Review of "Evaluation of simulated photochemical partitioning of oxidized nitrogen in the upper troposphere" by Henderson et al.

This paper presents a stochastic convective modeling approach used in conjunction with both in-situ chemical observations/modeling results and meteorological information in the upper troposphere as a way to compare estimates of air mass photochemical aging. Distributions from a suite of chemical simulations are used to compare various chemical mechanisms. In general, I find this approach very interesting and exciting. As far as I know, this is a new approach for airborne data analysis, and an important addition to the field. I did find several problems/issues within the paper, however, that must be addressed before I can recommend publication.

In the abstract, the authors cite observational based studies that are purported to show that models often over-predict upper tropospheric HOx, leading to too-fast conversion of NOx to HNO3. This claim is incorrect. None of the studies listed show this to be the case, and some, in fact, find the exact opposite (a significant model **under**prediction of upper tropospheric HOx). Bertram et al. (2007) only use observations of OH and HO2 in their analysis and make no claims or suppositions about model predictions of HOx. Ren et al. (2008) looks at INTEX-A data and find a very large model underprediction of HO2 above 8 km, with a median obs/calc ratio of 2.05 (~3 for the 10-12 km range). OH predictions were within 10% of observations in that same study. Olson et al. (2006) look at the tendency of models to **under**predict HOx in the upper troposphere with increasing NOx concentrations and small-scale variability.

Following this, (1. Introduction, top of p. 20128) the following statement is incorrect: "Recent chemistry evaluations suggest that chemical representation could contribute to underpredictions of NOx via overpredicting HNO3 formation rates (Olson et al., 2006; Bertram et al., 2007; Ren et al., 2008). These studies do not support this statement.

The authors need to state what value they use for NO. Table 2 does mention that calculated NO is used, rather than observation (and this is a good choice), but this should be brought out in the text as well. Additionally, Table 2 lists the HNO3 measurement used as from CIT (similar to NO, I would also suggest this be mentioned in the text).

The measurement of HNO3 from the UNH instrument is quite different than values from the CIT during INTEX-NA (median ratio between the two for 8-10 km is ~0.68). This would result in NOx/HNO3 ratios nearly 50% larger than those derived using CIT-HNO3. How would these differences impact your calculated mean air parcel lifetimes and conclusions? A discussion of the uncertainty in your results due to the choice of HNO3 measurement needs to be included.

Figure 5: In calculating the statistical differences for these species, did the authors include instrument uncertainty? For instance, though the statistics identified O3 from SAPRC99 as statistically distinct from observations for the mid-age and old subsets, I have a hard time believing that such a small difference (few ppb) is significant.

A few minor points:

1. Introduction

Nitric acid is not chemically stable at higher temperatures. Its reactivity is temperature dependent.

2.2 Observations

Line 16 (H2O<200 ppb) typo? Should this be ppm?

Whichever the case, this discrimination of stratospheric air confuses me. 200 ppm H2O is much too high to use to discriminate stratospheric air, but 200 ppb is extraordinarily low (i.e., when I obtain the 10s merged INTEX-A data between 8-10 km, I see no points that have water vapor (from the DLH instrument) less than 18 ppm). On the other hand, I find that 37% have H2O < 200 ppb, and of those points, 91% have O3 < 100 ppb (as low as 33 ppb), which is certainly not indicative of stratospheric influence. Why not use a combination of high O3 and low CO, along with the radioisotope ratios?