We thank anonymous referee #2 for their helpful comments on this work. Below, original comments are in italics and our responses are in bold.

P5370, lines 2-4: the sentence about averaging kernels seems a bit technical and out of place in the abstract, I think it would be beneficial for the reader to change this description to say something about the inversion sensitivity to the emissions.
We have added a definition of the averaging kernel in the abstract as well as in Section 5.4.

P5370, lines 12:15: I think the first two sentences would read better if merged into a single sentence
We prefer to keep this as two sentences, but have included the word "also" in the second sentence to improve the flow.

P5371, line 28: suggest changing "scale on which" to "scale over which"
We have made this change.

P5372, line 25: it would be useful to have a brief description of CARIBIC, including a definition of the acronym, on the previous page as is done for HIPPO. Would it also be useful to reference any previous studies of the CARIBIC N2O?
We have added a definition of the acronym here. Previous studies by Schuck et al. and Assonov et al. are cited later in Section 4.

P 5373, line 11: suggest moving GEOS-5 to go after the definition.
We have made this change.

P5373, line 12: clarify longitude and latitude for horizontal resolution?
We have made this change.

P5373, line 16: suggest using "anthropogenic" rather than "anthropogenic sources"
We have made this change.

P5374, line 9: define what MERRA means.
We have made this change.

P5374, Section 4: it might be useful to include a short paragraph on satellite observations of N2O either at the end of this section or in the introduction, especially as this is very briefly touched on at the end of the summary section.
Because we do not include any satellite-based observations of tropospheric N_2O in this study, we prefer not to include a paragraph describing them in the methods section. While there has been recent work developing tropospheric N_2O retrievals from satellite-based infrared sounders (such as the AIRS retrieval mentioned in the summary), these are currently not validated or publicly available.

P5379, line 7: suggest changing "based on" to "using"
We have made this change.

P5381, lines 1-7: it would be useful if the authors could comment on the sensitivity of the inversion to the vertical profile of the measurements here – it looks to me as though the HIPPO observations provide a stronger constraint because they extend throughout the troposphere to the surface, therefore, do the authors have any sense as to the altitude range at which the constraint breaks down? This would be especially useful in the context of Section 5.5 and would maybe make a strong statement on the importance of in situ profiles from aircraft as part of the global observing system.

This is a good point. We have added some text to mention that the HIPPO observations include profiling from the boundary layer to the upper troposphere, which is likely the major reason why it provides a stronger constraint than CARIBIC. Additionally, we point the reviewer to the error reduction results in Figure 9, which show that significant error reduction is achieved with CARIBIC only in the vicinity of Frankfurt where flights in the lower troposphere occurred.

P5381, line 23: should "sources as well as sinks" be "sinks as well as sources"? the previous subsection has dealt with sources and this one talks more about the sink.

We have made this change.

P5382, line 7: stating that the second year is the final year seems a bit unnecessary, also the statement that "the inversion does not significantly affect the observations" seems to be the wrong way around to me – isn't it the impact of the observations on the inversion which is being assessed?

We choose to keep the statement that the second year is the final year as a reminder to readers. We have reworded the latter statement as we agree it was confusing as is. We have changed the text to say "Stratospheric loss of N_2O in the second (i.e. final) year of the inversion does not significantly affect the N_2O mixing ratios at the observation locations" meaning that biased stratospheric loss will not have a measurable signal in the troposphere during the second year of a two-year simulation. Because of this the (pseudo) observed mixing ratio and model mixing ratios are very similar, and there is insufficient forcing to correct a prior bias.

P5383, line 1: it's hard to tell from the Figure that there is any significant change in the a posteriori compared to the a priori.

The changes between a priori and a posteriori in the right-hand panels of the figure are indeed very subtle. We have zoomed in the y-axis range to try to address this (from 0 to 1.5 instead of 0 to 2) and also added some text to note that the deviation from the a priori is very slight.

P5383, lines 24-26: it would be useful if the authors could comment here on any vertical correlations that might help to address this, or which limit the impact of aircraft observations measured at cruise altitude

We have made this addition.

P5385, lines 1-2: it might be useful if the authors could comment briefly on how the box model results relate to transport across the tropopause on different timescales (e.g. tropical vertical mixing vs. isentropic mixing in the extratropics)
The disclaimer on P5384, lines 5-7 indicates that the box model results do not capture seasonal effects or spatial gradients within the stratosphere/troposphere. It would also not capture stratosphere-troposphere transport mechanisms that occur on different timescales. We have added this to the text.

P5386, line 2 and P5387, line 2: please check the consistent use or a priori or prior and a posteriori and posterior.
We have modified the text to include the use of a priori and a posteriori throughout.

P5391, line 25: clarify that the surface observations are both in situ and flask measurements.
We have made this change.

P5392, line 19: is this statement on model transport specific to GEOS-5? It might be useful to add a comment on model transport issues based on other studies using relatively long-lived constituents (e.g. CO2 or CO)
Model transport issues are not specific to GEOS-5. In a TransCom model intercomparison, Thompson et al. (2014, ACP) found the N_2O gradient across the tropopause to vary significantly among eight different chemical transport models, presumably due to differences in modeled vertical transport and rates of strat-trop exchange. We added a citation of that paper at this point in the text.

P5393, lines 5-15: I think that it would be of benefit to the reader if the authors could rephrase the description of the averaging kernels to what it means in terms of the sensitivity of the inversion to the emissions (similar to my comment for the abstract) – by all means this should then reference the averaging kernel values but would be easier to understand the broader significance.
We have added some text to Section 5.4 defining the averaging kernel as the sensitivity of the inversion to emissions in each grid square. However, we feel we have already highlighted the broader significance of the averaging kernel, which is that it tells us about the observational constraints on emissions achieved in a particular location. In instances where the averaging kernel for a location is close to 1.0 at that location and close to 0 everywhere else, the local emissions are well-resolved by the observations in our inversion framework. Conversely, highly smeared averaging kernels indicate that emissions in a particular location are likely to be conflated with those in other parts of the world. This has important implications for how we interpret the inversion results.

P5393, line 13: is it an underconstraint or no constraint in the tropics?
We feel underconstraint is more appropriate than no constraint, because we have only shown rows of the averaging kernel for one tropical location, which does not necessarily mean there is zero constraint throughout the tropics.

P5393, line 21-23: could this also be linked to requirements for similar targeted measurements for other atmospheric constituents, and greenhouse gases in particular? How do the findings here compare to similar studies for CO2 and CH4?
Targeted aircraft measurements in the tropics would likely also be useful for any species that has a significant tropical source. N_2O is somewhat unique in that elevated concentrations have been measured aloft in the tropics (e.g. Kort et al., 2011, GRL), indicating the presence of a large, episodic emission source in the vicinity of tropical convection. So, aircraft observation in this region may be particularly useful for N_2O, and would complement the use of surface observations. We have added some text to Section 5.5 to note this.

P5394, line 1: it would be useful to have a brief sentence on satellite observations in the introduction.
As in our response to the similar comment above, we choose not to include any additional information in the introduction since satellite observations are not used in this study.

P5409: clarify that HIPPO or CARIBIC are aircraft measurements.
We have made this change.

Figures 6 and 7 would benefit from some clearer labelling of the plots and reference in the captions.
We have edited the captions for both figures to be more explicit about where the labels are located and which lines refer to which label.