

General Comments

A very unique set of measurements obtained during the ATom-1 mission was analyzed throughout this study. Comprehensive data sets, both from in-situ and satellites, and a state-of-the-art global chemistry model were utilized. The analyses methods used here were reasonable and the results are presented in logical manner. However, it is not clear to me that what the scientific backgrounds are and what scientific questions the authors try to answer. What new and unique findings are discussed in this study? What can we learn from the results presented here? Specifically, why is it important for GEOS-5 to be able to simulate and forecast CO over the ocean? And why ATom-1 (and also the future deployment phases) was planned and how is it different from the previous campaigns, for instance, HIPPO? I think including more scientific background and contexts, and specific goals will make this article far more interesting.

We thank the referee for the thoughtful comments. We have expanded the introduction to include some of the goals of ATom and relate our modeling work to those goals. In particular, we now state: “Major goals of the Atom mission include identifying chemical processes that control the concentrations of short-lived climate forcers, quantifying how anthropogenic emissions affect chemical reactivity globally, and identifying ways to improve the modeling of these processes” and “GEOS-5 forecasts help determine the source regions and emission types that contribute to the trace gas and aerosol concentrations measured during ATom, which is directly relevant to the goal of quantifying how anthropogenic emissions affect global chemical reactivity.” We also clarify the importance of validating GEOS-5, adding “GEOS-5 supports numerous aircraft missions, and validation of the model forecasts is important for developing confidence in and understanding the limitations of chemistry forecasting for aircraft missions. The ATom dataset, which uses unbiased sampling rather than chasing plumes, provides a unique opportunity to validate the overall performance of the GEOS-5 model on a global scale.”

We also added a supplemental figure showing a map of the ATom flight tracks.

In the future, the ATom Mission PI will prepare an overview paper that contains more details on the scientific background, planning, goals, and overall accomplishments of ATom. An extensive discussion of these topics is thus beyond the scope of the present work.

Specific Comments

P1 (Abstract) – I think the abstract can be rewritten in order for it to serve as a brief summary of this study. The simplest way is to rearrange it. 1) What is the goal of ATom-1, 2) What are the observational findings shown in this study, 3) Why is the modeling study performed and 4) What the conclusion is.

We reorganized the abstract to begin with a description of ATom and better motivate our modeling study. We add “The first phase of the Atmospheric Tomography Mission (ATom-1) took place in July–August of 2016 and included flights above the remote Pacific and Atlantic oceans. Sampling of atmospheric constituents during these flights is designed to provide new insights into the chemical reactivity and processes of the remote atmosphere and how these processes are affected by anthropogenic emissions. Model simulations

provide a valuable tool for interpreting these measurements and understanding the origin of the observed trace gases and aerosols, so it is important to quantify model performance” and “We use GEOS-5’s tagged tracers for CO to assess the contribution of different emission sources to the regions sampled by ATom-1 to elucidate the dominant anthropogenic influences on different parts of the remote atmosphere”.

P2 (Introduction) – This section can be rewritten to add clarity as well. It should include 1) Scientific background and goals of the ATom-1 (and other phases) mission, 2) What are the scientific questions ATom-1 aimed to answer? 3) What did ATom-1 sample and accomplish? 4) What is the goal of this study?

We have expanded the introduction as described in the response to General Comments to include information on the goals of ATom and how our work relates to these goals. We also added a link to the ATom website, which includes additional information about ATom.

In addition, the information in the third and the fourth paragraphs (lines 42-53) provides general overview, instead of specific to this study.

We reorganized this section to better link the overview to the current study.

I would also like to know what the ATom-1 specific references are. Is Prather et al. (2017) the only reference relevant to ATom-1?

We also reference the doi for the ATom dataset. As noted above, an ATom overview paper is planned but not yet published. We added a reference to *Prather et al.* [2018] to Section 5.

Technical Comments

P3 (Section 2.1) – What are the vertical ranges of ATom-1 measurements? The accuracy or precision of the species measured during ATom-1 should be provided here.

We state that “Each of the 11 flights included sampling from the boundary layer to the top of the aircraft range (39 kft).” The QCLS observations have an accuracy and precision of 3.5 ppb and 0.15 ppb, respectively. We added that statement to this section.

P3 (Section 2.2) – The vertical range of MOPITT and MLS and their overlap have to be mentioned specifically. What pressure ranges do they overlap?

We now clarify that we are using the CO column and 700 hPa CO retrievals from MOPITT and the 215 hPa level for MLS. We also add the following about the overlap: “The MOPITT averaging kernels include some sensitivity to the 200 hPa level, implying a small overlap between the MOPITT and MLS observations.”

How often did ATom-1 sample stratosphere? Are there any ground-based measurements to compare? It has to be mentioned why only the satellite data are used in the comparison. Also, why was the MODIS AOT used?

We use satellite observations for our analysis of interannual variability because they provide broad spatial coverage over the oceans, where most of the ATom sampling occurred. Ground-based measurements over the oceans are sparse, so we do not include

them in this analysis. We use MODIS AOT to examine aerosol interannual variability because MODIS provides a relatively long record. We added text clarifying these points: “We focus on satellite observations because they provide broad coverage over the oceans, where surface data is sparse. “ “We use MODIS data in this analysis because it provides a relatively long data record.”

P3 (MOPITT) – I believe the current version of MOPITT is 7 not 6.

This is correct. We used version 6 in this study because it was a well-established product at the time we conducted this analysis.

P4 (Section 2.3) – It has to be mentioned why the model is used in this study. What are the questions that can be answered through modeling? Is a global model the sufficient tool for this study?

