issues" section.

Page 4768, line 14: is that supposed to be "meteorological model input resolution"?

Response: No this is the CTM's horizontal resolution. We have revised the issue in the manuscript.

Page 4769, line 20: another reference here: Kelly et al, 2012, 5367-5390, 2012.

Response: Added.

Page 4770, line 19 to line 24: the sentence "In Markakis (2014)...NOx-limited conditions." Could use a little clarification: presumably each didn't work for the other condition, as well?

Response: We rephrased our sentence to make our point clearer. The sentence now reads: "*In Markakis et al. (2014) we showed that ozone formation occurs under a VOC-limited chemical regime in the 10-year simulations that used the bottom-up emission inventory. This result is consistent with previous studies over the Paris area (Beekman and Derognat, 2003; Beekman and Vautard, 2010; Deguillaume et al., 2008). On the contrary, when the regional top-down inventory was used instead, ozone formation occurred under a* NO_x*-limited chemical regime*."

Page 4771, first line: using RCP6, Kelly et al did this – a laborious process of linking the RCP recommendations with specific industries.

Response: This is indeed an oversight from our part. We revised the manuscript: "*Long-term projections are constrained by the evolution of large scale energy supply and demand and the link between global and regional scale projections is a laborious task (Kelly et al. (2012))*."

Page 4772, line 13: The authors should explain why this particular RCP was used in their work, rather than the other RCP scenarios available.

Response: This was an inevitable decision related to the availability of simulations on the larger scale that were used to provide the boundaries to our simulations.

Page 4774, line 1: might be worth mentioning here that the database in question does not explicitly consider point sources.

Response: We revise "Present-time emissions (as areas sources) are compiled…"

Page 4775, line 25: not clear why the same metrics were not being used for both O3 and PM2.5. Explain.

Response: We use MFB and MFE which are considered more suitable for fine particles evaluation based on EPA guidelines. In Page 4775, lines 6-9 we already include the EPA, 2007 reference.

Page 4788, lines 24-25: This was rather a surprise to me, though I may be used to more detailed emissions inventories in North America. When you say "no major point sources can be found within the urban area" does this mean that "no major point sources exist within the urban area of Paris" or "no major point sources are explicitly included within this inventory within the urban area of Paris"? In contrast, North American emissions inventories in Canada and the USA include tens of thousands of point sources (to the extent that one has to choose criteria for deciding which ones will be selected for plume rise calculations and which have minor enough emissions to be treated as area sources). So a line or two explaining whether the issue here is a lack of data in this inventory for this region, a lack of data in any inventory for this region, or if there really are no major point sources in this large city.

Response: The local emission inventory includes tenths of point source (Page 4774, line 13-15) but the major of these sources are in industrial areas outside the city. We revise: "*Following the AIRPARIF post-processing (ANN) all urban emissions are released in the surface layer because according to the local point source emission database no major industrial units are found within the urban area.*"

Page 4789, lines 10-13 and section 4.5 in general. You might also be interested in having a look at Makar et al, Geoscientific Model Development, 7, 1001-1002, 2014, since there are several points of overlap between the authors' paper and that one: the reference looks at several stages of emissions improvement and how two different off-line CTMs responded to those changes. Both temporal and spatial changes in emissions were evaluated and the impacts on O3 and PM2.5 predictions evaluated.

Response: The paper is very interesting and we have added a discussion in Page 4778, line 13: "*Makar et al. (2014) investigated the response of modeled concentrations to the refinement of the spatial and temporal allocation of input emissions and found that the model was as sensitive to these improvements as to the vertical mixing parameterization. Also they conclude that the temporal distribution of emissions in particular, could be very important in stable urban atmospheres and that this sensitivity is reduced with increased mixing conditions.*" We also add at the end of section 4.5: "*We note here, that recent work has pointed out that the sensitivity of modeled concentrations the spatiotemporal resolution of the emission inventory is model-dependent (Makar et al., 2014).*"

Page 4791, line 15-16; see the above discussion; I think that this last sentence currently ending "especially taking into account the large increase of model resolution from 50 to 4km" should be "for the range of meteorological and CTM horizontal resolutions attempted here. A stronger impact of resolution may occur at even higher resolutions (c.f. Leroyer et al, 2014)."

Response: This sentence is under the "effect of CTM's resolution" section therefore we revise the statement according to that feature alone e.g.: "*We may conclude that the benefit of increasing the CTM's resolution is insignificant for both ozone and* $PM_{2.5}$ *especially taking into account the large refinement attempted here (*0.5^o *to 4km).*" The reviewer's suggestion as to the possible effect of refining the meteorological model resolution is already added in the text in accordance to the comment number 2 of the "major issues" section.

Page 4792, lines 25 through 27 seems to be saying "getting annual emissions totals right has a big impact on O3 results" while lines 12 through 14 seem to be saying "getting the annual emissions totals right has a very minimal impact on O3 results". This needs to be clarified.

Response: The statement in lines 12-14 regards only urban ozone. This is clarified: "*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 to the urban ozone responses is the annual emissions totals.*"

Page 4793, lines 10 through 12: the authors may wish to consider and discuss the potential for compensating errors in this regard. That is, (1) the PM2.5 levels in the urban regions are likely mostly controlled by primary emissions; (2) increasing the emissions inventory resolution will defacto concentrate the PM2.5 emissions into a smaller spatial extent of the urban area (the reverse side of the artificial dilution issue that the authors have already discussed); (3) if the emissions totals are themselves biased high, then the resulting error will only become apparent at higher resolution. That is, the conclusion should not necessarily be "the emissions resolution makes the PM2.5 worse", but "the emissions resolution may be showing us that the emissions totals are too high, and this only becomes apparent at high resolutions".

Response: "*the emissions resolution makes the PM2.5 worse*": this is by no means the message we intent to deliver. Table 6 only shows how the REG application would respond if selectively incorporated features of the local application. The message is: "*if the REG application uses the coarse inventory totals along with higher resolution modeling and refined spatial allocation of emissions this would result in a high overestimation of concentrations*" and that "*the REG application has to adopt all emission-related features of local scale to improve model scores*". We have added a discussion on possible error compensation in the modeling exercise presented in the paper as suggested by both reviewers (see also the 2nd reviewer's 1st comment).

Minor spelling/grammar mistakes to be corrected:

Response: Spelling mistakes were corrected in the text.

Page 4768, first sentence of abstract starts with a preposition, better to use "Previous research helped to....spatial scale effect, but our knowledge is limited....

Page 4768, line 21: "(same improvement" should be "(the same improvement"

Page 4768, line 23,24: "bias on" should be "bias of", and "associated to" should be "associated with"

Page 4769, line 2: "at urban scale" should be "at the urban scale", ditto, line 24.

Page 4770, line 2: "By principle" should be "In principle".

Page 4770, line 27: "to higher" should be "in higher". Reductions of what? O3?

Page 4771, lines 3 to 5: sentence is unclear.

Page 4771, line 7: “demands and that” should be “demand. This“

Page 4771, last line: “f)” should be “(f)”, ditto for “g)” should be “(g)” on the next page.

Page 4773, line 5: “sulfates, nitrates” should be “sulfate, nitrate”.

Page 4773, line 14: “vertical” misspelled.

Page 4773, line 19: “gasses” should be “gases”, I think. Might be British versus American spelling conventions, here.

Page 4774, line 5: “are available” should be “are also available”

Page 4776, line 9: “speed was the” should be “speed were the”

Page 4777, line 9-10: “due to the surface emissions, ozone concentrations in the afternoon peak hour had the second largest sensitivity after meteorology.” Should be “the sensitivity of ozone concentrations in the afternoon peak hour was the second largest after the sensitivity associated with meteorology.”

Page 4778, line 6: “scale” should be “scales”

Page 4784, line 3: “to modeled precipitation” should be “to the accuracy of modelled precipitation.”

Page 4784, line 20: “of the meteorological grid” should be “of the input meteorological grid” (I think).

Page 4787, line 5: “is in the “ should be “is on the”.

Page 4790, line 6: “model processes” should be “model process”. Line 18: “Fine particles” should be “Fine particle”

Page 4791, line 8: “is very little sensitive” should be “has relatively low sensitivity” Page 4791, line 24: “same source” should be “same sources”

Page 4792, line 5: end the sentence with a question mark and put quotations around “what are the main...or at least reduced?”

Reviewer 2

In this study, Markakis et al., conduct a series of sensitivity calculations using WRF and CHIMERE to downscale global inputs (IPSL-CM5A-MR an LMDz-INCA) to 0.44 and 0.11 degrees (for the met) and 4 km (for air quality). The sensitivity calculations conducted are used to assess the response of the modeling system to various different ways of processing inputs (e.g., resolution, which emissions, climate/reanalysis inputs). For those involved in such calculations the manuscript provides useful information.

General Comments:

(1) One issue I found throughout the manuscript is the struggle for what is the appropriate way to evaluate models when conducting such an exercise. There is always a temptation to suggest that the close the model results are to the observations, the better, which is typically true. However, then one has to decide upon a set of metrics, of which there are many. It is not always the case that finer resolution is better (for numerical and other reasons. . . e.g., potential spatial misalignment in inputs). Further, there is the issue of compensatory errors. Thus, one is left with the question of what should be the standard for comparison, and what is the "best" result. This manuscript does well at showing sensitivities, but also discusses bias, with the implicit assumption that a smaller bias means that the simulation is better. However, it could just be that it does a better job at having errors compensate for each other. The manuscript should better deal with this issue.

Response:
We added a discussion on possible error compensation on the conclusion section. The text added there reads: "*We note here that PM2.5 levels in the urban regions are likely mostly controlled by primary emissions; increasing the emissions inventory resolution will concentrate the PM2.5 emissions into a smaller spatial extent of the urban area (the reverse side of the artificial dilution issue taking place at coarse resolution); if the emissions totals are themselves biased high, then the resulting error will only become apparent at higher resolution. Therefore, the emissions resolution may be showing that the emissions totals are too high, and this only becomes apparent at high resolutions.*"

(2) They should also consider doing a dynamic analysis of the model responses (e.g., the work by Dennis et al., at the US EPA and as part of AQMEII).

Response: We quote the AQMEII Dennis et al. paper: "*This exercise requires historical case studies where known emission changes or meteorological changes occurred that could be confidently estimated. Dynamic evaluation also requires that these changes have a discernable impact on air quality*". To our best of knowledge we are not familiar with any such studies in IdF that would fit the prerequisites of the dynamical evaluation process.

(3) The authors would also serve the community by digging in to the results of the multiple simulations (and possibly conducting some additional simulations), to provide a more general understanding of the spatio-temporal patterns of model responses. At present, they give specific results for their set up, which may be all they are limited to at present. It would be great if they could better say "our results show that, in general, model resolution will tend to have the following effects on model results: (list of effects, with some indication of spatio-temporal trends). This will likely require looking at distributions of model results. Indeed, as demonstrated by the work from Harvard, when looking at air quality and climate impacts, distributions of air quality responses are very informative (e.g., Wu et al., (2008) JGR, DOI: 10.1029/2007JD008917). This is done, to some degree in Fig. 7, but that does not give a spatio-temporal understanding.

Response: As a general comment the authors would like to note that at this point we do not have the necessary computing power to perform additional simulations. Moreover we understand that the reviewer asks for a finer representation of spatiotemporal variation of the trends. It was in fact our initial decision not to provide too refined information regarding the temporal and spatial patterns of the sensitivities because that would produce a large number of contradicting messages amongst various seasons/areas of the domain. This stems from the fact that (and this is the originality of this work) we discuss results on a local high-resolution domain with very large spatial and temporal gradients of e.g., emissions. We have anticipated that the potential reader of the paper would seek a better understanding of more coarse spatiotemporal surrogates such as the ones we provide: urban, suburban, rural and summer (ozone period for ozone), winter, annual. Regarding the list of effects from the various sensitivities presented, our results reveal a consistent pattern in the trends of modeled ozone (decrease) when refined information is implemented. In contrast for fine particles this is not observed. In the initial version of the manuscript this was not clearly discussed. In the updated version the following discussion has been added to the conclusions section (page 4795, line 4):

"*Excluding the sensitivities having the smallest impact (roughly less than 2%, see Table 3) we observe a very consistent trend in ozone concentration: daily average and maximum ozone decrease as input data become more refined, namely passing from climate meteorology to reanalysis, increasing the resolutions of the horizontal and vertical CTM grid, of meteorology, of emissions and by using bottom-up emissions and post-processing instead of top-down. This decrease in ozone concentrations, from 2.5% up to 8.3%, is observed mainly in the urban and suburban areas and in all cases stems from enhanced NOx emission fluxes in the surface-layer leading to titration inhibition. Trends and the underlying changes in emissions are highly variable for PM2.5 with increase in PM2.5 concentrations that may be as low as 2% or as high as 30% for climate meteorology and resolution of the vertical mesh, meteorology and emissions and also cases where concentration decreases in a wide range of values from 3% up to 34% (annual emissions, model resolution) depending on the season.*"

We have also added in the abstract: "*In the case of modelled ozone concentrations, the implementation of refined input data results in a consistent decrease (from 2.5% up to 8.3%), mainly due to inhibition of the titration rate by nitrogen oxides. Such consistency is not observed for PM2.5.*".

