Editors comment paper acp-2015-118; Use of North American and European air quality networks to evaluate global chemistry-climate modeling of surface ozone by Schnell et al.

Dear authors, co-authors,

I finally found the time to carefully check your response to comments raised by the reviewers and the revised version of the manuscript having tackled those comments. This brings me to the decision to accept your paper for publication in ACP after giving you the opportunity to handle a number of last minor editors comments. These comments are mainly addressing the issue of what controls surface ozone levels, especially its temporal variability. I already raised this point in previous communications and have seen this comment is partly addressed but still deem that the analysis is relying to a large extent on the assumption that surface ozone concentration changes are dominated by the chemistry. Under the polluted conditions, the chemical tendency might indeed be large, in contrast to pristine conditions such as found over the tropical forest but according to my knowledge, also for the polluted conditions the combined role of surface deposition and boundary layer dynamics can have a profound impact on the diurnal and seasonal cycles in O_3. I am aware that it would be best to support this observation by showing some analysis of the O_3 process tendency for the polluted conditions (see one of the comments below). If you can make a compelling point that O_3 concentration temporal variability for the polluted conditions your study focusses on, you could consider to include this somewhere in the introduction also justifying the focus of the analysis on the role of chemistry in explaining this temporal variability. Below you can find some more detailed comments that especially address this issue.

Abstract: Having read the abstract I somehow miss at the end a strong concluding statement. You now end with still stating a result from the model evaluation but think that especially your abstract would benefit from a closing statement that makes a clear point what can be concluded from the presented model evaluation.

Introduction; reading over again the statement that explains what controls surface ozone levels I was surprised to see that the term deposition was missing there. I recall that I mentioned this in my response on the initial discussions on this paper and that you also then responded indicating that you indeed were also recognizing that this is an additional relevant process for surface ozone. I see that later on in the introduction you mention the importance of surface deposition and land use changes but think that you should already include this in the listing of the important features that determine surface ozone.

Description of the models, 2.2: although most readers should be familiar with the notations, to deal with acronym slang I suggest to properly introduce the terms CCMs and CTMs

Results: "**The shape of the diurnal cycle of O3 is driven primarily by sunlight, meteorology (e.g., temperature and variations in boundary layer mixing), and the daily cycle of precursor emissions**", also here add the deposition term.

But then does this statement "**The hour of the maximum phase h occurs when**

these factors align, usually in midafternoon" still holds?

"The underestimate of the summertime diurnal amplitude H by most ACCMIP models suggests that they either underestimate net daytime production or have too little nighttime loss of O3." I would add here that they might underestimate entrainment of free troposphere air masses or have too little nighttime loss of O_3 associated with chemical destruction or surface deposition.
You discuss hereafter that the representation of the nocturnal inversion layer (I would rather call it inversion layer than PBL) is the reason for the misrepresentation of the early morning increase in O_3. However, the growth of the PBL in the early morning is one of the essential processes in 1) mixing in residual and free troposphere air masses enhanced in O_3 and this growth of the PBL also determines the efficiency of mixing of the emitted species with the depositing species and, because of that, to some extent the efficiency of the chemical production of the O_3. This could be analyzed by comparing the turbulent transport tendencies versus the chemical tendencies, diagnostics that are generally unfortunately not available in global model simulations.
This sentence reads a little weird due to the combination of all abbreviations and propose the next small addition: "**This reduces daytime production and could partly explain why the models G and H consistently underestimate *H* more than others, however A overestimates *H*.**"

Regarding the statement: "**The boundary layer schemes may be responsible for these underestimates, however, Menut et al., (2013) notes that at least for one model, increasing its vertical resolution results in very small surface O3 changes**", it suggests that only increasing the vertical resolution might be the way to more properly resolve the role of boundary layer dynamics and turbulent transport in properly resolving the diurnal cycles in O_3. However, if the models would all fail on properly simulating the surface energy partitioning, which might be likely with the models e.g., not considering the urban tile/surrounding areas differences in energy balance, then increasing vertical resolution would definitely not help in improving the models performance on PBL dynamics.

Page 9, lines 14-19:" **The lack of hourly variation of emissions may account for the overall underestimates of H by the ACCMIP models, since NO emissions can be lost heterogeneously at night, less effectively than those during the morning and afternoon peaks in traffic. In addition, if the early morning peak in transport NOx emission was included, the modeled morning rise in O3 would most likely be augmented, thus yielding larger values of H.**"

Reading through this statement I fully appreciate the observation about the potential importance of missing temporal variability in the emissions that might be relevant for explaining some of the issues on the representation of the diurnal cycle in O_3. But also recognizing the fact that you deal with large-scale models that do not generally distinguish the urban tile and simply apply grid average emissions, you would not consider the representation of diurnal exchange processes which combines the role of anthropogenic and biogenic emissions but also dry deposition as a function of the large contrasts in nocturnal and daytime turbulent exchange. Those emissions are now generally assumed to be directly into the models surface layer (or above when emission height profiles are considered) whereas most emissions, with the suppressed nocturnal and early morning mixing conditions might happen in the "urban" of

vegetation canopy (BVOCs, NOx) and where consideration of such subtle features of the actual location of emissions and role of turbulent transport might result in very different diurnal cycles in the effective emissions into the surface layer. Another point of discussion is if the emissions are included in the form of NO or NO_2?

Page 9; lines 28-29: "..and which will be done/conducted after submission of this manuscript."

Page 10, "In northern mid-latitudes, processes that drive the shape of the annual cycle are similar to those of the diurnal cycle (i.e., sunlight, temperature, and precursor emissions)". Again making my point about dry deposition, what about the role of seasonal cycles in dry deposition due to large differences in biomass and stomatal uptake?

Page 11, line 5: I would suggest "The amplitude M is controlled by both meteorology and photochemistry and dry deposition"

Having read section 3.2 I realize that an essential aspect on the analysis of the 24 hour versus the MDA8 values might be the actual reference heights that are used for the models and the observations. I guess that you use the model simulated O_3 concentrations at the middle of the surface layer, which is let's typically about 30-40m height. What is the measurement height of most of the observations?? If this is much closer to the surface, you might get large differences between the measurements and the models for stable nocturnal conditions. So, this difference would especially impact the 24 hour comparison whereas it would not largely affect the comparison of the daytime only data. This is pointing again at the challenge that also for analysis of these large-scale model output on atmospheric chemistry we need to integrate some of the knowledge on nocturnal exchange phenomena.

Page 15: "**but above the 60th percentile (where UCI and observed**"

Page 19: "**It remains unclear whether such errors result from chemical or physical processes.**". If you would indeed agree on my statements on the role of surface deposition and boundary layer dynamics than I would suggest to modify this statement (but also previous ones) on the role chemical versus non-chemical processes in O_3 temporal variability to chemical and biogeophysical processes (to consider the role of the biosphere and PBL dynamics).

Section 4.1: on this discussion on what kind of output would be required to more optimally diagnose the temporal variability, having worked myself with global chemistry and climate models I used generally the output at an output frequency on the order of 7/13/23 hours or so to sample at least in a month each time of the day to construct from this a monthly mean diurnal cycle. Also as mentioned, to differentiate between the role of chemistry versus surface deposition and vertical and horizontal transport, the output of process tendencies would be optimal but am aware that this is simply too much to get from these global model simulations. You could however consider to get this extra diagnostics for some of the particular locations where such more detailed analysis could be insightful, e.g. getting the process tendencies for strong contrasting regions of the tropospheric O_3 budget.