

## Replies to reviewer report 1:

We thank the reviewer for the detailed and supportive comments. Below are our point-to-point replies.

**This manuscript explores an important topic regarding the recovery of the stratospheric ozone layer, in particular the emission rates of halogenated very short-lived substances (VSLs) at the surface. These compounds can be rapidly transported into the stratosphere, especially over areas of active, deep convection, and affect stratospheric ozone levels, thereby delaying the recovery process. very few observations exist in key areas of the world where VSLs emissions and transport rates can be significant to the stratospheric budget of inorganic bromine, so the availability and use of new aircraft data sets constitutes a unique opportunity to test our models and evaluate our inventories. I will first provide some general comments to the manuscript, followed by some specific ones.**

We thank the reviewer for the insightful comments.

### General

**1. While the topic is of great relevance, it is unclear what the contribution of the manuscript is, as written. Similar flux calculations were done by previous studies (lines 171-173). Is the contribution of this manuscript related to the methodology used, to the new aircraft data set over the tropical Western Pacific, and/or to the new magnitudes of fluxes obtained in this study?**

We have applied a MAP approach to inferred  $\text{CHBr}_3$  and  $\text{CH}_2\text{Br}_2$  fluxes over tropical Western Pacific from the new CAST / CONTRAST experiments. Currently there are large differences in the distribution and magnitude of between existing  $\text{CHBr}_3$  /  $\text{CH}_2\text{Br}_2$  inventories (see, for example, the new Figure 6).

To our knowledge, we are the first to use the MAP approach to infer  $\text{CHBr}_3$  and  $\text{CH}_2\text{Br}_2$  surface fluxes over open oceans in tropical Western Pacific region, supported by new data from the CAST/CONTRAST campaigns. Our posterior estimates consistently show systematic deviations from the three independent prior inventories (see new Figure 6). These results have now been emphasized in the discussions (Page 11).

**2. The results presented in the manuscript are based on numerous model assumptions (e.g., lines 215–226). Were any sensitivity tests performed on the choice of values used? Are there any references to justify the choice of values used? Paragraphs 4 and 5 in the Introduction highlight how previous studies were based on several (different) assumptions and how those results need to be examined with caution. How can the results from this analysis, along with the assumptions used, be compared against previous studies?**

The reviewer is correct that we do introduce several assumptions to help infer surface fluxes from aircraft measurements. As with other top-down flux inversions, we assume prior knowledge and its uncertainty. In the revised manuscript, we include sensitivity tests to test these assumptions about the magnitude of uncertainty (Lines 223-235), and three different sets of prior fluxes (Figure 6). The results from these tests demonstrate the robustness of our  $\text{CHBr}_3$  fluxes over the main study domain where

observation coverage is relatively dense. Our posterior model simulations at two different spatial resolutions (revised Figure 2) are also in better agreement with observations than those based on prior inventories. However more independent data including the direct flux measurements are needed to fully evaluate our results in particular for CH<sub>2</sub>Br<sub>2</sub> fluxes (See discussion).

**3. The type of correlation between bromoform and dibromomethane is of importance. What is the rationale for a linear correlation used in several studies published earlier (e.g., lines 103-105)? Given that the new aircraft data set elucidates a different correlation between the two 2 compounds, elaborating some more on this topic will highlight one of the new findings from this study.**

Several previous studies have assumed the linear correlation between CHBr<sub>3</sub> and CH<sub>2</sub>Br<sub>2</sub> based on some observations as well as on the assumption about the shared biogenic sources. However other measurements and model studies suggest a rather complicated correlation between these two species. Our inversions also show no evidence to support such a simple linear relation. But our posterior flux uncertainties, in particular for CH<sub>2</sub>Br<sub>2</sub> (Figure 3), are too large for us to reach a definitive conclusion.

**4. There are several instances of missing punctuation marks such as commas and periods throughout the manuscript.**

During the revision, we have corrected the punctuation.

