Besides those minor text changes proposed by the reviewers, the major changes to the manuscripts are:

Page 5991: changed the title to lower stratosphere, and all other places mentioning it.

Page 5997: Fig. 2 and the text involved have been changed to have chemical lifetime included. *[starting on page 5 line 152 in updated manuscript]*

Page 5997 Line 18: added more discussion of the photochemical steady states of O_x. *[starting on page 6 line 172 in updated manuscript]*

Page 5997 Line 27, and Page 6001 Line 4: we have included discussion of the interaction between chemistry and circulation. *[starting on page 6 line 185 in updated manuscript]*

Page 6004: At the end of the first paragraph in Summary, we added a short discussion of extending the results to upper troposphere by using the ERAi circulation. *[starting on page 5 line 152 in updated manuscript]*

Page 6004-6005: We added more discussion of the different heating rates from reanalyses on the second paragraph. *[starting on page 5 line 152 in updated manuscript]*

Page 6005 Lines 9: we added more discussion of the Ploeger et al., 2012 paper and some other evidences that support our conclusion. *[starting on page 5 line 152 in updated manuscript]*

Page 6006: we added a paragraph discussing the limitation of this model. *[starting on page 5 line 152 in updated manuscript]*

Next are the replies to those reviewers separately.

Reply to Dr. Sun Wong (C434):

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.

Reply to Anonymous Referee #1 (C751):

We thank the Anonymous Referee #1 for the careful reading of the manuscript and the valuable comments. We have carefully considered each of the comments in our revision. Our responses are provided below inline in italics.

Suggestions:
1. Title may want to specifically state "Tropical" UTLS.

Reply:
Besides tropical region, we also showed some results in polar region. So it might be more appropriate to change the title to "... in the lower stratosphere (LS)". Noted that we could have results covering the entire UTLS but due to MERRA negative heating rates at lower levels that prevent parcels ascending to the stratosphere, we cannot initiate parcels at lower region.

2. Line 22, page 5994; insert "the" in front of 370K

Reply: *Done.*

3. Line 8, page 5995: states "re-entered the troposphere"....if you're concerned with the UTLS distributions, don't you need to only ignore parcels that can't impact the UT anymore? So, possible, a lower pressure level would be relevant?

Reply:
Those parcels travel back to 250 hPa and below have little impact on the lower stratosphere, so we just ignored them. An even lower boundary level would have no impact on our results because we mostly concern about parcels at and above 100 hPa.

4. Line 9, page 5995; need a "the" in front of ~2200K

Reply: *Done.*

5. Line 20, page 5995; change "waves" to "wave"

Reply: *Done.*

6. Question on heating rates: Does the reanalysis diabatic rates include some portion related to reanalysis increments? That is, if you did a radiative heating rate calculation with the gas and temperature profiles in the reanalysis, would you get the same diabatic heating rate that is stored in the reanalysis output? What causes the differences in Q

shown in Figure 1? Do the four reanalyses have dramatically different temperature and ozone/ch4/h2o/n2o inputs in their heating calculations?

Reply:
The details of differences in diabatic heating rates among reanalysis data sets are a complex issue, and have recently been discussed by Wright and Fueglistaler (2013). The differences in Q are mostly caused by the differences in the long-wave radiative heating, which is affected by temperature and ozone. For example, ERAi uses zonal-mean monthly mean climatology of ozone, MERRA and CFSR use prognostic ozone simulated by the underlying forecast model, and JRA25 uses daily offline calculations. Ploeger et al. (2012) shows that the ERAi heating rates in the tropical lower stratosphere are significantly larger than radiative heating calculation estimates, but this is just one aspect of the overall uncertainties. Part of our motivation for including trajectory model calculations based on both ERAi and MERRA was to test the sensitivity to heating rates for lower stratospheric transport. We have highlighted this in our revised Discussions.