(4) One of the most important questions is not addressed by this manuscript: that is, how do these changes (resolution, emissions processing, meteorology) impact how the model responds to emissions changes, e.g., how ozone, PM and NO2 respond to NOx, VOC and SO2 emissions changes.

Response: Actually, we think that this precisely what our paper is all about. With all the sensitivity studies we carried out we try to understand in what way (i.e. by which physical or chemical process) ozone and PM modeling is affected. We will agree though that in some places this is not clear. We have revised the manuscript and added relevant pieces in each sensitivity section. We list here the more striking examples from the manuscript to illustrate this:

From the sensitivity to vertical resolution:

"Interestingly, the impact of the refinement of the vertical grid on daily averaged Ox is much stronger that on ozone: Ox , changes by 0.9ppb in the urban and suburban areas. The change in Ox is reasonable since in VERT, NOx emissions are released within a surface layer thinner by 60% compared to REF (from 20m to 8m) leading to higher NOx concentrations."

From the sensitivity to the annual emission totals:

"Changes in modeled urban daily average ozone concentrations are small ($|\Delta c|$=0.8ppb 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). This is due to the fact that when passing from the AIRPARIF to the ECLIPSE inventory (see also Fig. 2) NOx emissions decrease (weakening titration) and NMVOCs increase (intensifying production)."

From the sensitivity to emission post-processing:

"Modeled PM2.5 sensitivity is significant for both summer and wintertime ($|\Delta c|$=3.4μg/m3 or 24.8% and 4.6μg/m3 or 18.3% respectively) (Table 3). POST wintertime bias is almost two times higher than ANN (Fig. 9b). This is because the coarse resolution annual post-processing coefficients weight towards allocating more of the annual emissions into the winter period significantly influenced by the residential sector emissions which are overstated in the ECLIPSE inventory."

Finally in the conclusions we summarize these effects, please refer to the response of comment 3.

Specific Comments:

(1) The base air quality model was run with 8 layers, with a sensitivity run at 12 layers. Many model are now run, as a base, with significantly more layers (e.g., Simon et al., Environ Sci Technol. 2013 Mar 5;47(5):2304-13. doi: 10.1021 use 24 layers). It would have been of interest to have an even higher vertical resolution analysis.

Response: We agree with the reviewer that it would have been interesting to investigate an additional vertical distribution case. Our decision was a compromise between the computational demands and insights of previous work conducted in the region. Menut et al., 2013b studied the same effect (although using an 1-day episode) and found that the observed change of implementing a refined vertical mesh with 20 layers never exceeded 3μg/m^3 both for ozone and PM$_{10}$ compared to an 8-layer configuration. In any case the increase from 8 to 12 layers might appear small by as we state in the paper: "*VERT implements a 12 vertical σ-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.*" A 23 layer configuration would refine the vertical mesh in the boundary layer but the surface layer to which the study of concentrations is the main focus of this paper would be of similar depth. This is important because our work suggests that emissions are the key input in our simulations and these are almost exclusively released in the surface layer.

(2) The article is opaque at times, e.g., "As regards PM2.5 modeling . . . regional realization canno tselectively incorporate any combination of local scale features. . ." is tough to parse. "By principle" is not standard English. Likewise, many areas still in need of editing for grammar.

Response: We have corrected the aforesaid grammatical mistakes. We have also carefully inspected the manuscript to identify errors in grammar and we have revised accordingly.

(3) They should make clear what they mean by "top-down" vs. "bottom-up" emissions inventories. Some people might use "top-down" to refer to using observations.

Response: We already define the top-down methodology in the introduction (page 4770, line 12): "*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.*". For the bottom-up definition we add in page 4770, line 19: "*In Markakis et al. (2014) we showed that the implementation of bottom-up emissions (e.g., compiled using source specific activity information) in a decade simulation over Paris…*"

(4) Also, the discussion of the weaknesses of various emissions inventory approaches predate Markakis 2010, and should be referenced (e.g., look back at the NARSTO reports, as well as Gilliland, JGR, v. 108).

Response: We referenced the following papers:

Russell, A., and Dennis, R.: NARSTO critical review of photochemical models and modelling. Atmos. Environ., 34, 2261-2282, 2000.

Gilliland, A.B., Dennis, R.L., Roselle, S.J., and Pierce, T.E.: Seasonal NH_3 emission estimates for the eastern United States based on ammonium wet concentrations and an inverse modeling method, J. Geophys. Res., 108(D15), 4477, doi:10.1029/2002JD003063, 2003.

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

K. Markakis[1], M. Valari[1], O. Perrussel[2], O. Sanchez[2], C. Honore[2].

[1] {Laboratoire de Meteorologie Dynamique, IPSL Laboratoire CEA/CNRS/UVSQ, Ecole Polytechnique, 91128 Palaiseau Cedex, France}

[2] {AIRPARIF, Association de surveillance de qualité de l'air en Île-de-France, 7 rue Crillon, 75004, Paris, France}

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

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 10yr model simulation over the greater Paris (France) area at 4km resolution and anthropogenic emissions from a 1km 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 (~8%), the post-processing method (6.5%) and the horizontal resolution of the ~~CTM~~ air quality model (~5%) while annual $PM_{2.5}$ levels are particularly sensitive to changes in their primary emissions (~32%) and the resolution of the emission inventory (~24%). The air quality model horizontal and vertical resolution have little effect on model predictions for the specific study domain. In the case of modelled ozone concentrations, the implementation of refined input data results in a consistent decrease (from 2.5% up to 8.3%), mainly due to inhibition of the titration rate by nitrogen oxides. Such consistency is not observed for $PM_{2.5}$. In contrast this consistency is not observed for $PM_{2.5}$. In addition we use the results of these sensitivities to explain and quantify the discrepancy between a coarse (~50km) and a fine (4km) resolution simulation over the urban area. We show that the ozone bias of the coarse run (+9ppb) is reduced by ~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~~ of $PM_{2.5}$ concentrations follows a more complex pattern with the positive ~~bias~~ values associated with~~to~~ the coarse run (+3.6μg/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~~ its 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; Kelly et al., 2012; 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 we face 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 related uncertainties as well as to CTMs numerical resolution. 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 (Gilliland et al., 2003; Markakis et al., 2010, 2012; Russell and Dennis, 2000). In Markakis et al. (2014) we showed that ozone formation occurs under a VOC-limited chemical regime in the 10-year simulations that used the bottom-up emission inventory. This result is consistent with previous studies over the Paris area (Beekmann and Derognat, 2003; Beekmann and Vautard, 2010; Deguillaume et al., 2008). On the contrary, when the regional top-down inventory was used instead, ozone formation occurred under a NO_x-limited chemical regime. 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. Long-term projections are constrained by the evolution of large scale energy supply and demand and the link between global and regional scale projections is a laborious task (Kelly et al. (2012)).

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 air-quality ~~atmospheric composition~~ studies over finer scales due to different factors such as emission and meteorological input as well as the CTM's horizontal and vertical resolution. We use as reference run a 10yr long simulation (1996-2005) over the Ile-de-France region in France (IdF) at 4km resolution, using the high-resolution (1km) 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 0.5º over Europe, where the 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 versus 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 versus top-down emissions; e) AIRPARIF versus 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 improve~~ing~~ 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° east and 47.89–49.45° north 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^o 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^o and 0.44^o (Vautard et al., 2013). The vertical resolution of the meteorological input consists of 31 σ-p layer extending to 500hPa.

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^o resolution grid (~50 km) over Europe (EEA, 2104) and subsequently at 4km 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 is been benchmarked in the past in a number of model inter-comparison experiments (see Menut et al. (2013a) 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 σ-p hybrid vertical layers from the surface (999hPa) up to approximately 5.5 km (500hPa) with the surface layer being 25m 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 and ~~, including~~ aerosols and was used to drive the ~~global scale~~ LMDz-INCA simulations at the global scale.

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 (as areas sources) 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° x 0.5° resolution grid.

Present-time (2008) emission estimates for the IdF region are also available in hourly basis over a 1km 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_x, Non-methane Volatile Organic Compounds (NMVOCs), SO_2, PM_{10} and $PM_{2.5}$ with a monthly, weekly and diurnal -source specific- temporal resolution. Emissions from point sources are inputted 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 from~~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_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 1km resolution emissions were aggregated to the 4km resolution modeling grid.

2.3 Data and metrics for model evaluation

Model results from the different sensitivity runs are compared against observational data for O_3, NO, NO_2 and $PM_{2.5}$. 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 sites. To benchmark model performance we use the skill score *S* which is based on the equations of Mao et al. (2006):

$$S = \frac{1}{2}\left(1 - \left|\frac{BIAS}{MGE}\right| + \left|\frac{MGE}{RMSE}\right|\right) \qquad (1)$$

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~~ to predict ~~predicting~~ present-time decadal air-quality with respect to emission and meteorological input as well as the CTM's horizontal and vertical resolution. For that purpose we conduct five sets of 10yr long simulations

(1996-2005) over a 4km 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~~of 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 were 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^o 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 (8 σ-p layers). Modeled meteorological fields are further interpolated over the 4km-resolution IdF grid for the air-quality simulation. We note here, that interpolating the 0.44^o resolution meteorology over the 4km 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^o (ERA01) and compare with the ERA05 run.

The third sensitivity test addresses the issue of the CTM’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 the CTM’s 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 σ-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. We note that because the WRF meteorology (resolved in 31 layers) is interpolated to the CTM's vertical grid, technically, increasing the number of vertical layers in CHIMERE from 8 to 12 will result in a refinement of the meteorological input used for the chemical simulations as well.

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 the sensitivity of ozone concentrations in the afternoon peak hour due to surface emissions was the second largest after the sensitivity associated with 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 4km resolution emission dataset over the IdF domain (ANN) in which annual emission fluxes match the ECLIPSE emissions (0.5° resolution) but are downscaled spatially and temporally to obtain 4km-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 by the CTM's chemical mechanism~~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 than AIRPARIF. 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 particle~~s~~ emissions are 2.4 times more in ECLIPSE, which probably stems from the use of a population proxy to spatially allocate wintertime 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~~ should be made clear that ECLIPSE dataset is not meant to accurately represent emissions at such fine scales.

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. Makar et al. (2014) investigated the response of modeled concentrations to the refinement of the spatial and temporal allocation of input emissions and found that the model was as sensitive to these improvements as to the vertical mixing parameterization. Also they conclude that the temporal distribution of emissions in particular, could be very important in stable urban atmospheres and that this sensitivity is reduced with increased mixing conditions. For ~~this~~ our 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 4km-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 on pollutant concentrations 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 30-50km to 4-12km (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 36km to 12km improved the 1h 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

2.5km resolution domain over a 15km 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. This~~ese~~ 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 0.5º 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 4km resolution from the POST run spatially averaged over the 0.5º 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 0.5º resolution (MET_ERA05) and iii) meteorology downscaled from reanalysis data at 0.11º (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 2m-temperature, 10m-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 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°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 8-hour running means (MD8hr) 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 $O_x=O_3+NO_2-0.1*NO_x$ (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 metric~~s~~ 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 +1.8ppb) 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 1ppb. 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 (Szopa et al., 2009; 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 (+5ppb). The small bias in O_x (+0.6ppb) suggests that the problem stems from the representation of local titration and more specifically daytime titration; the daily average ozone bias drops to +3.9ppb while O_x is accurately represented in this case (-0.2ppb). 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=+2.8ppb) and a good skill score (0.73). ~~We~~ The ~~identified~~ two major downwind locations in the IdF domain ~~and found that they represent~~which present the lowest biases (less than 0.1ppb and 1.1ppb for the south-west and north-east directions respectively). The bias of the daytime average reaches +2.1ppb.