**5. Some of the listed uncertainties are significant (e.g., line 262, line 342). What is the impact of these uncertainties on the conclusions of this study?**

Our current estimates, in particular for CH<sub>2</sub>Br<sub>2</sub> still have large uncertainties, limited by the observation quality and coverage, as well as by transport model errors. We believe they will only be addressed by more coordinated fluxes/concentrations measurements.

### **Specific**

**6. Abstract, line 25: An r value of 0.38 does not really qualify as “reasonably consistent” correlation.**

Here the consistence is about the agreement with vertical distributions. We have clarified the text.

**7. Abstract, line 36: which a priori inventory was used for the comparison?**

It is from Ordonez et al. (2012). We have clarified this in the abstract.

**8. Introduction, line 47: “The wide range of ... lifetimes allows for ...”**

Changed the sentence as suggested

**9. Introduction, line 55: “There is a wide range of...”**

Changed the sentence as suggested

**10. Data, paragraph 1: Are there any references available for the CAST and**

## **CONTRAST instruments?**

See Andrews et al. (2016).

**11. Data, lines 146-149: What are “mean absolute percentage errors”? Which data set is higher? Are the differences uniform with height? How is this metric used in the analysis and how does it impact the results?**

Their deviations change with altitudes (see Andrews et al 2016). By design, our inversions depend on observed horizontal and vertical gradients in the boundary layer, mainly observed by CAST. Most CONTRAST measurements are at much higher altitudes and hence less sensitive to local sources.

Our sensitivity experiments ), in which we introduce a bias between CONTRAST and CAST data that we infer in our inversion, show very similar results to our control run. We have included this in the main text.

**12. Data, lines 146-149: The second half of the statement is confusing as stated. WAS refers to the collection method and GC/MS to the analysis technique. Each campaign provided one data set. It might be simpler to state “...between the CAST and CONTRAST instruments”, instead.**

Good suggestion. We have changed accordingly.

**13. Data, line 157-158: Is there a reference available to support the statement?**

See Butler et al (2016)

**14. Data, lines 163-164: Are the referenced data from NOAA’s ground network collected at the surface? Given that this study examines data at higher altitudes as well, are there any model comparisons with data at higher altitudes?**

Yes. They are NOAA surface measurements. Unfortunately we cannot find independent aircraft measurements to evaluate our prior or posterior model simulations.

**15. Data, line 186-187: Is a 6-month spin-up enough time and seasonally appropriate?**

It is appropriate because of the short lifetime (<4 months) of the species. In the inversion, we are also only focused on January and February, 2014.

**16. Results, line 248: Even with higher a priori ocean fluxes, the model still depletes bromoform much faster between the surface and 2 km than the observations show. Is this a result of chemistry, transport, and/or something else within the model?**

We believe the higher lapse rate is likely related to the issues with model vertical transport, as it has also been found for posterior model simulation (Figure 2) even at finer model resolution.

**17. Results, line 274-278: The right panel of Figure 2 shows that the model’s vertical distribution of bromoform is practically the same when run at coarse and fine spatial resolutions. This suggests that sub-grid convection, assuming**

**that the model resolves some events at the finer scale used, does not play a significant role in the modeled vertical profile. Is this result expected for a tracer with a relatively short lifetime and over a region of active, deep convection?**

The coarse and fine model simulations show different atmospheric lapse rate in the boundary layer for both prior and posterior surface fluxes. Even with our fine-scale model simulation (spatial resolution of ~25 km) there are sub-grid scale processes that are unaccounted. The role of model error is the subject of ongoing work.

**18. Results, line 293-294: How were the 50% and 30% chosen? How sensitive are the results of the analysis to these percentage choices?**

These percentages are chosen just to demonstrate observation constraints. The error reduction is insensitive to these values. Also, as shown in Figure 6, our inversion results are not sensitive to *a priori* fluxes over regions with proper observation coverage.

**19. Figure 1, line 445: Suggest using “15°S–25°N”**

Thanks. We have changed the latitude range as suggested.