7. Question on chemistry: When you pick out production and loss rates from WACCM and apply them to the reanalysis based trajectories, are you considering any variation in season, or with temperature? If the WACCM temperatures aren't the same as the reanalysis temperatures, does that introduce error? Just a bit more explanation of how the production and loss rates are applied to the trajectory model would be helpful (Section 2.2).

Reply:
We did consider the seasonal variations of production and loss rates from WACCM. In Page 5996 lines 14-17 we mentioned that we calculated production rates and loss frequencies as a function of latitude, altitude, and climatological months to cover the annual (constant) cycle of chemical behavior. However, these calculations will not accurately handle the situation where chemical losses are linked to meteorological behavior, such as intense ozone losses during unusually cold polar winter stratosphere. We have included a brief discussion of this in the revised Discussion section. We have also included more details as to how we applied the production and loss rates to the trajectory model in the revised manuscript.

8. Line 2, page 5998; what kind of MLS climatology is used? (annual, monthly averaged, daily averaged, averaged over what time period?)

Reply:
Thanks for reminding us. We used the climatology of O_3 and CO averaged over August 2004 to December 2012 to set the initial chemical abundances when parcels are initialized. This means that parcels are initialized with a constant annual cycle of O_3 and CO. We have addressed this carefully in the revised manuscript.

9. Line 24, page 5998; change to "differences from MLS retrievals"

Reply: Done.

10. Section 3.1, page 5999: Define "reasonable agreement". I would conclude a spread from looking at Figure 4.

Reply:
In Fig. 4a we see reasonably good agreement between TRAJ_MER and MLS. Noted that the x-axis is in log scale to highlight the differences. Given very low concentrations of O_3 at this altitude (100-50 hPa), the discrepancy between TRAJ_MER and MLS is rather small so we think they are in reasonable agreement.

11. Can you compare WACCM heating rates to those in the reanalyses? There seems to be a significant difference between the WACCM and trajectory runs at 68 hPa. Is that a consequence of a difference between upwelling computed in WACCM versus that in the reanalyses?

Reply:
We could compare the WACCM heating rates to reanalysis, but this would not be especially insightful. Three-dimensional transport within WACCM is based on a finite-volume semi-Lagrangian scheme, and is not directly related to radiative heating rates in the model.

12. Line 7-9, pg 6005; why do you suspect that the ERAi upwelling may be too strong? Please add a sentence summarizing Ploeger's conclusions.

Reply:
Ploeger et al. [2012] performed a radiative calculation based on CHAMP temperature and HALOE ozone and water vapor data, which shows that ERAi heating rates is about 40% too high. This conclusion is consistent with Schoeberl et al. [2012], which shows that trajectory simulations of water vapor tape recorder signal based on ERAi heating rates is at least ~30% too fast compared with MLS observations. Thanks for the reviewer's reminding. Now we have clarified this in the revised manuscript.

13. Line 26, pg 6005; change discussions to discussion

Reply: *Done.*

14. Summary and Conclusions: Could you add a few sentences as to how the trajectory approach presented here provides more information than analyzing an SD WACCM run?

Reply:
In the last version we have mentioned some potential strengths of using trajectory modeled chemical species in the last paragraph. Now we modified this part to make it more obvious to the reader. Thanks for pointing it out.

[**References**]

Ploeger, F., Konopka, P., Müller, R., Fueglistaler, S., Schmidt, T., Manners, J. C., Grooß, J.-U., Günther, G., Forster, P. M., and Riese, M.: Horizontal transport affecting trace gas seasonality in the Tropical Tropopause Layer (TTL), J. Geophys. Res., 117, D09303, doi:10.1029/2011JD017267, 2012.

Wright, J. S. and Fueglistaler, S.: Large differences in reanalyses of diabatic heating in the tropical upper troposphere and lower stratosphere, Atmos. Chem. Phys., 13, 9565–9576, doi:10.5194/acp-13-9565-2013, 2013.

Reply to Anonymous Referee #2 (C1128):

We thank the Anonymous Referee #2 for those valuable comments. We have carefully considered each of the comments in our revision. Our responses are provided below inline in italics.

