

## Author Additional Comments

Referee comments in black and author comments/actions in red.

Further amendments to the manuscript have been made in accordance with an internal review procedure within the Canadian Centre for Climate Modelling and Analysis department of Environment Canada, of which is the host institution of co-author *David A. Plummer*. These amendments are listed below:

“P1, L32: With the model biased high in the lower strat and underproducing photochemical O<sub>3</sub> in the troposphere, the conclusion that stratospheric O<sub>3</sub> intrusions are larger than previously thought is cast in doubt, since that may just be because of the model biases.

Can more explanation be added here to explain why the authors still feel confident in that conclusion? For example, was a correction applied to the model before coming to that conclusion?”

Changed “larger than previously thought” to “significant”. Although we can be confident that the real influence of the stratosphere is larger than that found by Lamarque et al. (1999), which is the main study we base this assertion on, we have to appreciate it is a twenty year old study that was based on a much simpler model. We therefore include new references throughout the manuscript as suggested by the reviewers and tone down statements such as that above. We see no reason to alter the estimated influence exceeding 50 % in the wintertime high latitudes however, as both model show this (it was only CMAM which had the high (low) bias in the lower stratosphere (troposphere) and if anything the inverse was found for EMAC). However, it is an important point and does suggest there is some uncertainty still in our understanding of the stratospheric influence.

“P2, L25: Fiore et al (2002, JGR) and references therein could be added as older (more seminal) references of the increase in background tropospheric O<sub>3</sub>.

Fiore, A. M., D. J. Jacob, I. Bey, R. M. Yantosca, B. D. Field, A. C. Fusco, and J. G. Wilkinson (2002), Background ozone over the United States in summer: Origin, trend, and contribution to pollution episodes, *J. Geophys. Res.*, 107(D15), ACH 11–1–ACH 11-25, doi:10.1029/2001JD000982.”

Done.

Here's another reference for transpacific transport of O<sub>3</sub>:

“P2, L28-29: Zhang, L., et al. (2008), Transpacific transport of ozone pollution and the effect of recent Asian emission increases on air quality in North America: An integrated analysis using satellite, aircraft, ozonesonde, and surface observations, *Atmos. Chem. Phys.*”

Reference added.

“P4, L7: Lines 5-11 read as a little dismissive of the many excellent tropospheric O<sub>3</sub> measurements from satellites such as TES, OMI/Trop-OMI, TOMS, MLS, etc - many of which have \*long\* term (> decade), global datasets, and are well validated by ground-based remote sensing, in situ, and inter-satellite comparisons.

If just trying to motivate the use of models to fill in the gaps in measurements, one may mention that the satellite measurements are usually limited to just a couple of overpass times per day at each location, and often have large errors in retrieved O<sub>3</sub> (but are those uncertainties more than those from the mode??).

However, the second point about assessing and quantifying the causes and processes of trop O<sub>3</sub> is good”.

These lines have been rewritten to sound less dismissive, with greater acknowledgement of the value such satellite datasets provide and their contribution to our understanding of tropospheric ozone. The paragraph as a whole has also been modified to highlight the value of CCMs in terms of providing insights and interpretation (mechanisms, feedbacks and quantification of the stratospheric influence) that cannot be inferred from observations alone, as opposed to the filling in of gaps in the measurements, which is not such a valid point for such a reason you mention.

“P4, L7-8: add references”

Reference added.

“P4, L16-17: do you have a reference or other proof (e.g., your own calculations) of emissions being the largest source of model uncertainty?”

No, but we add in a reference which supports this claim.

“P4, L27: what does F stand for? Is O<sub>3</sub>F the total O<sub>3</sub>? ...Ok, coming back to this, I think I understand what O<sub>3</sub>F is, and "F" may stand for "fraction", perhaps?

Ideally O<sub>3</sub>S and O<sub>3</sub>F would be changed to something clearer like: O<sub>3</sub>\_strat and O<sub>3</sub>\_%strat (use subscript instead of "\_")

The F indeed refers to fraction. This has been made clearer in the manuscript and indeed should have been defined as such here (the first instance it is mentioned).

“P6, L9: EMAC also extends to 0.01 hPa, but above you haven't given the altitude. Either move to first mention of 0.01 hPa, or remove the km altitude altogether, since that's highly variable depending on latitude”.

Removed 95 km reference.

“P13, Fig. 3: The RSD from the ozonesondes looks to be much lower than that from OMI (and that from the models too), but this is not discussed in Section 3.2.”

The scale range has been revised to show more structure in the RSD plots, however the ozoneosnde RSD is significantly lower than OMI in particular with few exceptions. This will require further investigation.

“P14, L25-29: I think the ppbv goes with Fig S1, but this method of using brackets is confusing b/c there is also left, middle, and right in brackets in the same sentence. I suggest describing Fig 4 in one sentence, and then adding a second sentence, saying that Fig S1 shows the same thing, but absolute differences of VMRs instead of percent differences.”

Correct but agreed that the use of brackets is excessive in this sentence so have revised this and split into two sentences in accordance with your suggestion. Many thanks.

"P15, L31: change "should be" to "are""

Done.

"P17, L23: It's not clear to me what O3S is. From the wording, it sounds like it is just O3. But from the equation given in the caption, O3S seems to be the tagged stratospheric O3. If the latter, this sentence needs to be changed to say "Seasonal composites of the monthly mean, zonal-mean vertical distribution of stratosphere-originating ozone (O3S)" or "...tagged-stratospheric ozone (O3S)".

