Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2019-990-RC2, 2020

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Characterizing sources of high surface ozone events in the southwestern U.S. with intensive field measurements and two global models" *by* Li Zhang et al.

Anonymous Referee #2

Received and published: 7 February 2020

This paper leverages intensive field measurements from the Fires, Asian, and Stratospheric Transport-Las Vegas Ozone Study (FAST-LVOS) and two global chemistry models to attribute high ozone events in the Las Vegas area to various causes: wildfires, stratosphere to troposphere ozone intrusions and transport from Asia. For many of the events the disagreement between the two models is notable.

Overall, I found this to be an interesting, well-written paper with well-crafted figures. After the authors address a few points I would recommend publication.

Major Comments:

I. It is unclear how the different events are attributed to different sources. A more clear

discussion of attribution is necessary.

(i) As the author's state: "Identifying the primary source of the high-O3 events solely based on observations is challenging; additional insights from models thus needed as we demonstrate below." Thus, how are the events attributed in Table 1? Is the measured data in section 4.1 used to support the modeled attribution? Or is it the primary attribution mechanism? Are both models used to attribute a particular event or just one? Is the event attribution through the preponderance of evidence?

(ii) On a more philosophical note it seems the authors often make good qualitative arguments for a particular type of event but is this sufficient to attribute an event to a particular cause? In particular, as air from various sources is mixes together how does one determine the cause of an ozone exceedance? Quantitatively, doesn't one need to know the average ozone contribution from a particular source and if ozone from that source is elevated sufficiently from its average one can attribute the ozone exceedance to that source (even if it is not the dominant source)? Or does one require a particular source to be the dominant source to attribute an event to it? At any rate more detail should be given as to what it means to attribute an event to a source.

II. The difference between the models in Figure 6 is deeply disturbing to me. While the authors concentrate on the differences in the stratospheric ozone intrusion, the figures overall are very different, not only in their ozone but in the isentropes. Is this a resolution problem, an interpolation problem, a meteorological analysis problem or possibly a result of differences in the advection algorithm? This seems quite important to determine as one of the main differences between the models seem to be in their handling of stratospheric intrusions. It seems as a minimum the authors could look at: (i) potential vorticity and potential temperature surfaces in the native resolution of the two meteorological datasets. Are these the same or different? (ii) Then examine potential vorticity and potential temperature in the resolution of the two model grids. Does changing the model grid do something to the fields? (iii) Finally they could examine the ozone differences between the figures. Is the ozone similar in the stratosphere in general in the two models which would point to greater numerical diffusion in GEOS-CHEM diluting the stratospheric intrusion? Or perhaps GEOS-CHEM has less ozone in the stratosphere. The paper seems to imply that the simplified stratospheric chemistry and dynamics in GEOS-CHEM is the reason for the discrepancy between the two models. What evidence do the authors have for this assertion?

III. When comparing differences between the AM4 and GEOS-CHEM model it would be useful to know the extent to which these differences might be due to differences in emissions. At a minimum the authors should discuss some of the emission differences, and the extent to which these might contribute to the difference in the USB ozone. Ideally these type of studies should be done with the same emissions. However, the authors should take any emission differences into account in their analysis, or at the least discuss that these may be a source of uncertainty in analyzing the differences between the models.

IV. One of the important aspects of this study seems to be in the attribution of exceedances. I think the authors should quantitatively assess the skill of the models in diagnosing exceedances (perhaps over a longer period of time than the FAST-LVOS timeframe itself, if possible). There are many measures of this skill in the literature. What percent of time when an exceedance is measured do each of the models predict the exceedance? And how often do the models get a false positive (predict an exceedance when one doesn't occur)?

V. I feel the conclusions could be made stronger. The authors give for the most part a detailed comparison between the two models. I think the paper would be strengthened if the authors stepped back a little.

(i) First, it would be interesting to discuss model skill in simulating the exceptional events as discussed in the first paragraph in the paper. Giving the difference between the two model simulations to what extent are we confident that they can screen exceptional events? What is the skill of the models in assessing extreme events (see

comment IV).

(ii) Second, in the case of an exceptional event, to what extent can these models be used to attribute the event to a particular cause?

(iii) The authors claim that much of the pattern of USB in AM4 is due to STT and the ability to simulate STT is an important difference between the models. This seems like an important conclusion, but how strong is the evidence? It would be good if the authors would summarize the reasons they conclude this. Do the differences in USB coincide with locations where the stratospheric tracer is high? Some more analysis might be beneficial to really make this point.

(iv) While the authors point out some differences in USB it would be nice to quantify the uncertainty here. I think a difference map in the USB between the two models would be very helpful. Do we know USB within 10% or 20% (at least based on these two models) and how important is this for policy considerations.

Minor Comments:

1. P2, l47 “contribute”. Should this be contribute episodically?

2. P2, l58 “independent”. This seems a bit strong. As has been shown in climate models (e.g., Knutti et al., 2013) models are really not independent from each other due to the sharing of information and algorithms across groups. For example the two models in this paper share the MEGAN scheme, but probably also share other aspects. Thus, I would delete the word “independent” here and in other locations.

3. Fig. 6, Please make the vertical and horizontal scales identical so these figures are easier to compare.

4. P6, l46 Please give the frequency of measurements used here and elsewhere.

5. P7, l188 What is GEOS-FP meteorology?

6. P8, l229 What is the standard representation of lightning?

7. P9, l235 Are the emissions the same in AM4 and its predecessor? If not to what extent is the comparison between them simply a matter of the different emissions. At the minimum the authors should mention these comparisons use different emissions and the extent to which these differences can explain the differences between the models.

8. P12, l319. PVU is not a unit. Please give the mks units for a PVU.

9. P14, l394. The authors suggest excessive lightning NOX in GEOS-chem causes excessive ozone. They cite a number of older papers to make their point. What is the evidence in this study for excessive lightning NOX? For example, P18 l496 the authors state the overestimate is likely due to lightning NOX. On P20 l569 the author categorically state it is the abundance of lightning NOx that results in higher background ozone in GEOS-chem. Without more analysis it seems lightning NOX is a possible explanation. However, if they authors claim this is the likely explanation they need to give some more evidence.

10. P14, section 4.3: The authors made a number of sensitivity simulations with respect to the simulation of fire plumes. I did not get a sense as to which of these sensitivities improved or degraded the simulation. Please give some overall conclusions.

11. P15, l14 “dominant source”. Could you clarify? If the local emissions are 20-30 ppb and simulated emissions are over 60 ppb, why are local emissions the dominant source?

12. P17 l481 How were the STT events diagnosed?

13. P17 l486 “underestimates the magnitude of STT”. The authors show that GEOS-CHEM underestimates the ozone concentrations in stratospheric folds, at least the ones they examined. First this sentence needs to be qualified as to where and when. Secondly while GEOS-CHEM may underestimate the ozone in the stratospheric intrusions this does not mean it underestimates STT (the model might simply be excessively

diffusive while still simulating the same exchange).

14. P18 l519, and p19 l20: “Many of the standard O3 events. . .” Could you quantify this? From the figure it appears to be less than half the events?

15. P19 l549 “contributing to $\sim$30ppbv to surface ozone”. I believe this is a modeling result. Please state this.

16. P20 l555 “wildfire event”. This would be a good place to summarize whether any of the sensitivity tests resulted in a better capture of the wildfire event.

17. P18, l500 “likely reflect”: what are the arguments for this? It might be interesting to show a difference map for USB, getting at the uncertainty in USB between two state-of-the-art models.

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-990, 2019.