Major Points

1. The chemical lifetime of CO in the model domain ranges from 1-3 months at the lower boundary to orders of a day or less at the top of the domain. The chemical lifetime of ozone at the bottom of the model domain is rather large, but it also decreases in the tropics to about 10 days at 10 hPa and less than a day at the top of the domain. The chemical loss and production rates are imprinted from the WACCM simulation. That means that the model results relax to the WACCM results with the time constant given by the chemical lifetime. Therefore it is clear why the results, especially in the upper model domain are virtually identical. In order to understand this problem and to interpret the model results, it would be necessary to show the chemical lifetimes e.g. similar as given in figure 2. As the purpose of the paper is not the validation of the WACCM model, the focus of the plot should more clearly be the regions in which the transport time scales are faster than the chemical time scales.

Reply:
We agree with the reviewer. Now Fig. 2 has been updated to include the respective O_3 and CO lifetimes, as shown below. It is apparent that transport dominates at the lower UTLS region, where chemical species has longer lifetime. This is the region of the focus of our paper. At upper air (especially above 10 hPa) chemistry dominates so the chemical lifetimes are very short.

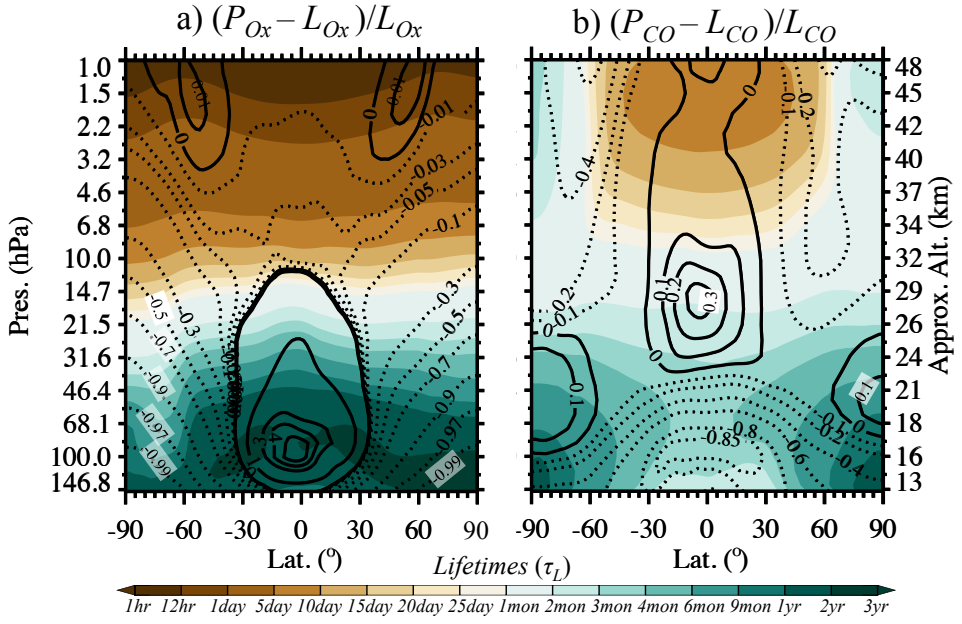

Fig. 2. The ratio of chemical net tendency (production rate minus loss rate) to loss rate from WACCM for **(a)** O_x and **(b)** CO. Negative numbers are dashed to highlight the net chemical decrease and positive numbers indicate net chemical increase, while zero lines indicate comparable amount of production and loss. For reference, the respective O_x and CO lifetimes are contoured in color.

2. The model setup does not consider any mixing (if I understand it correctly). It may not be so important for most results shown here which are mostly averages, but it is not clear, how in general the neglect of mixing influences the results. Especially in correlations like those displayed in fig 10, the process of mixing should change the results.