Ozone daily maxima in the urban and rural stations are underestimated by 10% (-4.2ppb) and 7% (-3.2ppb) 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 sensitive to temperature compared to the daily averages as shown in Menut at al. (2003). This could also be due to the fact that 4km 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 1μg/m^3). In annual basis the bias is also small while a larger underestimation is predicted for the summertime season (bias=2.8μg/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 t~~This case study~~ is to~~ estimates 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~~ could lead to decreased~~ in the~~ reaction rates, less biogenic emissions and consequently to less ozone. The lower bias in 10m wind speed under MET_ERA05 is bound to increase surface concentrations through ~~lead to less~~reduced 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 (not shown) 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~~ ($|\Delta c| \sim 1$ppb or 3.4%). ~~suggesting a~~ The small improvement of model performance with the reanalysis dataset (ozone decreases through higher NO$_x$ emissions following the PBL scheme described above) is due ~~which stems from~~to 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 ($|\Delta c|=0.1$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 concentrations is 3.1μg/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 ($|\Delta c|=1.4$μg/m^3 or 9.4%). The use of the 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 the accuracy of 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 ($|\Delta c|\sim0.5$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. The 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 ($|\Delta c|=0.3\ \mu g/m^3$ or 3.4%) and on annual basis ($|\Delta c|=0.6\mu g/m^3$ or 4%), but stronger during the winter season ($|\Delta c|=1.3\ \mu g/m^3$ or 6.8%). In fact, wintertime statistics suggest that model bias actually increases with the refinement of the ~~input~~ meteorological grid as a consequence of the reduced modeled precipitation (less scavenging)~~, and wind speed~~ ~~(weaker dispersion)~~ and PBL by 20% (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 $PM_{2.5}$. Never the less this result could reflect the local area's characteristics (flat terrain, situated away from the coast) confirming previous studies (Menut et al., 2005; Valari and Menut, 2008). In regions with more complex topography or close to the coast the resolution of the meteorological input could have a profound effect on the simulated meteorological conditions (Leroyer et al., 2014). We note here that the refinement in the resolution of the meteorological model from 0.5^o to 0.1^o may not be sufficient for the CTM to simulate noticeable concentration responses. For example Leroyer et al. (2014) (see also references therein) observed that substantial changes in vertical and horizontal transport in an urban environment occurred mostly in the transition from resolutions of 2.5km to 1km and even higher (250m).

4.3 Sensitivity to the resolution of the CTM's vertical grid (REF vs. VERT)

This study addresses the impact of the resolution of the CTM's vertical mesh and more specifically of the thickness of the first CTM 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 0.5ppb (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%). 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 0.9ppb 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 20m to 8m) 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) 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 look~~ed~~ how modeled sensitivities change as a function of these parameters (Fig.7).

NO_2 concentrations increase with the refinement of the first vertical layer of the CTM for all vertical mixing conditions (Fig. 7a). However it is only under low vertical mixing ($1< K_z<5$

m^2/s) that ozone sensitivity becomes positive (Fig. 7b). Under stronger turbulence ($K_z > 5\ m^2/s$), 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 of the CTM, enhances deposition rate. These two conflicting effects explain the small impact of the CTM's vertical resolution on $PM_{2.5}$ concentrations.

We conclude that b~~B~~oth 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 versus 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 (|Δc|=0.8ppb 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). This is due to the fact that when passing from the AIRPARIF to the ECLIPSE inventory (see also Fig. 2) NO_x emissions decrease (weakening titration) and NMVOCs increase (intensifying production). This is also seen in the weaker s~~S~~ensitivity ~~is weaker for~~of O_x (0.4ppb 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 (|Δc|= 1.1ppb or 3.2%) and of the same order of magnitude as the sensitivities to climate model driven meteorology and to the resolution of the meteorological input. The weaker change in suburban O_x (|Δc|=0.1ppb or 0.3%) suggests that this area benefits more than the urban area from the improvement in the titration process. The skill score associated to the REF run is also higher by 8% (Fig. 8a). 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|= 0.3ppb or 1%). It is noteworthy that the absolute change in modeled ozone concentrations is ~~on~~ in the order of 1ppb or less despite the large differences in ozone precursors' emissions between the local and the regional inventory.

Changes in the daily average fine particle concentrations in summertime, wintertime and in the annual basis ~~daily average~~ are much stronger than ozone (|Δc|=4.1μg/m^3 or 33%, 6.6μg/m^3 or 33.8% and 5.5μg/m^3 or 31.9% respectively). $PM_{2.5}$ 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 +5.8μg/m^3 showing that fine particle emissions from the ECLIPSE inventory are overestimated (see also Fig. 2). The main source of primary wintertime $PM_{2.5}$ 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 the finer scale. This is consistent to the fact that the summertime bias in the ANN run is much lower (+1.4μg/m^3). In fact, in this case the ANN bias is even smaller than the REF bias (-2.8μg/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 in REF is higher than in~~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 4km-resolution domain ~~and~~ as well as different matrices for their chemical speciation. The ANN dataset uses the AIRPARIF bottom-up approach whereas the EMEP methodology is applied ~~on~~ to 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 the (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~~ sectoral ECLIPSE raw emissions quantified in SNAP level are treated with the respective sectoral coefficients ~~SNAP level EMEP information~~ that stems from the EMEP inventory having a very 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 concentrations over the urban area is considered moderate (|Δc|=1.9ppb or 6.4%) (Fig. 9a and Table 3). Model bias is reduced from +4.5ppb in POST to +2.6ppb 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_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~~ identical ~~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_x emissions than POST enhancing ozone titration. This stems from the fact that the annual emission totals are allocated in the CTM's vertical layers very differently. Following the AIRPARIF post-processing (ANN) all urban emissions are released in the surface layer because according to the local point source emission database no major industrial units are 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_x emissions over~~in~~ Paris in the first model layer.

Another important piece of information ~~of the post-processing of emissions regards their~~id the diurnal variation of emissions. 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 can be to the final modeled concentrations. Fig. 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 10:00LT and 15:00LT, ANN underestimates ozone concentrations due to too much NO emissions, enhancing titration and this is maximized in the local peak (15:00LT) where NO concentrations are overestimated by a factor of 2 (not shown). The daily maximum concentration shows the highest sensitivity in the emission post-treatment among all the presented cases (|Δc|=2.2ppb). This is consistent with Menut et al. (2003) who also found that the afternoon peak concentrations at a typical summertime episode in Paris are~~is~~ very sensitive to the NO emissions change. In the evening (after 15:00LT) ANN deviates from the observations 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 (|Δc|=3.4μg/m^3 or 24.8% and 4.6μg/m^3 or 18.3% respectively) (Table 3). POST wintertime bias is almost two times higher than ANN (Fig. 9b). This is because the coarse resolution annual post-processing coefficients weight towards allocating more of the annual emissions into the winter period significantly influenced by the residential sector emissions which are overstated in the ECLIPSE inventory. 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 20:00LT) are related to the residential heating activity (Fig. 10b).

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. We note here, that recent work has pointed out that the sensitivity of modeled concentrations the spatiotemporal resolution of the emission inventory is model-dependent (Makar et al., 2014).

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 ~~|Δc|=~~2.8ppb or 8.3% (Fig. 11a). The change in daily average O_x is smaller but comparable (|Δc|=1.2ppb or 2.9%) suggesting that ozone titration is not the only model process 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). 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 (AVER). This leads to lower titration rates and therefore, higher ozone levels. Therefore the increase in the resolution of the emission input leads to a reduced positive bias from +7.3ppb (AVER) to +4.5ppb (POST). AVER overestimates ozone peaks by 0.8ppb while POST underestimates them by -1.2ppb. 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 (|Δc|=0.7ppb), with scores practically unchanged (Table 3).

Fine particle concentrations are also very sensitive to the resolution of the emission input, especially in wintertime (|Δc|=7.1μg/m^3 or 30%), with higher concentrations modeled with the refined emission inventory in POST (Table 3). Similarly to ozone t~~T~~his 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 $PM_{2.5}$ concentrations to ~~model~~ the CTM's horizontal resolution. We compare the simulations of~~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° resolution (Fig. 12). REG, models higher ozone concentrations than AVER over the urban area ($|\Delta c|$=1.7ppb 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 has ~~relatively~~ low sensitivity to model resolution ($|\Delta c|$=0.5ppb or 1.4% and 0.2ppb 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~~ increase performance~~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 as a result of the lower primary emissions in REG.

We may conclude that the benefit of increasing the CTM's resolution is insignificant for both ozone and $PM_{2.5}$ especially taking into account the large refinement attempted here (0.5° to 4km).

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 4km resolution using the AIRPARIF bottom-up inventory of anthropogenic emissions (REF) and a regional scale run at 0.5° resolution where emissions stem from the ECLIPSE top-down inventory (REG). Both realizations implement identical climate driven meteorology (at 0.44° resolution) and an 8-layer vertical mesh therefore are susceptible to the same sources of error due to climate model driven meteorology, the resolution of the meteorological input and the resolution of the CTM's 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 9ppb over the city of Paris while the bias of REF is only +1.8ppb (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 $|\Delta c|$=7.2ppb of ozone modeled with the REG run compared to REF (Table 4); 26.4% ($|\Delta c|$=1.9ppb) is related to the post-processing of the annual emissions totals which are based on the EMEP factors, 11.1%

(|Δc|=0.8ppb) to the annual emission totals in the ECLIPSE inventory, 23.6% (|Δc|=1.7ppb) to coarse model resolution and 38.9% (|Δc|=2.8ppb) 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 to the urban ozone response is 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 2.7ppb. Even largest improvement comes through the better spatial representation of ozone precursors emissions in the local emission inventory (|Δc|=2.8ppb) 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 1.7ppb, 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 (|Δc|=2.4ppb) 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, these are better simulated with the REF run with a bias of -0.8μg/m^3 and a high skill score of 0.78 compared to a strong positive bias of +3.6μg/m^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 performance of $PM_{2.5}$ with increase in the bias by 7.1μg/m^3. This only means that if the emission totals from ECLIPSE are used over Paris in the coarse REG application then refining the resolution will only accumulate additional emissions in the city augmenting the modeled concentrations. The remaining features have also a positive effect; model resolution reduces the bias by 0.4μg/m^3, annual emission totals by 6.6μg/m^3 and post-processing of the annual totals by 4.5μg/m^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 a bottom-up post-processing technique would result in an overall bias of the regional application of -0.9μg/m^3.

6 Conclusions

In the present paper we assess the sensitivity of ozone and fine particle concentrations with respect to emission and meteorological input with a 10yr 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 modeling the average daily ozone concentrations was observed in the urban areas primarily due to the resolution of the emission inventory (|Δc|=2.8ppb or 8.3%) and secondly to the post-processing methodology applied on the annual emission totals (|Δc|=1.9ppb 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 (|Δc|=2.2ppb 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 (|Δc|=0.8ppb or 2.5%). Also modeled ozone is fairly insensitive to the use of climate model or reanalysis meteorology. 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~~ in the regional-scale totals of~~in~~ the ECLIPSE inventory;~~.~~ u~~U~~sing 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 the coarser inventory smoothens 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 the CTM'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. We also note the weak sensitivity of modeled concentrations to the increase in the CTM's and the meteorological model's horizontal resolution at least for the area and the range of resolutions studied here.

Excluding the sensitivities having the smallest impact (roughly less than 2%, see Table 3) we observe a very consistent trend in ozone concentration: daily average and maximum ozone decrease as input data become more refined, namely passing from climate meteorology to reanalysis, increasing the resolutions of the horizontal and vertical CTM grid, of meteorology, of emissions and by using bottom-up emissions and post-processing instead of top-down. This decrease in ozone concentrations, from 2.5% up to 8.3%, is observed mainly in the urban and suburban areas and in all cases stems from enhanced NO_x emission fluxes in the surface-layer leading to titration inhibition. Trends and the underlying changes in emissions are highly variable for $PM_{2.5}$ with increase in concentrations that may be as low as 2% or as high as 30% for climate meteorology and resolution of the vertical mesh and also cases where concentration decreases in a wide range of values from 3% up to 34% (annual emissions, model resolution) depending on the season.

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 areas. 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 ($|\Delta c|=7.2$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 (4km or 0.5°) and 11.1% is associated to the annual emissions ~~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 the downscaling coefficients derived from the existing bottom-up modeling systems which combined could reduc~~eing~~ the bias of the regional application by 37.5%. We note here that $PM_{2.5}$ levels in the urban regions are likely mostly controlled by primary emissions; increasing the emissions inventory resolution will concentrate the $PM_{2.5}$ emissions into a smaller spatial extent of the urban area (the reverse side of the artificial dilution issue taking place at coarse resolution); if the emissions totals are themselves biased high, then the resulting error will only become apparent at higher resolution. Therefore, the emissions resolution may be showing that the emissions totals are too high, and this only becomes apparent at high resolutions.