We added the purpose of the model to the first sentence, and clarify that a global model is necessary because CO is transported globally. We added the following text about why we use GEOS-5:

“to quantify the contribution of different emission sources to the observed CO distribution and to identify the origin of observed plumes. A global model is necessary for this analysis since CO is transported globally.” and “We use the FP system in our study because it is the system used to generate forecasts that are used during ATom and other aircraft missions, and is thus relevant to future mission and flight planning.”

P4 (L125) – What could be learned from the chemical forecasts during a field campaign, such as, ATom-1? How many pollution or non-pollution events occurred during the ATom-1 period?

We add a discussion of what chemical forecasts tell us to the first paragraph of Section 3.

Pollution plumes are shown in figures 1 and 2. We add this text to Section 3:

“The chemical forecasts provide the ATom team with a preview of the chemical environments that the flight is expected to sample, including the location of pollution, biomass burning, or dust plumes; regions of substantial but well-mixed anthropogenic pollution; and cleaner regions. The forecasts also provide a broader spatial context for the observations, since the 3-dimensional model output shows the spatial extent of features that intersect the flight track.”

P5 (L160) – Is the bias in the emissions responsible for the underestimation in the model? What about the spatial resolution of the model?

These are both possible explanations. We add the following clarification:

“Either biases in emissions or insufficient vertical or horizontal model resolution may thus be responsible for the underestimate.”

P5 (L167) – ‘As expected from the basic chemistry and seen in previous observations’ can be replaced by ‘In the stratosphere’.

We now say “As expected from stratospheric chemistry and seen in previous observations”

P5 (L169) – ‘Underestimating the observed decrease’ can possibly be replaced by ‘showing higher CO’ or something similar.

We changed it to “underestimating the observed gradient” since we want to convey that the change in CO as the flight enters the stratosphere is underestimated.

P6 (L184)- The meaning if this sentence is unclear. So, using analysis and forecast wind fields makes no difference? Does this mean that the wind fields are the same or the results are not sensitive to the winds? And what is the reason for this?

It means that the plume timings are already captured by the forecast winds and don't change substantially with the analysis winds. We reword this sentence to say: “Comparison of Fig. S4 with Fig. 2f shows that the impact of using analysis versus forecast wind fields is small for this flight since the forecasts already capture the timing of the plumes.”

P6 (L193 & L203) – What determines the placement of the plumes in the model? Do models misplace the plumes often?

We added “extent” as well as placement of the plumes. In general, the good agreement in the timing of plumes in the model compared to observations suggests that the model usually places the plumes correctly. However, the placement of plumes is sensitive to vertical and horizontal transport, and to the injection height of the biomass burning emissions. The injection height is a source of uncertainty. In addition, global eulerian models tend to diffuse plumes too quickly. We added the following text to explain this: “The placement and strength of simulated plumes is sensitive to the injection height of the biomass burning, which is a source of uncertainty. In addition, plumes in models tend to dissipate more quickly than in observations due to the numerical effects of limited model resolution (*Eastham and Jacob, 2017*).”

P7 (Figure 4 description) – It is not easy to compare the data and the model simulations in this figure. I wonder what the differences (data-model, or %) would look like if it's plotted like Fig. 4a and 4b.

We added a third panel plotting the data – model difference.

P9 (L306) – Full name for GFED version 4 and a citation should be included.

We added the full name and the reference to *van der Werf et al., [2017]*

P9 (L312) – ‘its relative variability’ – what does this mean?

Relative variability means the variability relative to the mean. We add this clarification to the text.

P10 (L320) - A reference is needed for Siberian biomass burning here.

We provide references later in this paragraph for large fires in western Russia and Siberia.

P11 (L353-354) – Is this expected? And why?

Both CO and aerosols have a major source from biomass burning, so it makes sense that regions influenced by biomass burning, such as the tropical Atlantic, would have high values of both CO and aerosols. We add the following:

“This similarity is consistent with the importance of biomass burning emissions for both CO and aerosols.”

P11 (L366-367) - I wonder if and how the MOPITT V6 bias drift is accounted for in this study.

We do not attempt to account for the MOPITT drift, except to mention it in the text. Since the drift in the CO column is almost negligible [Deeter *et al.*, 2014], we do not expect it to influence our CO column results. There may be a small impact from the drift on the 700 hPa results, but those results are generally consistent with the CO column results.

P12 (L402) – A general website for the ATom mission will be useful to include here.

We added a link to the ATom website in the introduction.

P12 (references) – Formatting of the references should be corrected to be consistent with the publication standard.

We have formatted the references in this way.

P17 & 18 (Tables 1 & 2) – A map showing all the geographical regions (with specific latitude and longitude) used in Tables 1 & 2 might be useful. I am also curious to know if it would be easier to compare the observations and the model in a form of a bar graph instead of a table (Table 2).

We added a supplemental figure showing maps of the geographical regions corresponding to the tagged tracers. We added another supplemental figure showing the ATom flight tracks on a map and labeling the start and end points of the flights, so that the reader can visualize the locations discussed in Table 2. We prefer to keep Table 2 as a table so that we can more easily and quantitatively compare the standard deviations as well as the mean values.

P18 (Fig. 1) – I wonder what would Figs. 1d-1f and Figs 1g-1i would look like if they have the same vertical ranges. The current plots give a false impression that CO is higher between Kona-Pago Pago-Christchurch than the first leg.

We considered this, but putting panel e on the same scale as panel d squashes the detail in panel d.