Reply to Dr. Sun Wong:

We would like to thank Dr. Sun Wong for insightful comments that helped us to improve the manuscript. We have carefully considered each of the comments in our revision. Our responses are provided below inline in italics.

1. The Production P and Loss frequency is parameterized as climatological monthly values for O_3. So the interannual variability seen in the model is mainly driven by the interannual variability in the circulation, as discussed in the text and Fig. 13. When there is interaction between circulation and chemistry, this may not be reflected in the climatological Loss frequency. For example, during stratospheric warming in late winter early spring, warmer temperature results in a warmer and looser polar vortex which is not in favor of the ozone hole depletion. The interannual variability of the depth of ozone hole caused by this reason may not be reflected in the model. The authors may consider discuss this or how to remedy such problem. Similarly, the variability in ozone hole related to QBO caused by changes in zonal-mean flow and wave interaction may also be missing in the model.

Reply:
The reviewer is correct. We admit that our trajectory model cannot account for interactions of circulation and chemistry properly. This is most likely important in winter polar regions, where ozone loss chemistry is temperature-dependent. While our model does a reasonable job at simulating the time-average behavior of polar regions, we cannot simulate the detailed behavior in any particular (cold) year with large chemical ozone losses due to the climatological loss frequency adopted from WACCM. We have included a brief discussion of this in our revised Discussion section. There are no similar circulation-chemistry interactions for CO. Thank the reviewer to bring this up.

2. When comparing satellite data with model data, I am happy to see the authors apply averaging kernels to make decent comparison. However, have the comparison also done with the same sampling? For example, MLS data have quality flags. When the MLS flags a data at a location and time as bad quality so the data are not used in averaging, do you discard the corresponding location and time's data in the model? How may this sampling issue influence your results?

Reply:
We have only included MLS data with good quality flags. We didn't sample our trajectory results at MLS observations, but rather made comparisons with our gridded MLS data. In our analysis, we performed the comparison on either monthly or climatological basis; because of the dense space-time coverage of MLS, the sampling issue will not have large influence to our results.

3. A minor comment is on Fig. 2. It is known that the maximum O3 production is at around the tropical 30 hPa (may migrate with season.) The plot (P-L)/L here shows the

maximum net production is at the tropopause level instead of 10 hPa. Is that because the loss of ozone is also big at 30 hPa, or is that because the effect of dividing by L so that the peak becomes at a lower altitude?

Reply:

The (P-L)/L is the ratio of net chemical changes to loss – it represents the balance between production and loss. At tropopause level, maximum ratio indicates net increases of O_3 due to production in excess of loss; at 30 hPa and above the close-to-zero ratio indicates a balanced production and loss, i.e., loss of ozone at this altitude is comparably large to production. In order to clarify the overall chemical behavior (and in response to another reviewer), we have included a new Fig. 2 including O_3 and CO lifetimes, in addition to (P-L)/L ratios.