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 3.6μg/m^3 during wintertime compared to the fine scale simulation (REF) showing a bias of -0.8μg/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 (REG-REF=+7.1μg/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 10.7μg/m^3), neither the implementation of the annual emissions of the bottom-up inventory alone (REG-REF=-6.6μg/m^3) which would generate an overall negative bias of 3μg/m^3. A simpler action would be to integrate the post-processing bottom-up technique (REG-REF=-4.5μg/m^3) giving an overall bias in REG of -0.9μg/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

AIRPARIF, Evaluation Prospective des emissions et des concentrations des pollutants atmospheriques a l'horizon 2020 en Ile-De-France – Gain sur les emissions en 2015. Report available at http://www.airparif.asso.fr/_pdf/publications/ppa-rapport-121119.pdf, 2012.

Amann, M., Bertok, I., Borken-Kleefeld, J., Cofala, J., Heyes, C., Höglund-Isaksson, L., Klimont, Z., Nguyen, B., Posch, M., Rafaj, P., Sandler, R., Schöpp, W., Wagner, F., and Winiwarter, W.: Cost-effective control of air quality and greenhouse gases in Europe: Modeling and policy applications, Environ. Model. Software, 26, 1489-1501, 2011.

Arunachalam, S., Holland, A., Do, B., and Abraczinskas, M.: A quantitative assessment of the influence of grid resolution on predictions of future year air quality in North Carolina, USA, Atmos. Environ., 40, 5010-5026, 2006.

Beekmann, M., and Derognat, C.: Monte Carlo uncertainty analysis of a regional-scale transport chemistry model constrained by measurements from the Atmospheric Pollution Over the Paris Area (ESQUIF) campaign, J. Geophys. Res., 108(D17), 8559, doi:10.1029/2003JD003391, 2003.

Beekmann, M., and Vautard, R.: A modeling study of photochemical regimes over Europe: robustness and variability, Atmos. Chem. Phys., 10, 10067-10084, 2010.

Bessagnet B., Seigneur, C., and Menut, L.: Impact of dry deposition of semi-volatile organic compounds on secondary organic aerosols, Atmos. Environ., 44, 1781–1787, 2010.

Brasseur, G. P., Jeffrey, T. K., Muller, J.-F., Schneider, T., Granier, C., Tie, X., and Hauglustaine, D.: Past and future changes in global tropospheric ozone: impact and radiative forcing, Geophys. Res. Lett., 25, 3807–3810, 1998.

Cohan, D.S., Hu, Y., and Russell, A.G.: Dependence of ozone sensitivity analysis on grid resolution, Atmos. Environ., 40, 126–135, 2006.

Colette, A., Granier, C., Hodnebrog, O., Jakobs, H., Maurizi, A., Nyiri, A., Rao, S.**,** Amann, M., Bessagnet, B., D'Angiola, A., Gauss, M., Heyes, C.**,** Klimont, Z., and Meleux, F.: Future air quality in Europe: a multi-model assessment of projected exposure to ozone, Atmos. Chem. Phys., 12, 10613-10630, 2012.

Colette, A., Bessagnet, B., Vautard, R., Szopa, S., Rao, S., Schucht, S., Klimont, Z., Menut, L., Clain, G., Meleux, F., Curci, G., and Rouïl, L.: European atmosphere in 2050, a regional air quality and climate perspective under CMIP5 scenarios, Atmos. Chem. Phys., 13, 7451–7471, doi:10.5194/acp-13-7451-2013, 2013.

Clarke, L., J. Edmonds, H. Jacoby, H. Pitcher, Reilly, R., and R. Richels.: Scenarios of Greenhouse Gas Emissions and Atmospheric Concentrations. Sub-report 2.1A of Synthesis and Assessment Product 2.1 by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research. Department of Energy, Office of Biological & Environmental Research, Washington, 7 DC., USA, pp. 154, 2007.

Deguillaume, L., Beekmann, M., and Derognat, C.: Uncertainty evaluation of ozone production and its sensitivity to emission changes over the Ile-de-France region during summer periods, J. Geophys. Res., 113, D02304, doi:10.1029/2007JD009081, 2008.

Dufresne, J.-L., Foujols, M.-A., Denvil, S., Caubel, A., Marti, O., Aumont, O., Balkanski, Y., Bekki, S., Bellenger, H., Benshila, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., de Noblet, N., Duvel, J.-P Ethé, C., Fairhead, L., Fichefet, T., Flavoni, S., Friedlingstein, P., Grandpeix, J.-Y., Guez, L., Guilyardi, E., Hauglustaine, D., Hourdin, F., Idelkadi, A., Ghattas, J., Joussaume, S., Kageyama, M., Krinner, G., Labetoulle, S., Lahellec, A., Lefebvre, M.-P., Lefevre, F., Levy, C., Li, Z. X., Lloyd, J., Lott, J., Madec, G., Mancip, M., Marchand, M., Masson, S., Meurdesoif, Y., Mignot, J., Musat, I., Parouty, S., Polcher, J., Rio, C., Schulz, M., Swingedouw, D., Szopa, S., Talandier, C., Terray, P., Viovy, N., and Vuichard, N.: Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5, Clim. Dynam., 40, 2123–2165, 2013.

EEA Technical Report, Towards integrated assessment of air quality and climate policies, in review, 2014.

EPA, Guidance on the Use of Models and Other Analyses for Demonstrating Attainment of Air Quality Goals for Ozone, PM2.5 and Regional Haze. EPA -454/B-07-002, April, 2007.

Fiore, A.M., Naik, V., Spracklen, D.V., Steiner, A., Unger, N., Prather, M., Bergmann, D., Cameron-Smith, P.J., Cionni, I., Collins, W.J., Dalsøren, S., Eyring, V., Folberth, G.A., Ginoux, P., Horowitz, L.W., Josse, B., Lamarque, J.-F., MacKenzie, I.A., Nagashima, T., O’Connor, F.O., Righi, M., Rumbold, S.T., Shindell, D.T., Skeie, R.B., Sudo, K., Szopa, S., Takemura, T., and Zeng, G.: Global air quality and climate, Chem. Soc. Rev., 41, 6663-6683, 2012.

Flagg D.D., and Taylor, P.A.: Sensitivity of mesoscale model urban boundary layer meteorology to the scale of urban representation, Atmos. Chem. Phys., 11, 2951-2972, 2011.

Forkel, R., and Knoche, R.: Regional climate change and its impact on photo-oxidant concentrations in southern Germany: Simulations with a coupled regional climate–chemistry model, J. Geophys. Res., 111, D12302, doi:10.1029/2005JD006748, 2006.

Forkel, R., and Knoche, R.: Nested regional climate–chemistry simulations for central Europe, CR Geosci., 339, 734-746, 2007.

Gilliland, A.B., Dennis, R.L., Roselle, S.J., and Pierce, T.E.: Seasonal NH3 emission estimates for the eastern United States based on ammonium wet concentrations and an inverse modeling method, J. Geophys. Res., 108(D15), 4477, doi:10.1029/2002JD003063, 2003.

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.

Hauglustaine, D. A., Hourdin, F., Jourdain, L., Filiberti, M.-A., Walters, S, Lamarque, J.F., and Holland, E.A.: Interactive chemistry in the Laboratoire de Meteorologie Dynamique general circulation model: Description and background tropospheric chemistry evaluation, J. Geophys. Res., 190, D04314, doi:10.1029/2003JD003957, 2004.

Hauglustaine, D.A., Balkanski, Y., and Schulz, M.: A global model simulation of present and future nitrate aerosols and their direct radiative forcing of climate, Atmos. Chem. Phys., 14, 6863-6949, 2014.

Hodzic, A., Jimenez, J.L., Madronich, S., Canagaratna, M.R., DeCarlo, P.F., Kleinman, L., and Fast, J.: Modeling organic aerosols in a megacity: potential contribution of semi-volatile and intermediate volatility primary organic compounds to secondary organic aerosol formation, Atmos. Chem. Phys., 10, 5491-5514, 2010.

Hogrefe, C., Lynn, B., Civerolo, K., Ku, J.-Y., Rosenthal, J., Rosenzweig, 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.

Im, U., Markakis, K., Poupkou, A., Melas, D., Unal, A., Gerasopoulos, E., Daskalakis, N., Kindap, T., and Kanakidou, M.: The impact of temperature changes on summer time ozone and its precursors in the Eastern Mediterranean, Atmos. Chem. Phys., 11, 3847–3864, 2011.

Im, I., Markakis, K., Koçak, M., Gerasopoulos, E., Daskalakis, D., Mihalopoulos, N., Poupkou, A., Tayfun, K., Unal, U., and Kanakidou, M.: Summertime aerosol chemic al composition in the Eastern Mediterranean and its sensitivity to temperature, Atmos. Environ., 50, 164-173, 2012.

Jacob, D.J., and Winner, D.A.: Effect of climate change on air quality, Atmos. Environ., 43, 51-63, 2009.

Jimenez, P.A., and Dudhia, J.: Improving the Representation of Resolved and Unresolved Topographic Effects on Surface Wind in the WRF Model, J. Applied Meteorol. Climatol., 51, 300-316, 2012.

Jones, P.D., Lister, D.H., Osborn, T.J., Harpham, C., Salmon, M., and Morice, C.P.: Hemispheric and large-scale land-surface air temperature variations: An extensive revision and an update to 2010, J. Geophys. Res., 117, 2012. http://dx.doi.org/10.1029/2011JD017139.

Katragkou E., Zanis, P., Kioutsioukis, I., Tegoulias, I., Melas, D., Krüger, B.C., and Coppola, E.: Future climate change impacts on summer surface ozone from regional climate-air quality simulations over Europe, J. Geophys. Res., 116, D22307, doi:10.1029/2011JD015899, 2011.

Kelly, J., Makar, P.A., and Plummer, D.A.: Projections of mid-century summer air-quality for North America: effects of changes in climate and precursor emissions, Atmos. Chem. Phys., 12, 5367–5390, 2012.

Klimont, Z., Kupiainen, K., Heyes, C., Cofala, J., Rafaj, P., Höglund-Isaksson, L., Borken, J., Schöpp, W., Winiwarter, W., Purohit, P., Bertok, I., and Sander, R.: ECLIPSE V4a: Global emission data set developed with the GAINS model for the period 2005 to 2050. Key features and principal data sources, 2013. Available at http://eccad.sedoo.fr/eccad_extract_interface/JSF/page_login.jsf.

Klimont, Z., L. Hoglund, Heyes, Ch., Rafaj, P., Schoepp, W., Cofala, J., Borken-Kleefeld, J., Purohit, P., Kupiainen, K., Winiwarter, W., Amann., M , Zhao, B., Wang, S.X., Bertok, I., and Sander, R.: Global scenarios of air pollutants and methane: 1990-2050. In preparation.

Knowlton, K., Rosenthal, J.E., Hogrefe, C., Lynn, B., Gaffin, S., Goldberg, R., Rosenzweig, C., Civerolo, K., Ku, J.-Y., and Kinney, P.L.: Assessing Ozone-Related Health Impacts under a Changing Climate, Environ. Health Perspect., 112 (15), 1557-1563, 2004.

Kotlarski, S., Keuler, K., Christensen, O.B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D., van Meijgaard, E., Nikulin, G., Schär, C., Teichmann, C., Vautard, R., Warrach-Sagi, K., and Wulfmeyer, V.: Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble, Geosci. Model Dev., 7, 1297-1333, 2014.

Lam, Y.F., Fu, J.S., Wu, S., and Mickley, L.J.: Impacts of future climate change and effects of biogenic emissions on surface ozone and particulate matter concentrations in the United States, Atmos. Chem. Phys., 11, 4789–4806, 2011.

Langner, J., 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, 2005.

Langner, J., Engardt, M., Baklanov, A., Christensen, J. H., Gauss, M., Geels, C., Hedegaard, G. B., Nuterman, R., Simpson, D., Soares, J., Sofiev, M., Wind, P., and Zakey, A.: A multi-model study of impacts of climate change on surface ozone in Europe, Atmos. Chem. Phys., 12, 10423-10440, 2012.

Lattuati, M.: Contribution a l’etude du bilan de l’ozone tropospherique a l’interface de l’Europe et de l’Atlantique Nord: modelisation lagrangienne et mesures en altitude, Phd thesis, Universite P. M. Curie, Paris, France, 1997.

Leroyer, S., Bélair, S., Husain, S.Z., and Mailhot, J.: Sub-Kilometer Numerical Weather Prediction in an Urban Coastal Area: A Case Study over the Vancouver Metropolitan Area, App. Met., and Clim., doi:10.1175/JAMC-D-13-0202-2013, 2014.

Liao, H., Chen, W.-T., and Seinfeld, J.H.: Role of climate change in global predictions of future tropospheric ozone and aerosols, J. Geophys. Res. 111, doi:10.1029/2005JD006862, 2006.