This comes back to my earlier comment that O3F and O3S are not described clearly enough, and was confusing to me (and possibly other readers). Ideally, "O3F" and "O3S" would be replaced by "O3\_%strat" and "O3\_strat"

This has been clarified and sentence broken up to avoid excessive use of brackets and enhance readability.

"P17, L24: How is Fig. S2 different from Fig 5? I don't see any corresponding bracketed statement that goes with S2 rather than Fig 5...

Please add the explanation in these brackets, and make it clearer in the caption for Fig S2. For example, the caption to Fig S2 could be "Figure S2 - same as Fig 5, but for ...[whatever is different]."

...Ok, it took me a while, but now I see that the above sentence referring to O3 corresponds to Fig 5, and the bracketed "(O3S concentration)" corresponds to Fig S2. This method of using brackets to try to save wording makes it unnecessary confusing for the reader, and could be written just as efficiently as follows:

"Seasonal composites of [...] ozone concentrations (is it concentrations or VMR? clearly state one without using brackets) from 1000-80 hPa are shown in Figure 5 for EMAC (a), .... together with ... . The same is shown for the stratospheric-tagged O3 (O3S) in Figure S2."

This has been remedied following the action taken immediately above.

P17, L33: There are a lot of these bracketed inverse statements that I think should just be restated for clarity. For example, here it doesn't cost too many extra words to say:

"with the former clearly a greater influence near the surface, and the latter in the upper troposphere."

I would reword most, if not all, cases like this in the paper to improve clarity when it can be done so efficiently."

Done. We appreciate your point and make such revisions where necessary elsewhere to enhance readability and avoid reader confusion.

“P19, L25: Although this is a straightforward sentence using the bracket method, rewording to "when O3\_%strat reaches a maximum in winter and minimum in summer." only adds one word ("and") and is clearer to the reader.”

Done.

“P20, Fig. 6: what is the gray? Please mention what it is in the caption.”

The grey shaded regions represent where surface pressure values are lower than the plotted pressure level (i.e. where each pressure level would be below ground). This is now indicated in the figure caption.

“P23, L1: "in the present study””

Phrase added.

“P24, L11: what is meant by "even lower"? Meaning the models are biased even lower? Or meaning even lower in altitude? This should be clarified b/c if interpreted the first way, then a low-biased tropospheric O3 will give erroneously high biased O3F in the troposphere.

I would remove "even lower" from the sentence.”

We were referring to the influence exceeding 50 % as far down as the lower troposphere and hence we revise the sentence to make this clearer and avoid reader misinterpretation.

“P24, L19: just say "modelled O3 VMRs", remove the word "concentration" b/c VMR is not a concentration.”

Done.

“P24, L19: again I'm confused: does Fig S5 correspond with O3 VMR and Fig 8 correspond to O3 concentration? The brackets are confusing and unnecessary.”

Fig. S5 refers to the seasons DJF/JJA (also brackets at end of sentence). We agree that this method could confuse the reader so have clarified this.

“P24, L21: re-write to clearly state that Fig S6 is for DJF/JJA - if that's the case.”

Again, this has now been clarified.

“P24, L22: (and Fig. S7 for ...)”

Again meaning for DJF/JJA. This is now clearer.

“P24, L23: ditto”

Also for DJF/JJA but for the cross-sections. This is now made more obvious for the reader.

“P24, L29: by what measure? ...The paired t-test p-value threshold should be interpreted with caution, and I suggest the authors add a caveat (e.g., reference to Wasserstein & Lazar paper that I mention in a different comment).”

We now reference this citation and make the reader aware that such stippling needs to be interpreted with caution.

“P25, Fig. 8: A word of warning about interpreting the t-test threshold this way. The American Statistical Association (ASA) has thrown out the idea of using p-value thresholds to confirm or deny the null hypothesis, saying that you have to look at the broader picture and additional data to determine significance. Please see: “The ASA’s Statement on p-Values: Context, Process, and Purpose”, by Wasserstein, R.L., and Lazar, N.A., The American Statistician, 70:2, 129-133, 2016. <https://www.tandfonline.com/doi/full/10.1080/00031305.2016.1154108> ...and consider revising wording in this paper to be less definitive about statistical significance based on the  $p < 0.05$  threshold.”

Thank you for bringing this to my attention. We add in the necessary caveats when discussing the stippled regions in relation to statistical significance.

“P30, L9: explain what’s in Fig S8 in the brackets”

Specified the two different seasons presented in each of Fig. 11 (MAM/SON) and Fig. S8 (DJF/SON) for clarity.

“P31, L31: ....4-6 ppbv over the Northern Hemisphere and 2-6 ppbv over the Southern Hemisphere subtropics...”

Revised for greater clarity.

“P32, L3: too many brackets! Reword to remove as many as possible.”

Most of these brackets have now been removed.

“P33, L6: why would a complex chemistry scheme cause a high bias? Do you mean “inaccuracies in the complex chemistry scheme”?”

You are right, it should not unless inaccuracies exist but I would not know if this is the case. The implication here is that the emission inventories will have an error attached and this might manifest more prominently in EMAC’s modelling of ozone due to the complexity of the chemistry scheme. This would favour a high bias in this model if such inventories are an overestimate of the truth as implied by Hoesly et al., 2018. Sentence has been revised to reflect this.

“P33, L6: but both models used the same emissions, no?”

Correct, this has now been stated as so.

“P33, L16: Is there a paper on OMI’s long-term performance that you can reference?”

We have added a reference that discusses the long-term performance of OMI – Levelt et al. (2018).

“P34, L4: due to ... ?”

Further detail from the Bonisch et al. (2011) reference has been added regarding the cause of reduced transit times in the lower stratosphere.