Reply:
During the trajectory integration we didn't consider mixing of parcels, which allows us to trace parcels back or forth to see the full history or future evolution of parcels. However, there is an effective 'mixing' when many parcels are averaged within grid boxes to be compared with either observational or Eulerian model results. The mixing in extra-tropical tropopause is very important, but we mainly focused our results around the tropical tropopause, where the strong vertical gradients of chemical species indicate less mixing occurring. In fact, it is because Lagrangian models producing non-diffusive transport and thus are especially accurate in regions where there are strong tracer gradients (e.g., the edge of polar vortex, the tropopause).

Fig. 10 shows the tracer relations at 68-hPa, where both O_3 (Randel et al., 2007) and CO (Abalos et al., 2012) exhibit strongest vertical gradients. At this level our results agree with MLS very well (see Fig. 4 and Fig. 9), so Fig. 10 might state the true atmospheric tracer relations.

3. The model air parcels are initialized at the 370-K level from the MLS climatology. This is typically between 100 and 150 hPa in the considered range (40◦N-40◦S). A comparison of the model results with data at 100 hPa is close to just comparing the initial conditions. Differences at 100 hPa (figs 4b, 9b) are potentially more due to vertical interpolation of the data than due to any process reflected by the model. What is the typical age of the trajectories at 100 hPa in the plot?

Reply:
We totally agree with the reviewer on this point. Parcels at tropical 100-hPa are usually very young with ages of 2-3 months. The reason that we chose 370-K as the initial level is based on the total diabatic heating shown in Fig. 1, in which MERRA's negative heating at 150-130 hPa makes a transport barrier that prevents air ascending to the stratosphere. Therefore, we have to choose a relatively higher initialization level (370-K) to avoid this negative heating altitude. Because this paper presents comparison of using both MERRA and ERAi circulation, for fair comparison we ought to use the same initialization level.

However, we could have initiated parcels at lower altitude, such as 355-K, when using ERAi circulation (shown in figure below). The brown line marked "TRAJ_ERAi i355K"

indicates the ERAi run with parcels initialized at 355-K level. Compared with the ERAi run initiated at 370-K (orange) it is clear that the 100-hPa O_3 and CO are well represented even it is close to the initialization level. We have added this in the Discussion.

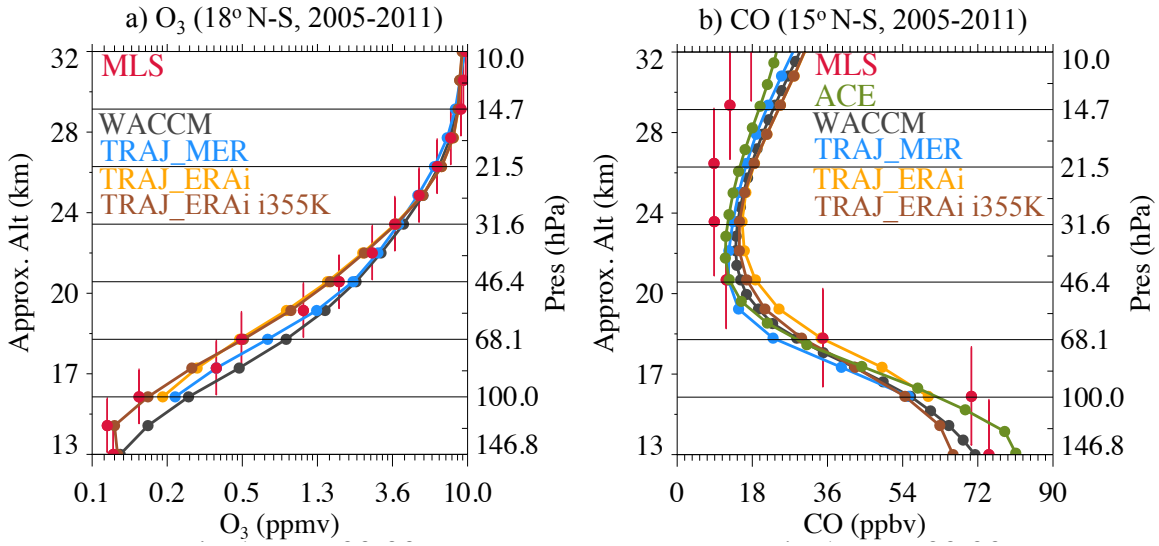

Figure. Tropical vertical profile of MLS, WACCM, and trajectory modeled (a) O_3 and (b) CO driven by MERRA and ERAi wind (initiated at 370-K isentrope), averaged over the deep tropics from 2005 to 2011. A test run driven by ERAi wind and initiated at 355-K isentrope is shown in brown as reference. Vertical bars in red indicate the MLS vertical resolutions at each of the MLS retrieval pressure levels.