**20. Figure 1 and Figure 2: Are the in situ data shown in these figures an average of both aircraft data sets?**

Figure 1 shows the altitudes of the CAST/CONTRAST measurements, and Figure 2 shows the CHBr<sub>3</sub> concentrations at the boundary layer, which are mainly observed by CAST.

## Reviewer report 2

We thank the reviewer for providing a second round of reviewer comments. Below are our point-to-point responses to the reviewers' comments (denoted in bold). In particular, following the reviewer's suggestion, we have conducted additional inversion experiments using Liang's inventories as our a priori, and include the results to the main text (Section 3 and Section 5). In short, we find that our posterior flux estimates are robust against using different prior emission inventories, subject to data coverage.

### General comments

**This manuscript describes an inverse modelling study designed to infer biogenic, oceanic emissions of two short-lived bromocarbons (bromoform and dibromomethane- CHBr<sub>3</sub> and CH<sub>2</sub>Br<sub>2</sub>). The authors use an emission inversion setup consisting of a global chemistry transport model, a priori emission inventories for CHBr<sub>3</sub> and CH<sub>2</sub>Br<sub>2</sub>, and aircraft measurements from two separate campaigns measuring both compounds over the tropical Western Pacific. The authors also carry out a short observing system simulation experiment to retrieve a set of known idealised emissions as a means of proving the efficacy of their inversion modelling setup. The authors conclude that the a priori emissions of CHBr<sub>3</sub> and CH<sub>2</sub>Br<sub>2</sub> are too high over this region, and find the a posteriori emissions of CHBr<sub>3</sub> and CH<sub>2</sub>Br<sub>2</sub> to be lower than the a priori emissions by 40% and 20%, respectively. They also conclude that assumptions in previous studies of a correlation in the emissions of CHBr<sub>3</sub> and CH<sub>2</sub>Br<sub>2</sub> cannot be supported based on the findings of this work.**

**The subject matter of the article sits well within the frame of ACP. In addition, the objectives of the scientific study and the design of the experiment (on the whole) mean that this work provides a useful scientific contribution on a topic (i.e. the biogenic, oceanic emissions of bromocarbons) where we have relatively poor understanding. It is welcome to see studies moving away from heuristic methods for inferring CHBr<sub>3</sub> and CH<sub>2</sub>Br<sub>2</sub> emissions towards using more robust methods. I therefore find that this paper is a welcome and much needed scientific contribution.**

**Overall, I find the article to be well written and organised. The scientific ideas are laid out in a clear and logical manner, and consequently one can follow the flow of ideas easily. The article also does quite well at justifying the methodological choices, although there is one major issue here that I will highlight below in the specific comments section. Unfortunately, this issue does have a direct bearing on my recommendation for publication. Separately, and as a more minor issue, I did find that the authors stopped short somewhat of some deeper discussion that I feel would help strengthen the article, and I will explain this in more detail below.**

**I therefore find that this has the potential to be a very good scientific article. However, I cannot recommend publication until the issues outlined in the specific comments section are addressed.**

We thank the reviewer for their supportive comments. We originally chose the dibromomethane and bromoform ocean flux estimates from Ordóñez et al (2012). as our prior because they provided more detailed spatial patterns than Liang et al (2010). As we show in the new Figure 6, using two alternative prior inventories (Ziska et al (2013) and. Liang et al. (2010)) does not significantly impact our results subject

to coverage provided by the aircraft data

We broadly agree with previous studies (but for a larger geographical region) that the Ordóñez inventory overestimated bromoform emissions over the Western Tropical Pacific region by nearly 40%. In our study, we also found large differences between open ocean fluxes and coastal (island) sources, contrary to Ordóñez et al, and also different from the simple spatial pattern over open oceans suggested by the inventory from Liang et al. (2010).

However, we agree with the reviewer that additional inversions informed by alternative prior knowledge can further this study. We now present additional experiments that use Ziska's and Liang's prior bromomethane emission inventory. These are included in the main text but summarized below.

