Response to Reviewer 1

We thank the reviewer for their careful comments on the paper. We respond below to the specific comments made.

1. *The paper recognizes that issues with the measurement of the background current exist, but it still treats this quantity as a well-defined and well-measured quantity, despite the problems encountered during the experiment. The need to change solutions multiple times to decrease the background is only one indication that this quantity is not well defined. It would be very helpful, if the authors added a conservative estimate of the uncertainty of the background current encountered in the upper troposphere and use this uncertainty estimate in their comparisons. While some of their measurements may be consistent with uplift of ozone poor air from the boundary layer, I would expect that the uncertainty is sufficient that uplift of free tropospheric air with higher ozone concentrations cannot be excluded. I would therefore urge the authors to improve the discussion of their uncertainties, which has important implications on the interpretation*

   Estimating the accuracy of the background current correction, as the reviewer acknowledges, is very difficult. We agree with the statement that the background current is not well defined, and will make this point earlier in the paper. The choice of correction method is therefore to some extent arbitrary. Nevertheless, making no correction is not an option. In the absence of a clear, well-understood methodology to correct for the background current, our approach to uncertainty was to examine as far as possible the internal consistency of the dataset, exploiting the fact that the sondes were all from the same batch and were all launched in a 23 day period. Looking at the uncontaminated sondes (nos. 15 onwards) we noted in the paper that

   (i) most of the measured background currents fell in the range 40 – 60 nA, i.e. within 10 nA of 50 nA, showing that the background current was reproducible to this accuracy from sonde to sonde

   (ii) the variation in background current between the beginning and end of the second preparation was again within 10 nA – showing that the background current was stable during preparation, again to this accuracy.

   We conclude that we could measure a stable and reproducible background current during preparation with an uncertainty of ±10 nA. Using this figure and assuming the background current is constant during an ascent we arrive at the estimate of ±3 ppbv for ozone in the TTL given in the paper. The next question is what changes occur in the background current during flight. Thornton and Niazy (1983) conducted laboratory experiments with ECC sondes suggesting that I_{bg} should be constant up to 100 mb, then decline logarithmically with pressure. In these experiments (unlike those of Vömel and Diaz (2010)) the sonde was not exposed to ozone at any point, and so were more representative of our preparations than Vömel and Diaz's. Our laboratory investigations on an uncontaminated sonde (fig A3) suggest a small decrease of around 5 nA in going from lab pressure to 100 mb, consistent with Thornton and Niazy's result within error limits. Taking this as an uncertainty (rather than a bias) in the variation of I_{bg} we estimate the uncertainty in TTL ozone below 100 mb to be ±5 ppbv.

   We have two other sources of information about the possible change in the background current. One is the comparison with the aircraft, where sonde and aircraft were consistent to within 3 ppbv, and the other is the consistency of the TTL ozone measurements from sonde to sonde. Figs 5 and 6 show that assumption of a constant background current for a 'clean' sonde leads to agreement with the aircraft within 3 ppbv, and the repeated measurement of minima in the range 12-14 ppbv in the TTL suggests consistency from sonde to sonde.

The referee contends that the TTL measurements are sufficiently uncertain that they could result from uplift of free tropospheric air. This is what we expected to see and it was a surprise to us that our TTL measurements were so low. Free tropospheric concentrations during the period of the lowest TTL ozone (19-23 Feb) were in excess of 20 ppbv up to 12 km (Fig 11). For the TTL ozone value to be in error by 9 ppbv, the error in background current would need to be around 30 nA, i.e. the actual background current in the TTL would be around 20 nA rather than 50 nA – close to the 25 nA that the Vömel and Diaz (2012) correction would imply and well below anything measured on the ground. We can see the effect of using Vömel and Diaz's correction in figures 5 and 6, where the resulting ozone profile is 5 ppbv higher than the aircraft measurements at 150 mb; the profile is clearly not consistent with the aircraft. There is no evidence here of a significant decrease in the background current with altitude.

Of course we cannot verify with this dataset whether the air that entered the deep convective complex over the ocean east of Manus contained the same low-level ozone values as those measured by the sonde, and so the hypothesis of uplift of near-surface air to the tropopause remains just that, a hypothesis.