Loomis, D., Grosse, Y., Lauby-Secretan, B., El Ghissassi, F., Bouvard, V., Benbrahim-Tallaa, L., Guha, N., Baan, R., Mattock, H., and Straif, K.: The carcinogenicity of outdoor air pollution. (On behalf of the International Agency for Research on Cancer Monograph Working Group IARC, Lyon, France.) The Lancet Oncology, 2013.

Mailler, S., Khvorostyanov, D., and Menut, L.: Impact of the vertical emission profiles on background gas-phase pollution simulated from the EMEP emissions over Europe, Atmos. Chem. Phys., 13, 5987-5998, 2013.

Makar, P.A., Nissen, R., Teakles, A., Zhang, J., Zheng, Q., Moran, M.D., Yau, H., and diCenzo, C.: Turbulent transport, emissions and the role of compensating errors in chemical transport models, Geosci. Model Dev., 7, 1001-1024, 2014.

Mao, Q., Gautney, L.L., Cook, T.M., Jacobs, M.E., Smith, S.N., and Kelsoe, J.J.: Numerical experiments on MM5-CMAQ sensitivity to various PBL schemes, Atmos. Environ., 40, 3092–3110, 2006.

Markakis, K., Poupkou, A., Melas, D., and Zerefos, C.: A GIS based methodology for the compilation of an anthropogenic PM_{10} emission inventory in Greece, Atmos. Poll. Res., 1 (2), 71-81, 2010.

Markakis, K., Im, U., Unal, A., Melas, D., Yenigün, O., and İncecik, S.: Compilation of a GIS based high spatially and temporally resolved emission inventory for the İstanbul Greater Area, Atmos. Poll. Res., 3, 112-125, 2012.

Markakis, K., Valari, M., Colette, A., Sanchez, O., Perrussel, O., Honore, C., Vautard, R., Klimont, Z., and Rao, S.: Air quality in the mid-21st century for the city of Paris under two climate scenarios; from the regional to local scale, Atmos. Chem. Phys. 14, 7323-7340, 2014.

Menut, L.: Adjoint modeling for atmospheric pollution process sensitivity at regional scale, J. Geophys. Res., 108(D17), 8562, doi:10.1029/2002JD002549, 2003.

Menut, L., Coll, I., and Cautenet, S.: Impact of meteorological data resolution on the forecasted ozone concentrations during the ESCOMPTE IOP2a and IOP2b, Atmos. Res., 74, 139–159, 2005.

Menut, L., Goussebaile, A., Bessagnet, B., Khvorostiyanov, D., and Ung, A.: Impact of realistic hourly emissions profiles on air pollutants concentrations modeled with CHIMERE, Atmos. Environ., 49, 233-244, 2012a.

Menut, L., Tripathi, Om P., Colette, A., Vautard, R., Flaounas, E., and Bessagnet, B.: Evaluation of regional climate simulations for air quality modeling purposes, Clim. Dynam., 40 (9-10), 2515-2533, 2012b.

Menut, L., Bessagnet, B., Khvorostiyanov, D., Beekmann, M., Colette, A., Coll, I., Gurci, G., Foret, G., Mailler, S., Monge, J.-L., Turquety, S., Valari, M., Vautard, R., and Vivanco, M.G.: Regional atmospheric composition modeling with CHIMERE, Geosci. Model Dev., 6, 203-329, 2013a.

Menut, L., Bessagnet, B., Colette, A., and Khvorostiyanov, D.: On the impact of the vertical resolution on chemistry-transport modeling, Atmos. Environ., 67, 370-384, 2013b.

Nenes, A., Pilinis, C., and Pandis, S.: ISORROPIA: A new thermodynamic model for inorganic multicomponent atmospheric aerosols, Aquatic Geochem., 4, 123–152, 1998.

Nolte, C.G., Gilliland, A.B., Hogrefe, C., and Mickley, L.J.: Linking global to regional models to assess future climate impacts on surface ozone levels in the United States, J. Geophys. Res., 113, D14307, doi:10.1029/2007JD008497, 2008.

Pascal, M., Corso, M., Chanel, O., Declecq, C., Badaloni, C., Cesaroni, G., Henschel, S., Maister, K., Haluza, D., Martin-Olmedo, P., and Medina S.: Assessing the public health impact of urban air pollution in 25 European cities: results of the Aphekom project, Sci. Total Environ., 449, 390-400, 2013.

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., van Weele, M., and Wild, O.: Fresh air in the 21st century?, Geophys. Res. Lett., 30, 1100, doi:10.1029/2002GL016285, 2003.

Punger, M., and West, J.: The effect of grid resolution on estimates of the burden of ozone and fine particulate matter on premature mortality in the USA, Air. Qual. Atmos. Health, DOI 10.1007/s11869-013-0197-8, 2013.

Queen, A., and Zhang, Y.: Examining the sensitivity of MM5–CMAQ predictions to explicit microphysics schemes and horizontal grid resolutions, Part III - The impact of horizontal grid resolution, Atmos. Environ., 42, 3869-3881, 2008.

Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N., and Rafaj, P.: RCP 8.5-A scenario of comparatively high greenhouse gas emissions, Clim. Change, 109, 33-57, 2011.

Russell, A., and Dennis, R.: NARSTO critical review of photochemical models and modelling, Atmos. Environ., 34, 2261-2282, 2000.

Sadanaga, Y., Shibata, S., Hamana, M., Takenaka, N., and Bandow, H.: Weekday/Weekend Difference of Ozone and its Precursors in Urban Areas of Japan, Focusing on Nitrogen Oxides and Hydrocarbons, Atmos. Environ., 42, 4708–4723, 2008.

Scire, J.S., Strimaitis, D.G., and Yamartino, R.J.: Model formulation and user's guide for the CALPUFF dispersion model. Sigma Research Corp., Concord, MA, 1990.

Sillman, S., Logan, J.A., and Wofsy, S.C.: A regional scale model for ozone in the United States with subgrid representation of urban and power plant plumes, J. Geophys. Res. 95, 5731–5748, 1990.

Sillman, S., and Samson, F.L.: Impact of temperature on oxidant photochemistry in urban, polluted rural and remote environments, J. Geophys. Res., 100 (D6), 11, 497–11, 508, 1995.

Skamarock, W.C., and Klemp J.B.: A time-split non-hydrostatic atmospheric model, J. Comput. Phys., 227, 3465–3485, 2008.

Solazzo, E., Bianconi, R., Pirovano, G., Volker, M., Vautard, R., Appel, K. W., Bessagnet, B., Brandt, J., Christensen, J. H., Chemel, C., Coll, I., Ferreira, J., Forkel, R., Francis, X. V., Grell, G., Grossi, P., Hansen, A., Miranda, A. I., Moran, M. D., Nopmongcol, U., Parnk, M., Sartelet, K. N., Schaap, M., Silver, J. D., Sokhi, R. S., Vira, J., Werhahn, J., Wolke, R., Yarwood, G., Zhang, J., Rao, S. T., and Galmarini, S.: Operational model evaluation for particulate matter in Europe and North America in the context of the AQMEII project, Atmos. Environ., 53, 75–92, 2012.

Stroud, C.A., Makar, P.A., Moran, M.D., Gong, W., Gong, S., Zhang, J., Hayden, K., Mihele, C., Brook, J.R. Abbatt, J.P.D., and Slowik, J.G.: Impact of model grid spacing on regional- and urban- scale air quality predictions of organic aerosol, Atmos, Chem. Phys., 11, 3107–3118, 2011.

Szopa, S., Hauglustaine, D.A., Vautard, R., and Menut, L.: Future global tropospheric ozone changes and impact on European air quality, Geophys. Res. Lett., 33, L14805, doi:10.1029/2006GL025860, 2006.

Szopa, S., and Hauglustaine, D.: Relative impacts of worldwide tropospheric ozone changes and regional emission modifications on European surface-ozone levels, CR Geosci., 339, 709-720, 2007.

Szopa, S., Foret G., Menut, L., and Cozic, A.: Impact of large scale circulation on European summer surface ozone and consequences for modeling forecast, Atmos. Environ., 43, 1189-1195, 2009.

Tagaris, E., Liao, K.-J., DeLucia, A.J., Deck, L., Amar, P., and Russell, A.G.: Potential impact of climate change on air pollution-related human health effects, Environ. Sci. Technol., 43, 4979-4988, 2009.

Tie, X., Brasseur, G., and Ying, Z.: Impact of model resolution on chemical ozone formation in Mexico City: application of the WRF-Chem model, Atmos. Chem. Phys., 10, 8983–8995, 2010.

Valari, M., and Menut, L.: Does an Increase in Air Quality Models' Resolution Bring Surface Ozone Concentrations Closer to Reality?, J. Atmos. Ocean. Technol., 25, 1955-1968, 2008.

van Loon, M., Vautard, R., Schaap, M., Bergström, R., Bessagnet, B., Brandt, J., Builtjes, P. J. H., Christensen, J. H., Cuvelier, C., Graff, A., Jonson, J. E., Krol, M., Langner, J., Roberts, P., Rouil, L., Stern, R., Tarrasón, L., Thunis, P., Vignati, E., White, L., and Wind, P.: Evaluation of long-term ozone simulations from seven regional air quality models and their ensemble, Atmos. Environ., 41, 2083-2097, 2007.

van Vuuren, D., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S., and Rose, S.: The representative concentration pathways: an overview, Clim. Change, 109, 5-31, 2011.

Vautard, R., Bessagnet, B., Chin, M., and Menut, L.: On the contribution of natural Aeolian sources to particulate matter concentrations in Europe: testing hypotheses with a modeling approach, Atmos. Environ., 39, 3291–3303, 2005.

Vautard R., Builtjes, P.H.J., Thunis, P., Cuvelier, C., Bedogni, M., Bessagnet, B., Honore, C., Moussiopoulos, N., Pirovano, G., Schaap, M., Stern, R., Tarasson, L., and Wind P.: Evaluation and intercomparison of Ozone and PM10 simulations by several chemistry transport models over four European cities within the CityDelta project, Atmos. Environ., 41 (1), 173-188, 2007.

Vautard, R., Moran, M. D., Solazzo, E., Gilliam, R. C., Matthias, V., Bianconi, R., Chemel, C., Ferreira, J., Geyer, B., Hansen, A. B., Jericevic, A., Prank, M., Segers, A., Silver, J. D., Werhahn, J., Wolke, R., Rao, S.T., and Galmarini, S.: Evaluation of the meteorological forcing used for the Air Quality Model Evaluation InternationaI Initiative (AQMEII) air quality simulations, Atmos. Environ., doi:10.1016/j.atmosenv.2011.10.065, 2012.

Vautard, R., Gobiet, A., Jacob, D., Belda, M., Colette, A., Déqué, M., Fernandez, J., Garcia-Diez, M., Goergen, K., Güttler, I., Halenka, T., Karacostas, T., Katragkou, E., Keuler, K., Kotlarski, S., Mayer, S., Meijgaard, E., Nikulin, G., Patarcic, M., Scinocca, J., Sobolowski, S., Suklitsch, M., Teichmann, C., Warrach-Sagi, K., Wulfmeyer, V., and Yiou, P.: The simulation of European heat waves from an ensemble of regional climate models within the EURO-CORDEX project, Clim. Dyn., 41, 2555-2575, 2013.

Zanis P., Katragkou, E., Tegoulias, I., Poupkou, A., Melas, D., Huszar, P., and Giorgi, F.: Evaluation of near surface ozone in air quality simulations forced by a regional climate model over Europe for the period 1991-2000, Atmos. Environ., 45, 6489-6500, doi:10.1016/j.atmosenv.2011.09, 2011.

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

	Annual emission totals[a]	Air-quality model resolution	Emission inventory resolution	Emission post-processing[b]	climate/reanalysis meteorology and resolution	Number of layers in air-quality model
REF	AIRPARIF	4km	4km	Bottom-up	RCP-4.5 (0.44°)	8
REG[c]	ECLIPSE	0.5°	0.5°	Top-down	RCP-4.5 (0.44°)	8
Sensitivity simulation						
ERA05	AIRPARIF	4km	4km	Bottom-up	ERA (0.44°)	8
ERA01[d]	AIRPARIF	4km	4km	Bottom-up	ERA (0.11°)	8
VERT	AIRPARIF	4km	4km	Bottom-up	RCP-4.5 (0.44°)	12
ANN	ECLIPSE	4km	4km	Bottom-up	RCP-4.5 (0.44°)	8
POST[e]	ECLIPSE	4km	4km	Top-down	RCP-4.5 (0.44°)	8
AVER[f]	ECLIPSE	4km	0.5°	Top-down	RCP-4.5 (0.44°)	8

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

[b] Temporal, vertical allocation and chemical speciation.