### **Specific comments:**

**1) The most significant problem I find in this study relates to the choice of Ordóñez et al as the sole choice for priori emission inventory for both CH<sub>2</sub>Br<sub>2</sub> and CHBr<sub>3</sub> emissions.**

**a. My first point relates to the CHBr<sub>3</sub> a priori emissions. I fully recognise the challenges Ordóñez et al. faced in creating an emission inventory using heuristic methods in a global model, and I fully respect the useful contribution Ordóñez et al. have made to our understanding of VLSL emissions. However, we now have several studies (including this study and Ordóñez et al. itself) that show that the Ordóñez et al. CHBr<sub>3</sub> emissions in particular are over estimated in the Western Pacific region.**

**i. In fact Ordóñez et al. (2012, ACP) itself in Figure 7 (the PEM-Tropics A, PEM-Tropics B, and TRACE-P panels) shows that their own model over estimates CHBr<sub>3</sub> in the Western Pacific region when using their own emissions.**

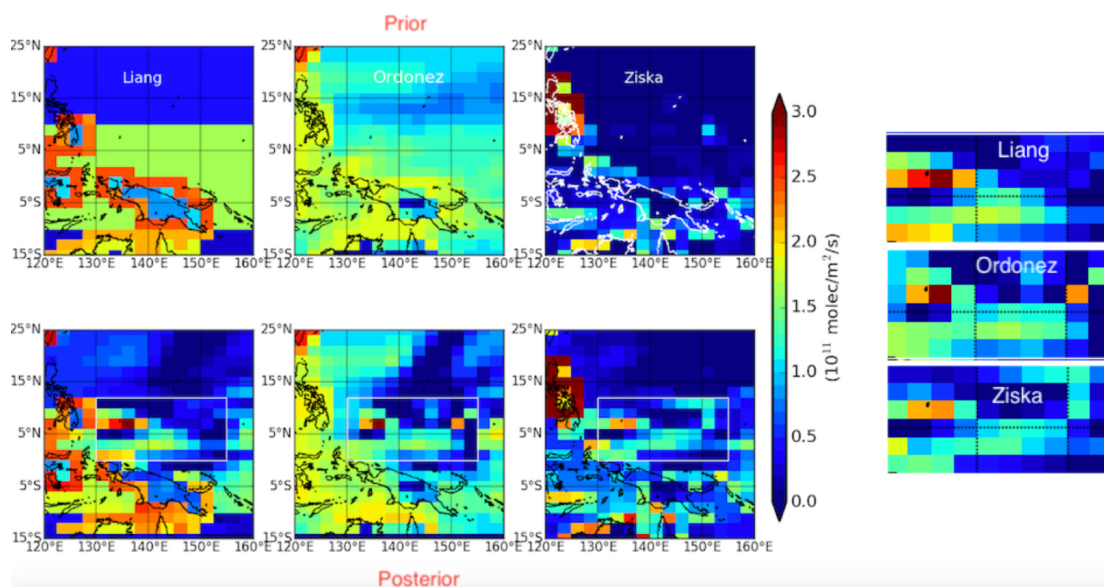
**ii. Ashfold et al. (2014, ACP) - another study employing a top-down method to infer VLSL emissions in the tropical Western Pacific - derived lower estimates of CHBr<sub>3</sub> emissions in the tropics than Ordóñez et al. Similarly, their retrieved western, northern and southern fringes. Can the authors please discuss how they think this issue affects their results.**

We agree that the aircraft observations do not uniformly cover the study domain, but some of these gaps are effectively filled by atmospheric mixing of surface sources. Our inversion system includes scaling factors not only for sources within the study region but also for neighbouring regions that lie outside our study regions and for the initial conditions at the beginning of the study period. Our sensitivity experiments reveal that our results are not sensitive to global priori inventory when the observation constraints are strong such as the for CHBr<sub>3</sub> emissions over the study domain between between 130°—155°E and 0°—12°N. This is discussed in section 4.

