Second review of ACP-2013-769, Markakis et al.

The authors have addressed my most important concerns. There are a few minor issues that remain that fall under the category of "minor revisions"; a few words here and there in the revised text. I've also tried to clarify a few of my earlier comments. The numbering follows the points in my original review.

2) I am a bit concerned about the use of a 10km resolution meteorological model to drive a 4km resolution air-quality model,…

…Some concerns: higher resolution is not the case in the study of Flagg, D.D., and Taylor suggested by the reviewer where the modelled city represents a complex multi-lake terrain.

Flagg and Taylor investigated the impact of the resolution of surface layer input data on model results for high resolution simulations of an urban area, and found that the model results were quite dependent on that resolution (e.g. root mean square differences of the heat flux between different input resolutions on the order of 20 to 30%). While both of the papers mentioned deal with cities by a coastal environment, the key issue I wanted to point out is that these models show one needs to go to very high resolution in order to capture the urban heat island circulation (Leroyer et al reference) and the effects of the surface layer changes (Flagg and Taylor reference) in an urban meteorological simulation. At the same time, I think the authors' counter-point is valid, that while higher resolution provides a better forecast in a theoretical sense, it may not do so in a practical sense. Flagg and Taylor's work implies that some of the lack of improvement may be the use of surface input data which is at lower resolution than the meteorological model, for example. As well, the downside of resolving plumes at a higher resolution is that small errors in the wind direction can result in decreases in correlation coefficient scores at higher resolution (since the lower resolution model plumes are spread out over a larger region, an error in wind direction will have less of an impact on the comparison to observations). I think the authors main arguments, once explained in the text, that they tried higher resolution meteorology in earlier work and found no improvement in the results, and that higher resolution had some additional computational overhead (though surely not as much as the air-quality model) are valid. To me, the first point suggests that more work is needed on the high resolution meteorological model - but that is beyond the scope of the current work. Discussing the issue is sufficient, here.

3) Given the relatively small size of the meteorological and air-quality model domains, more description is needed for the downscaling and the potential impact of boundary conditions

…WRF simulations were carried out on a 10 km resolution grid of 90x85 cells , (i.e. 900km x 850km) which is not a "very small" domain in our opinion compared to the size of the Ile-de-France region (156km x 128km).

On the meteorology side: yes, 900x850 km is bigger than 156 x 128 km, but what about the boundary conditions used for the meteorological model (global or regional analysis)? My question there is "to what extent are the local model results affected by the driving model boundary conditions?" For the air-quality model, I was thinking of studies such as the HTAP experiments (papers by Fiore, Dentener) where they show that wintertime O3 predictions within one continent are significantly impacted by emissions changes within another continent; the latter changing the downwind continent's O3 through advection of ozone. The sentence of the authors "Having performed the simulations..." would be better stated, "While boundary conditions may impact local scale model predictions, we focus here on the impact of local emissions through the use of a common set of boundary conditions, to ensure that the differences arise from the sensitivity to local emissions."

For the 50km resolution simulations, were the same emission data used as for the 4km resolution simulations?

No, emission data between the regional and local scale simulations are not the same.

Again, see my slight rewording suggestion above.

What boundary conditions were used for the outer 50km simulation, and where did they originate (if these were in the global coupled runs, was the model speciation the same or were there issues with matching them)?

The matching between LMDz-OR-INCA and CHIMERE species…

The chemical table surprises me - do neither of the models include biogenic hydrocarbons (isoprene, monoterpenes)?

Aside from that, what I think was needed at that point in the manuscript was a one sentence reminder to the reader that the 50km simulation boundary conditions come from a larger scale model simulation, with a slightly different chemical speciation.

5) Some aspects of the REF versus MIT scenarios and the relationship…

We believe that there might be a confusion regarding this issue…

This worked better. I wonder if this would be clarified further with a table with three columns going from left to right Global, Regional, Local and rows describing the different runs at each scale. It does help to have that change in the text, though.

Also, Figure 4 suggest that the relative impact…

The purpose of the figure is mainly to compare each future scenario…

My point here is that the dynamic range between maximum and minimum O3 in the rural area just outside of the IdF and within the IdF changed between the simulations; which I think is potentially interesting to mention - hence my suggestion. A few words of explanation of why the difference between rural and urban values has changed between the simulations (as opposed to focusing on the IdF in the core) would be useful.

The gradient between urban and rural O3 has greatly increased in the MIT scenario and this is worth pointing out,…

If my understanding is correct this is not true actually. O3 in the rural areas deceases much more in MIT than in REF….

Yes, but why is this the case? What in the scenarios has caused this change? My point here is that the range of maximum to minimum O3 across the grid has changed - its that difference that I find interesting - why have the differences between rural and urban O3 changed between the simulation. Asking it a slightly different way: why has the rural O3 in MIT decreased relative to the urban region much more than REF? Why has the difference between urban and rural O3 changed between the simulations? What

I was hoping for here is a sentence or two of explanation linking back to the emission scenarios employed and/or the boundary conditions.

Page 101, line 5, a comment: Actually ..

We thank the reviewer for this suggestion…

Agreed. This raises an interesting question as to whether the variability becomes larger at local scales, hence requiring a longer averaging period. In Kelly et al, we found 7+ years seemed to get convergence. Perhaps the higher variability associated with higher resolution models requires a longer time averaging period? Something for future work, perhaps.

Page 105: is the high bias of wind speed improved when WRF is run at higher resolution for urban regions? Given the LeRoyer et al and Flagg papers referenced above, they probably would be. See earlier comment on the resolution of the meteorological model simulations carried out.

Yes the 10km meteorology is able to resolve better urban scale wind speeds during winter but it does have the same performance during summer

I was referring to resolutions higher than 10km here – my question is: is there any evidence from other work that the wind speed bias improves when going to higher resolution? Perhaps in the reference quoted by the authors above?

Page 106, line 21: I don't follow the reasoning that short term meteorology would fail to result in 95th percentile peaks being simulated...

This is true but episodes…

This to me suggests more a problem with the accuracy of the meteorological model rather than the time span of the simulations. The terminology used was a bit imprecise in that it allows the reader to conclude either that the time span of the simulation (short term simulation) or meteorological model is the issue here. I think that this would be clarified by the authors porting the description in the response to the reviewer into the text.