[c] This simulation is used as boundary conditions for all local scale simulations.

[d] The ERA01 simulation is compared with the ERA05 not with the REF.

[e] The POST simulation is compared with the ANN not with the REF.

[f] This is not a standalone simulation. Concentrations modeled at 4km 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 Ile-de-France region. MET_CLIM dataset stems from a climate model and MET_ERA05, MET_ERA01 from reanalysis data at 0.5^o and 0.1^o resolution respectively. Absolute model bias is given in parenthesis.

Variable	***Obs***	***MET_CLIM***	***MET_ERA05***	***MET_ERA01***
Summer (JJA)				
T2 (^{o}C)	19.19	19.14 (-0.05)	18.28 (-0.91)	18.19 (-1.0)
WS10 (m/s)	2.9	4.0 (+1.1)	3.8 (+0.9)	3.8 (+0.9)
RH (%)	69.1	68.1 (-1.0)	68.3 (-0.8)	67.3 (-1.8)
PRECIP (mm/day)	0.076	0.108 (+0.032)	0.097 (+0.021)	0.098 (+0.022)
Winter (DJF)				
T2 (^{o}C)	4.3	4.0 (-0.3)	6.0 (+1.7)	5.8 (+1.3)
WS10 (m/s)	3.6	6.2 (+2.6)	5.7 (+2.1)	5.5 (+1.9)
RH (%)	85.0	80.3 (-4.7)	79.7 (-5.3)	79.5 (-5.5)
PRECIP (mm/day)	0.069	0.112 (+0.043)	0.089 (+0.02)	0.087 (+0.018)

Table 3. Absolute difference (and percentage in parenthesis) between daily averaged ozone (ppb) and PM$_{2.5}$ (μg/m^3) from two climate forced air-quality runs. The most influential factor for each sensitivity test is marked in bold.

Ozone	**Urban**	**Suburban**	**Rural**
Climate meteo (REF vs. ERA05)	1.0 (3.4%)	1.1 (3.2%)	**0.9 (2.5%)**
Meteo. resolution (ERA05 vs. ERA01)	0.2 (0.6%)	**1.4 (4.3%)**	0.3 (0.8%)
Vertical resolution (REF vs. VERT)	0.3 (1.2%)	<0.1 (0.2%)	<0.1 (1.5%)
Annual emis. totals (REF vs. ANN)	0.8 (2.5%)	1.1 (3.2%)	0.3 (1.0%)
Emission post-proc. (ANN vs. POST)	1.9 (6.4%)	0.1 (0.4%)	<0.1 (0.02%)
Emission resolution (POST vs. AVER)	**2.8 (8.3%**)	0.7 (1.9%)	0.2 (0.5%)
Model resolution (AVER vs. REG)	1.7 (4.7%)	0.5 (1.4%)	0.2 (0.5%)
PM$_{2.5}$	**Summer**	**Winter**	**Annual**
Climate meteo (REF vs. ERA05)	<0.1 (0.05%)	3.1 (17.6%)	1.4 (9.4%)
Meteo. resolution (ERA05 vs. ERA01)	0.3 (3.4%)	1.3 (6.8%)	0.6 (4.0%)
Vertical resolution (REF vs. VERT)	<0.1 (0.3%)	0.5 (2.2%)	<0.1 (0.2%)
Annual emis. totals (REF vs. ANN)	**4.1** (**33.0%**)	6.6 (**33.8%**)	**5.5** (**31.9%**)
Emission post-proc. (ANN vs. POST)	3.4 (24.8%)	4.5 (18.3%)	0.2 (0.7%)
Emission resolution (POST vs. AVER)	2.1 (20.3%)	**7.1** (30.0%)	4.3 (24.2%)
Model resolution (AVER vs. REG)	0.4 (4.1%)	0.4 (1.9%)	0.7 (0.5%)

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).

Ozone	**Ozone (ppb)**	**$PM_{2.5}$ ($\mu g/m^3$)**
REG (50km)	**+9.0**	**+3.6**
Model resolution	*-1.7*	*-0.4*
Emissions resolution	*-2.8*	*+7.1*
Annual emission totals	*-0.8*	*-6.6*
Emissions post-processing	*-1.9*	*-4.5*
REF (4km)	**+1.8**	**-0.8**

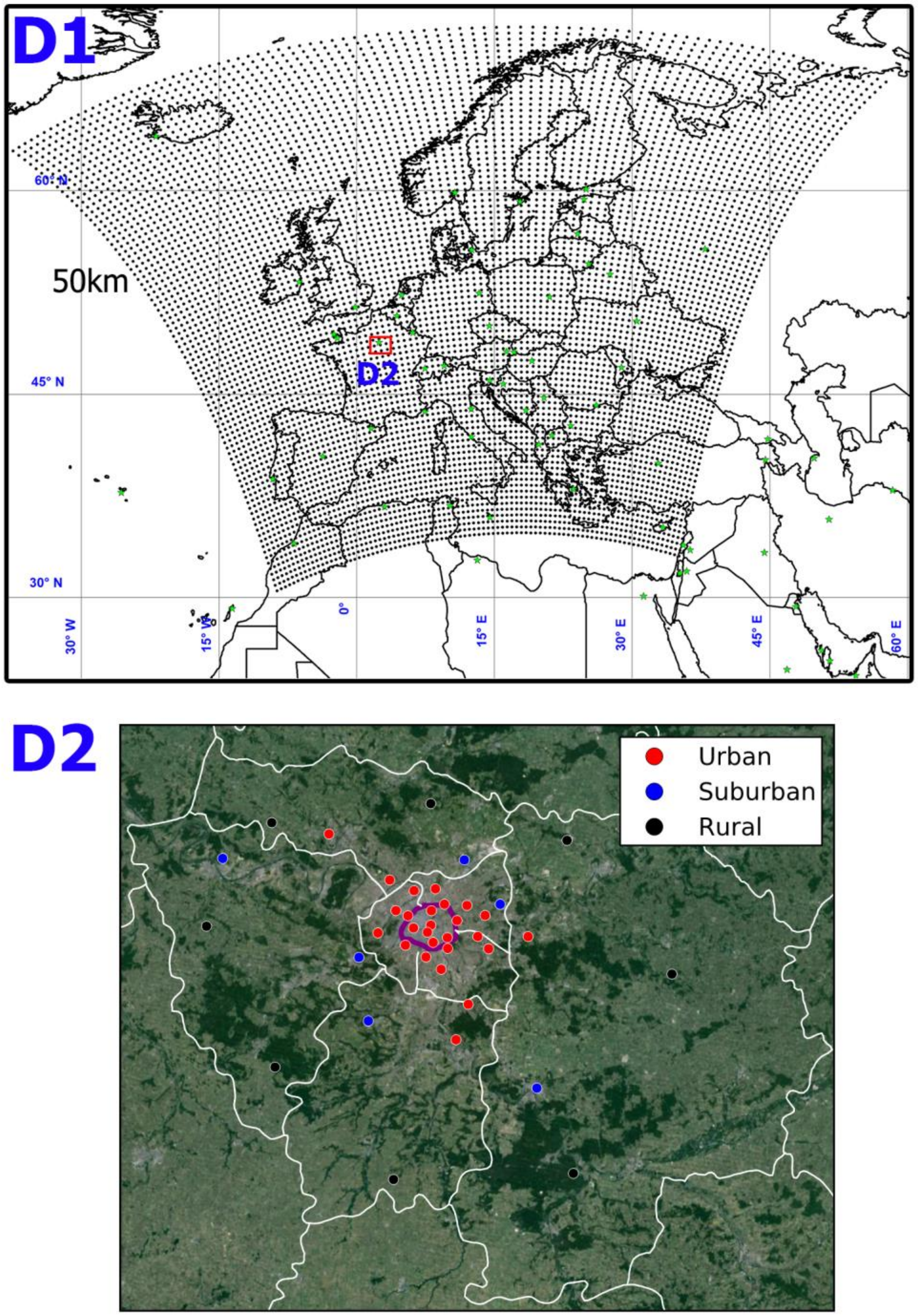

Figure 1. Overview of the coarse (D1 having 50km resolution) and local scale (D2, illustrated by the red rectangle having 4km resolution) simulation domains. In D2 the city of Paris in located in the area enclosed by the purple line. Circles correspond to sites of the local air-quality monitoring network (AIRPARIF) with red for urban, blue for suburban and black for rural.

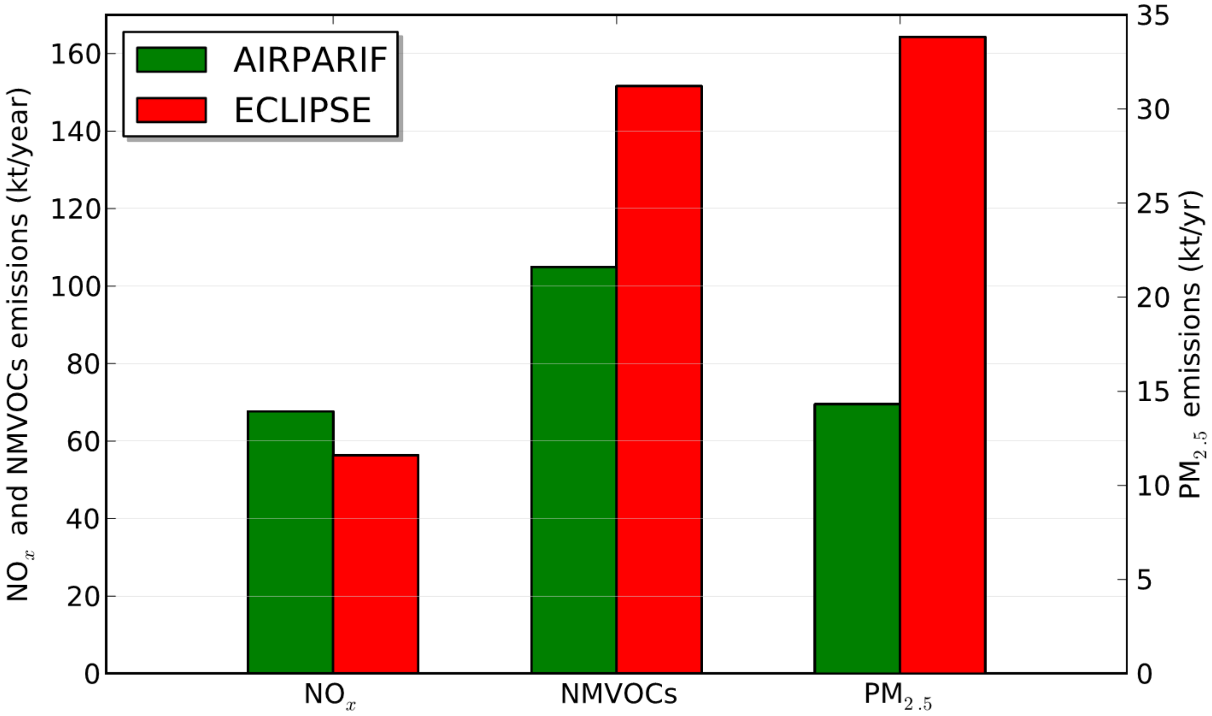

Figure 2. Domain-wide annual emissions of NO_x, NMVOC (left-axis) and $PM_{2.5}$ (right-axis) from the local (bottom-up) and the regional (top down) inventory (summed across the vertical column).

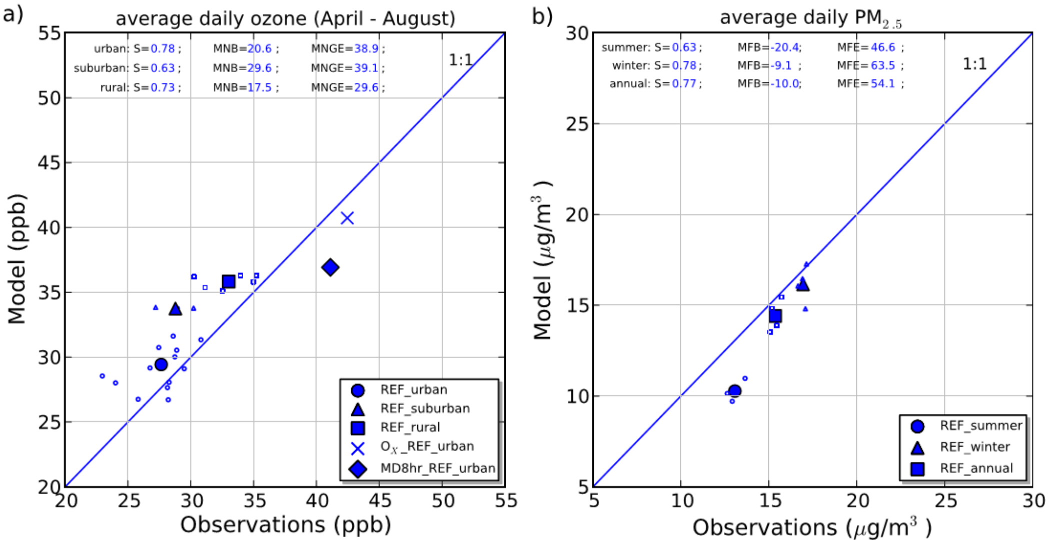

Figure 3. Panel a: Scatter plots and scores of daily average ozone concentrations at urban, suburban and rural stations from the REF simulation. Odd oxygen (O_x) and daily maximum values at urban locations are also shown. Panel b: daily average $PM_{2.5}$ concentrations in wintertime (DJF), summertime (JJA) and on annual basis over urban stations.