**3. Concluding my remarks on point 1), I strongly recommend, and as a condition of acceptance for publication, that in addition to running the emission inversion with Ordóñez et al. as the a priori for both compounds, that the authors also run their emission inversion algorithm with Ziska et al. (CHBr<sub>3</sub>) and Liang et al. (CH<sub>2</sub>Br<sub>2</sub>) for the two compounds. Comparing this work to that of Ashfold et al. (2014, ACP), one can see that Ashfold et al. (2014, ACP) undertook a variety of emission inversion experiments (including changing their a priori) to test the setup of their system. These aspects of Ashfold et al. (2014, ACP) strengthened their work, and, similarly, this manuscript would also benefit from a similar**

**effort.**

Figure 1 summarizes our results from using different prior inventories. Posterior flux estimates of CH<sub>3</sub>Br over the geographical region covered by CAST/CONTRAST aircraft data are remarkably similar. This supports the idea that the data are playing a significant role in determining the posterior flux estimates.



**Figure 1:** (Left upper panels) Prior and (left lower panels) posterior CHBr<sub>3</sub> flux estimates ( $10^{11}$  molec/m<sup>2</sup>/s) over the study region. The three prior inventories include Liang et al (2010), Ordóñez et al (2012), and Ziska et al (2013). The right panel is focused on the geographical region between (130°—155°E and 0°—12°N) between where the CAST/CONTRAST had the most information.

**2) Some key discussions seem to be missing including those of limitations of this study.**

**a. It would have been nice to see a discussion of the prevailing meteorology during the period of study and an explanation linking this to the error reductions that we see in the OSSE results in Figure 3. Presumably, the error reductions are a function of the location of the observations and the origin of air masses arriving at the observation locations. An analysis similar to what I am suggesting was carried out in Ashfold et al. (2014) in their Figure 2, which allowed them to determine where there inversion setup was able to retrieve emission values. I realise this is perhaps easier in the Lagrangian framework of NAME, but the authors could draw upon the information in their meteorological inputs for GEOS-Chem to create a climatology of the winds and then make a discussion that would add useful context to the results and strengthen the paper.**

We agree that it is of great interest to show the origins of air mass, which has been partially investigated by another study (Bulter et al, 2016). However, the complexity of the global CTM and its analyzed meteorological fields and the nature of the aircraft measurements precludes a simple and intuitive summary of the overall sensitivity of the CAST and CONTRAST observation to the underlying surface fluxes. In addition, we have included scaling factor for initial concentrations and for emissions from neighbouring regions, and as a result, the posterior flux estimates are more or less dependent on the difference between modelled and observed internal horizontal and vertical gradients, instead of single concentration values (which has also been

revealed by consistency in posterior fluxes when different a priori is used.

***b. It would be good to see the authors try to connect the results of the OSSE, i.e., the spatially limited pattern of the error reduction, to the areas in the a posteriori CHBr3 emissions where we see the largest reductions in absolute emission values relative to the a priori. Given the evidence I present in point 1) above, I do not believe that the similarity in the spatial patterns in the OSSE error reduction and the area of reduced a posteriori emissions is coincidental. I think this implies that with greater spatial coverage in the aircraft observations that we would see reductions in the a posteriori emissions covering a larger spatial area. The authors should discuss this point, and also conclude that the spatial extent of the aircraft observations provides a limitation for this study.***

This is an interesting point. Aircraft measurements are sensitive to a wider range of geographical regions than the error reduction suggests, e.g. Figure 2 and Butler et al, 2017. The inversion updates the ocean fluxes over a wide geographical domain but the error reduction is often small because of the low signal (contribution) to noise (observation error and model transport error) ratio.

***Further to point b., I do not find the a posteriori CHBr3 emission estimates outside of the region sampled by the observations (towards the N,S, W fringes of the domain) to be credible in light of the large reductions we see in the a posteriori compared to the a priori over the most sampled region. I am working on the assumption that the emissions are spatially correlated. Perhaps some discussion of this point in the context of the previous studies (e.g., those highlighted above in point 1) would help readers gauge the quality of the emission inversion in the areas on the N,S, W extremes of the domain where there is little information from the observations. This might also help readers understand the large gradients we see between the sampled and poorly sampled regions.***

See response to previous point.