4. Is in the comparison with ACE-FTS (e.g. fig 8) the latitudinal sampling taken into account? The plots could be either zonal mean cross sections (as indicated in the caption) or based on model interpolations onto the exact observation locations. The pattern of sampling times and latitudes of ACE-FTS may cause some of the shown difference.

Reply:
The ACE CO in Fig. 8 is a zonal mean cross section, and our comparisons are simply with zonal mean results from the trajectory model (not sampled like ACE). Park et al (2013) have recently shown similar model comparisons to ACE results, and demonstrated that there is almost no difference if the model is sampled identically to ACE measurement locations.

5. The critical point in the simulation is the method, how the diabatic heating rates are determined. It is said that they are determined including all radiation, latent heat etc. Please verify that this is the case for all reanalysis data sets. This is not trivial, since not all terms are equally saved in all data sets and must be reconstructed.

Reply:
In this study we took the total heating term directly available from each reanalyses. The

details of the diabatic heating results from different reanalyses have recently been discussed by Wright and Fueglistaler (2013), and this is highlighted in our revised manuscript.

Minor Points

1. 5999/ fig 4a: The error bars probably denote the vertical averaging kernel. From that it seems that one cannot decide whether vertical velocities derived from ERAi or MERRA are better. Error bar/uncertainty of the mixing ratio would also be interesting.

Reply:
The error bars denote the MLS vertical resolutions associated with each pressure level. They are obtained from the MLS data quality statement in Table 3.17.1 (http://mls.jpl.nasa.gov/data/v3_data_quality_document.pdf). We cannot decide whether ERAi or MERRA vertical velocities are better from this figure, but we can get some clue from Fig. 10.

We used to add error bars/shading to each datasets but it turned out to be too busy to read, and all those error bars actually didn't tell us more information. So we decided to only use profiles with MLS vertical resolutions as references.

2. Fig. 1: right y axis label (pressure) is not completely visible

Reply:
Thanks for reminding us. This is probably due to the formatting by the journal online version. We will make sure to remind the journal typesetting to adjust it in the next version.

3. 5995/fig1 caption: different latitude ranges are given. Is it 15 or 18 degrees?

Reply:
For the tropical vertical profiles we prefer averages within 18^o N-S. One exception is Fig. 9a, in which we added ACE CO for comparison. We obtained the ACE CO data directly from Park et al., [2013], in which the gridded CO is only available at 15^o N-S. For fair comparison, here we used all other datasets within 15^o N-S, too. Noted that there is barely any difference if we averaged over 18^o N-S.

[References]

Abalos, M., Randel, W. J., Kinnison, D. E., and Serrano, E.: Quantifying tracer transport in the tropical lower stratosphere using WACCM, Atmos. Chem. Phys. Discuss., 13, 13245-13283, doi:10.5194/acpd-13-13245-2013, 2013.

Park, M., W. J. Randel, D. E. Kinnison, L. K. Emmons, P. F. Bernath, K. A. Walker, C. D. Boone, and N. J. Livesey (2013), Hydrocarbons in the upper troposphere and lower stratosphere observed from

ACE-FTS and comparisons with WACCM, J. Geophys. Res. Atmos., 118, 1964–1980, doi:10.1029/2012JD018327.

Randel, W., Park, M., Wu, F., and Livesey, N.: A large annual cycle in ozone above the tropical tropopause linked to the Brewer–Dobson Circulation, J. Atmos. Sci., 64, 4479–4488, 2007.