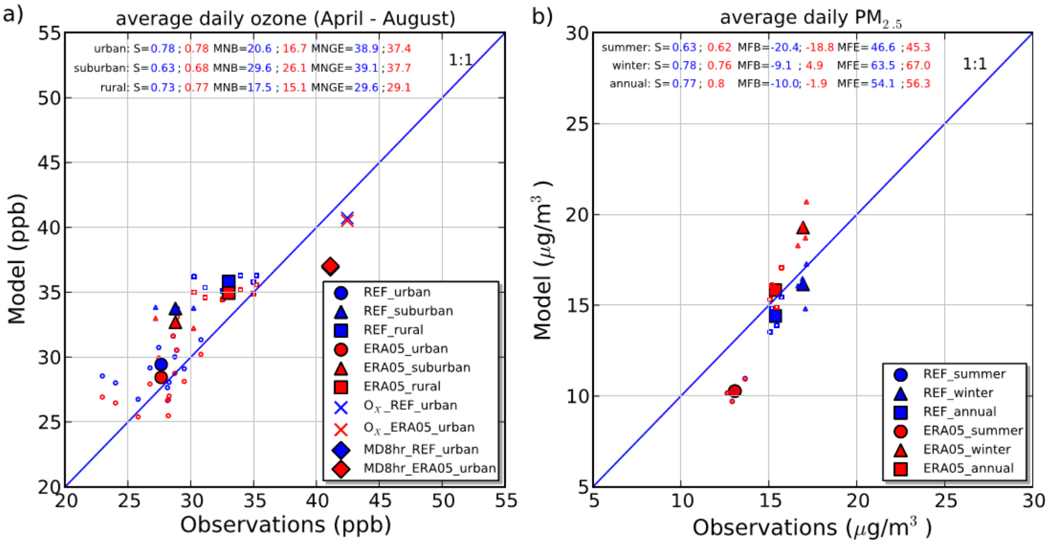

Figure 4. Scatter plots and scores for the sensitivity test on climate model driven meteorology for ozone and $PM_{2.5}$.

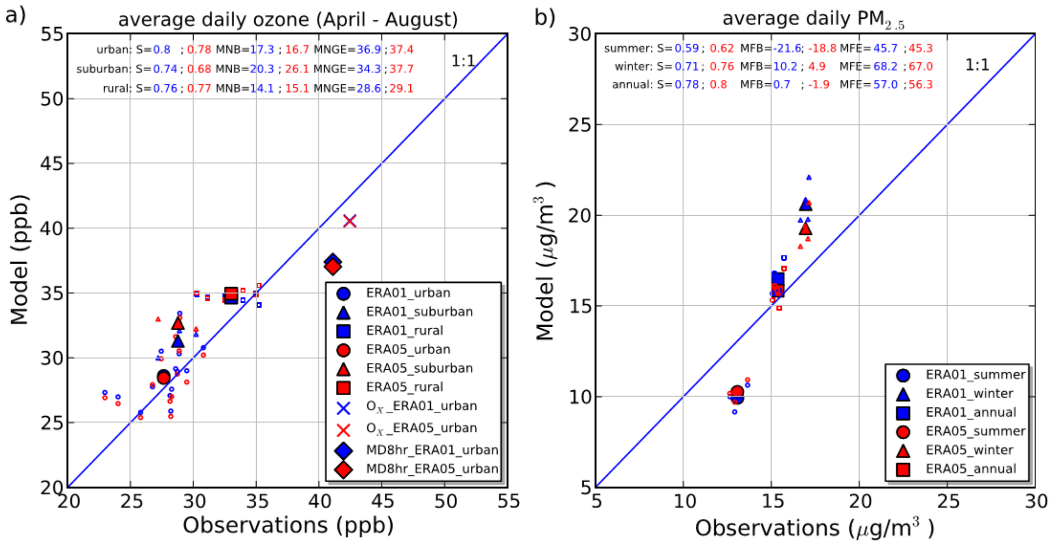

Figure 5. Scatter plots and scores for the sensitivity test on the resolution of meteorology for ozone and PM$_{2.5}$.

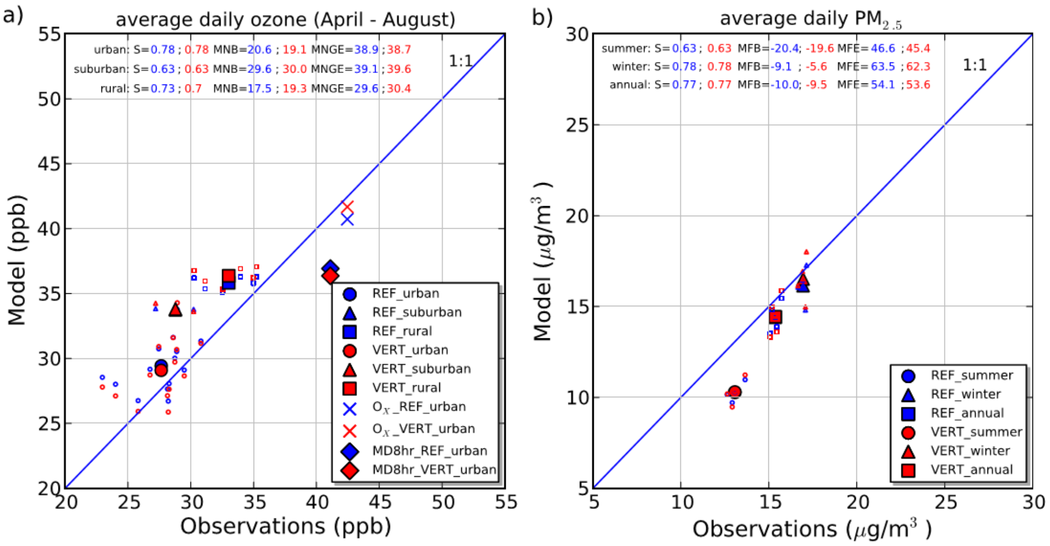

Figure 6. Scatter plots and scores for the sensitivity test on the ~~model~~ CTM's vertical resolution for ozone and $PM_{2.5}$.

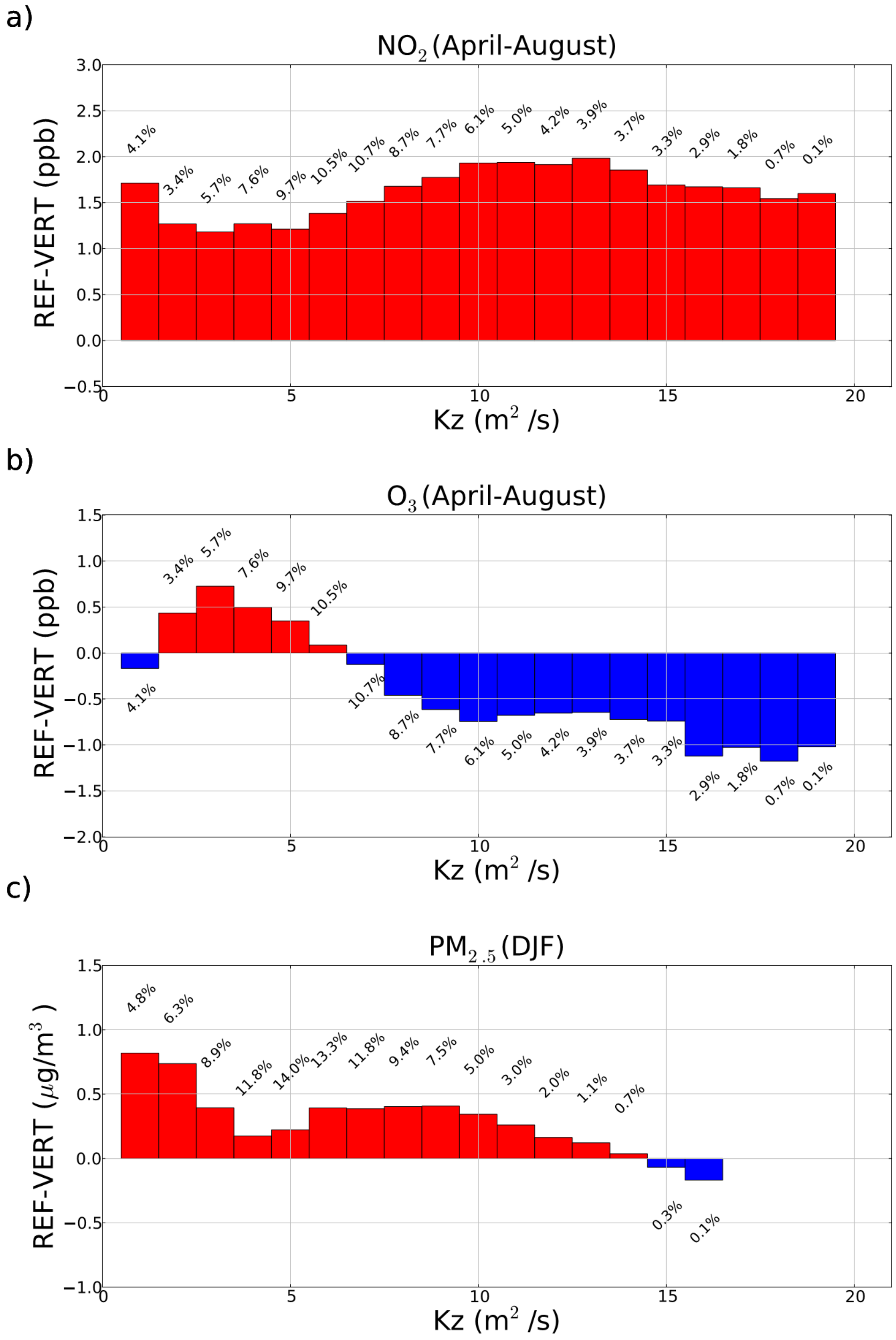

Figure 7. Difference in average daily simulated NO$_2$ (a), ozone (b) and PM$_{2.5}$ (c) concentrations between VERT (12 vertical layers) and REF (8 vertical layers) at urban areas per range of K$_z$ (bins of 1 m^2/s). Positive differences indicate that the refined vertical mesh leads to increased pollutant concentration and vice versa. The occurrence of sensitivity values within each K$_z$ range is also provided.

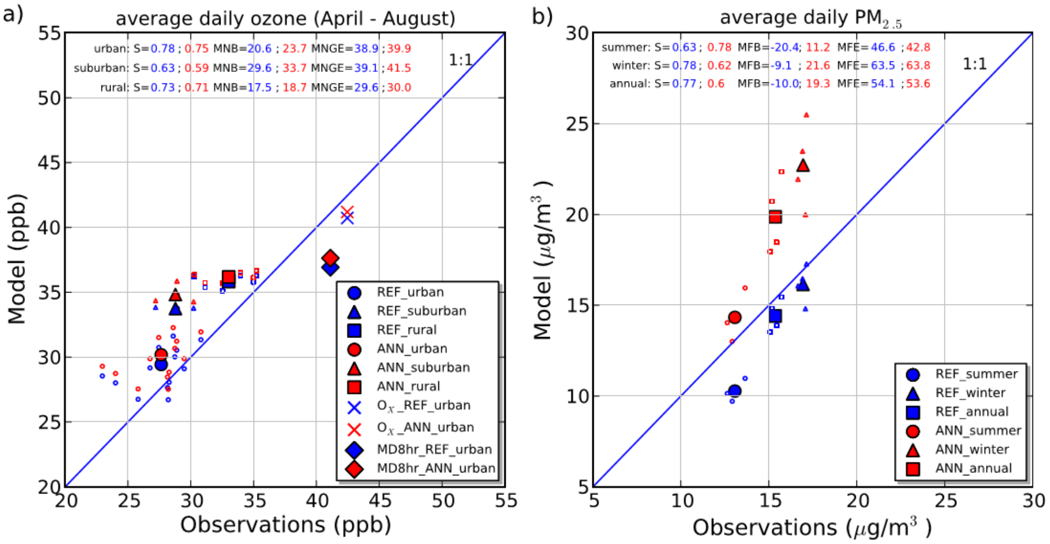

Figure 8. Scatter plots and scores for the sensitivity test on the annual emission totals for ozone and $PM_{2.5}$.