***Figure colours in Figure 4 need greater differentiation. I struggled to differentiate the monochrome orange/brown tones. A set of panels representing the relative differences between the a priori and a posteriori emissions would also be of help.***

Agreed. The manuscript (Figure 4) has been amended, accordingly.

***4) I think it is necessary for the authors to include a discussion of the conclusions of Russo et al. (2015, ACP). Russo et al. (2015, ACP) made two conclusions relevant to the work in this study:***

***a) That it is difficult to infer emissions using aircraft measurements and coarse global models in the case where the emission distribution is heterogeneous in regions of strong convective activity.***

***b). That model resolution can affect the simulated distributions of CHBr3 in cases where the emissions distribution is heterogeneous.***

***The authors should include some discussion of these points and should explain how they present limitations for the current work, or why this points are not relevant to the conclusions in this manuscript.***



***The authors should include some discussion of these points and should explain how they present limitations for the current work, or why this points are not relevant to the conclusions in this manuscript.***

Certainly, using a finer-scale model resolution would be preferable. Our forward simulation at the native model resolution of  $0.25^\circ \times 0.3125^\circ$  confirm that posterior fluxes result in a better agreement with observations than the prior as shown in the revised Figure 2. However, the resolution of estimates fluxes is determined by the quality, quantity, and distribution of available data. In this case, even if we used a finer resolution model it is likely we would need to aggregate model grid values to generate estimates that do not include large spatial correlations. We have included this discussion in the revised manuscript, following the reviewer's recommendation (Page 11).

***5) It is important to note that the a posteriori emissions are more heterogeneous than the a priori. Therefore, following from Russo et al. (2015, ACP) and the discussions in point 4) above, the issue of model resolution could affect the simulated distribution of CHBr3 more significantly for the a posteriori emissions than for the a priori emissions. The authors have tested the impact of model resolution on the a priori emissions and found no effect. However, it seems plausible that model resolution could change the distributions of CHBr3 in the atmosphere more significantly for the a posteriori emissions given their greater heterogeneity. I recommend that the authors test this in a separate sensitivity study and present their conclusions.***

Good suggestion. We have included such a comparison the revised manuscript (at revised Figure 2), which confirms that the posterior nested simulation at  $0.25^\circ \times 0.3125^\circ$  is very similar to the run at  $2^\circ \times 2.5^\circ$ .

***6) It isn't clear to me that the mean bias between the mole fractions of observed and modelled CH2Br2 decrease from the a priori to the a posteriori simulations. The paper states this, but as it is written the bias changes from 0.01 +/- 0.14 to -0.1 +/- 0.1. Can the authors please explain this result? Is this due to an overcompensation in the a posteriori emissions close the well observed region? According to the forward model section, a large fraction of the CH2Br2 originates from outside of the domain, and I imagine that in this case it is hard/impossible to infer those emissions with any reasonable specificity and overcompensation locally seems therefore to be a plausible explanation.***

We agree that it is due to an overcompensation caused by an uneven sensitivity (Jacobian) for measurements at different altitudes, and observation errors that are assumed to be proportional to mole fraction values by two campaigns. See Pages 10 and 11 for more discussions

***Technical comments:***

***Looking at Figure 4, it seems that the Ordonez et al and Ziska et al panels have been mislabelled in the caption whereby the Ziska emissions are described as being the Ordonez emissions and vice versa for the Ordonez emissions. Looking at Hossaini et al. (2013) ACP in figures 1 and 2 (and in fact the emission files themselves), I have checked the spatial patterns, and they seem to confirm this. Please can the authors check this themselves and confirm there is a mislabelling in the Fig. 4 caption? Please can the authors also check other instances of discussion of Ordonez and Ziska and verify that there a) there are no other mix-***

***ups in the naming and b) that this is just a technical naming error.***

We are grateful to the reader for spotting this error. We have checked and can confirm that it is just a plot labelling error.

***Russo et al. (2015, ACP) is included as a reference but is not cited. Please check for other articles referenced but not cited.***

Thanks. We have now checked the reference list and cite Russo et al in the revision (see above).