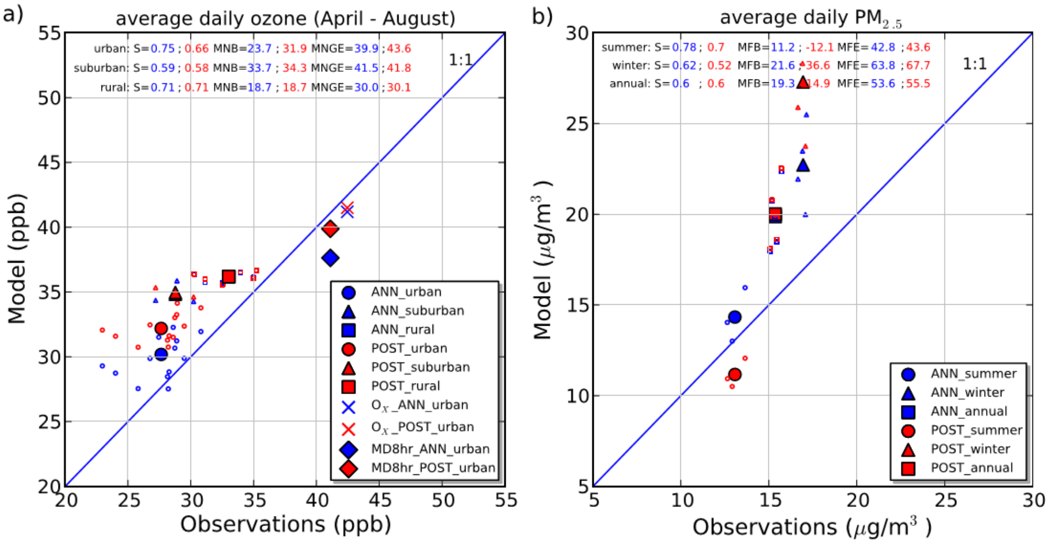

Figure 9. Scatter plots and scores for the sensitivity on the post-processing (temporal analysis and chemical speciation) technique applied on the annual emission totals for ozone and PM$_{2.5}$.

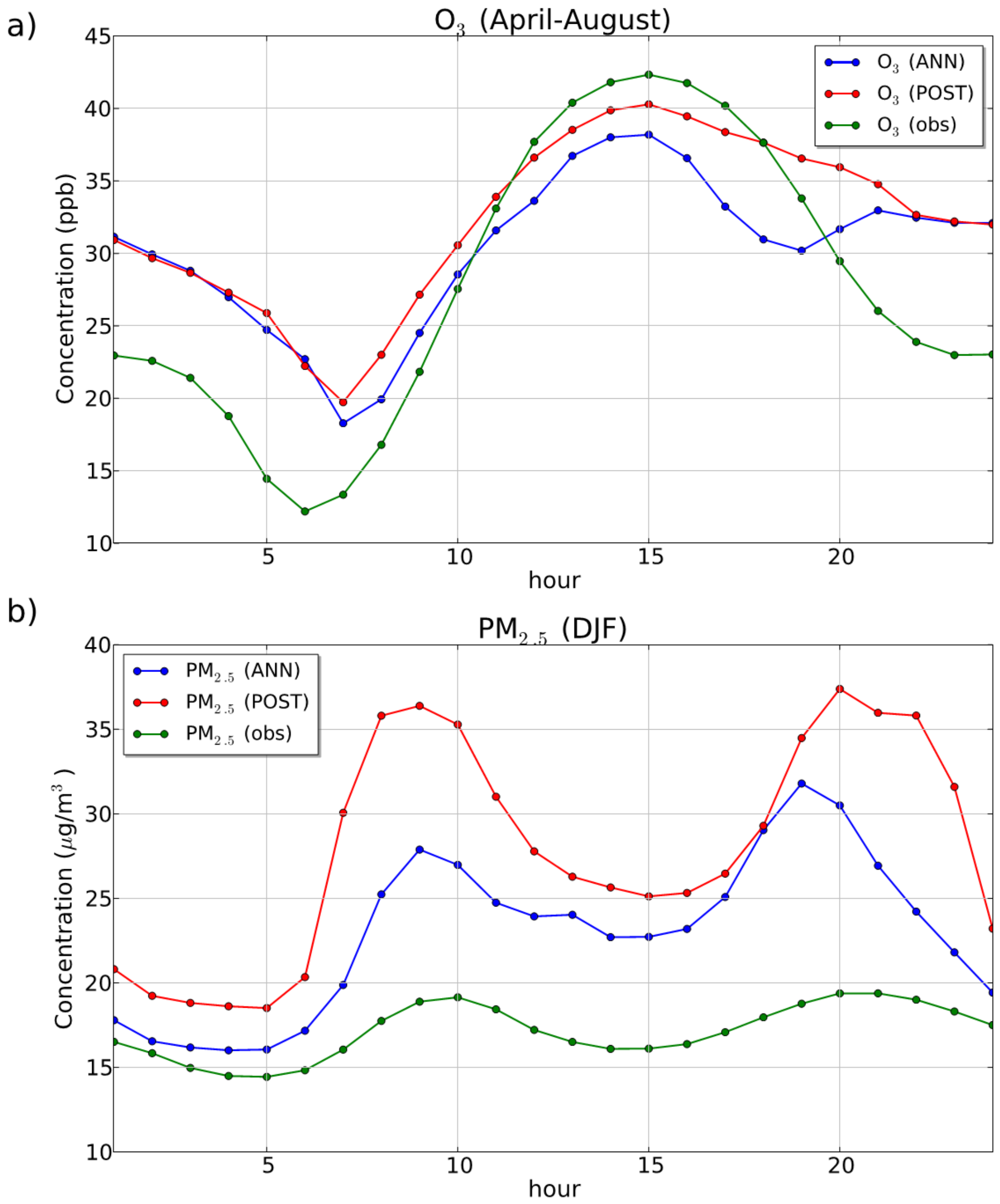

Figure 10. ~~Average~~ Mean diurnal variation of (a) ozone concentrations averaged over the ~~from~~ April-~~to~~ August period ~~(a)~~ and (b) wintertime PM$_{2.5}$ ~~(b)~~ concentrations in the urban area.

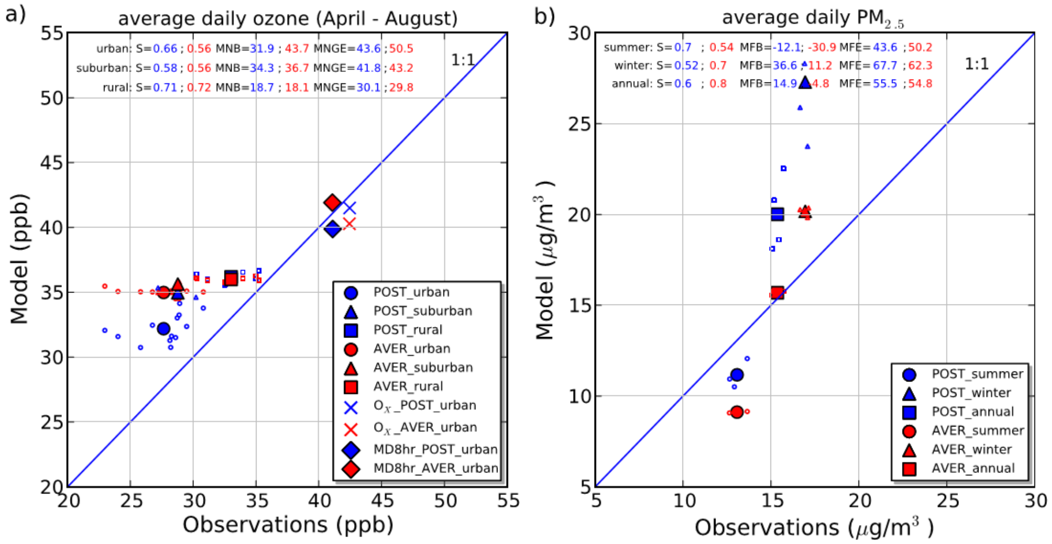

Figure 11. Scatter plots and scores for the sensitivity test on the resolution of the emission inventory for ozone and $PM_{2.5}$.

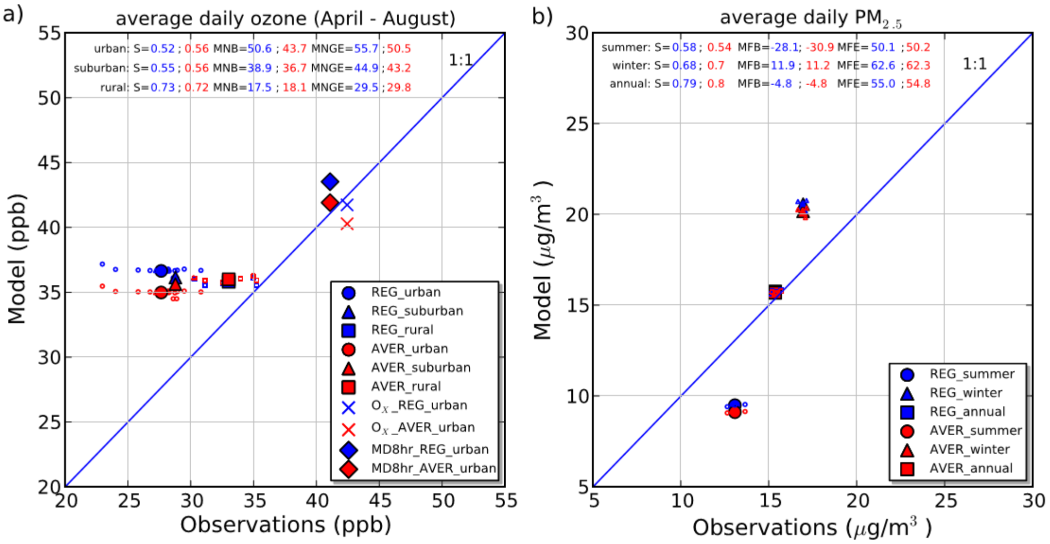

Figure 12. Scatter plots for the sensitivity test on model resolution for ozone and $PM_{2.5}$.

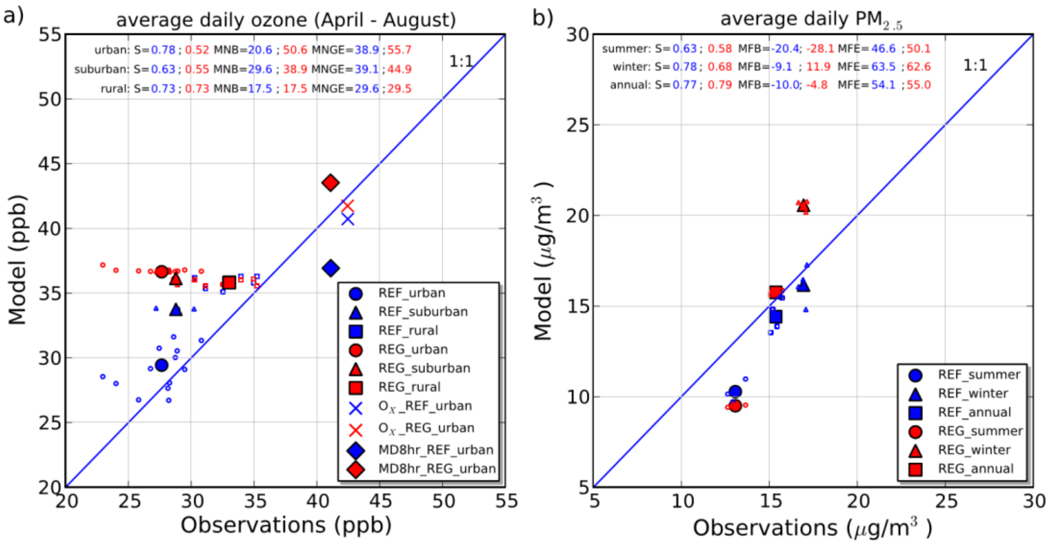

Figure 13. Panel a: Scatter plots of daily average ozone concentrations at urban, suburban and rural stations from the REF and REG simulations. The odd oxygen (O_x) and daily maximum at urban locations is also shown. Panel b: daily average $PM_{2.5}$ concentrations in wintertime (DJF), summertime (JJA) and on annual basis over urban stations.