While this paper was in review, results from the CONTRAST project were published in GRL. Our minimum TTL ozone of 12-13 ppbv is entirely consistent with the minimum value of 13 ppbv measured by the Gulfstream V during CONTRAST. Some additional text has been added to the paper (in blue) to include the CONTRAST findings. Text added in red are in response to reviewers' comments.

We have added the following text to the end of 2.1:

It is clear from previous work that the background current is not a well-defined quantity, and that there is uncertainty on the best way to measure it and its possible variation during flight. This is acknowledged by the Global Atmospheric Watch (GAW) report on ozonesondes (Smit et al, 2013), which calls for more fundamental research on this topic. We now describe in detail the ozonesonde preparation method in Manus, which departed from GAW standard procedures in a number of ways.

and the following to the end of section 2:

The spread in measured background current for the uncontaminated sondes was around 10 nA (0.01 µA, figure 2, sondes 15 onwards), with a similar difference between the values measured at the beginning and the end of the preparation, so it is reasonable to estimate an uncertainty in I_{bg} measured before flight of ~10 nA. If I_{bg} were constant during flight this would correspond to an uncertainty of ~3.4 ppbv in the TTL. According to Thornton and Niazy (1983) I_{bg} should remain constant up to 100 mb, then decline logarithmically with pressure. Our laboratory investigations on an uncontaminated sonde (fig A3) suggest a small decrease of around 5 nA in going from lab pressure to 100 mb, consistent with Thornton and Niazy's result within error limits. Taking this as an uncertainty (rather than a bias) in the variation of I_{bg} we estimate the uncertainty in TTL ozone below 100 mb to be ±5 ppbv. The cold-point tropopause during the campaign at Manus was always between 90 and 110 mb, with the ozone concentration increasing rapidly in this range: the minimum concentration was always found below 110 mb. Above 100 mb the use of a constant I_{bg} will tend to lead to an underestimate of ozone, but as ozone was generally > 50 ppbv above 100 mb, and increasing rapidly with height, this effect is only manifest in the stratosphere.

2. *The authors prepared their sondes not following standard recommendations by GAW. I can support the deviation of these standard recommendations, but the authors should try to comment on the impact of this deviation to other studies.*

We have added the following text (also responding to reviewer 2's comments):

The procedures used in Manus departed, as already mentioned, from the GAW recommendations. The most important deviation (a consequence of the malfunctioning calibration unit, see below) was that the majority of sondes were not exposed to ozone during preparation. This turns out to have been advantageous, as it avoided the decay in I_{bg} reported by Vömel and Diaz (2010). Smit et al. (2007) report that the background current measured 10 minutes after exposure to ozone in the final preparation exceeded that measured before exposure to ozone by 34 nA on average for a sample of five EnSci sondes. By contrast, for the uncontaminated sondes in Manus the average difference in I_{bg} measured at the beginning and end of the final preparation was only 6 nA (Figure 2). Together with changes in solution to ensure that I_{bg} fell to around 50 nA, not exposing the cell to ozone resulted in a stable I_{bg} during preparation, lending confidence to the subsequent assumption that it remained constant during flight. We examine this assumption further in the next section.

Other departures from GAW recommendations were:
- – the use of a 1% solution rather the 0.5% which leads to an oversensitivity to ozone and a bias of ~ +5% in ozone concentration (WMO, 2013)
- – measurement of T_{box} rather than the pump temperature, leading to an underestimate of ozone by ~ 3% since the pump temperature is higher by around 10°C (WMO, 2013)
- – use of a charcoal filter to provide ozone-free air rather than an ozone-free gas supply. The effect of this is difficult to quantify, but will be most serious in a laboratory with humid air and measurable concentrations of ozone. In this case the relative humidity of cabin air was around 50%, within the expected operational range of the filter. On occasion a sonde was allowed to sample laboratory air without the filter attached, but this made no difference to the measured current. This means either that the laboratory was essentially ozone-free or that the filter was not working. When the sonde was taken outside and the filter removed, an increase in signal was measured, so we conclude that the filter was working correctly and that laboratory air was essentially ozone-free.
- – correction to pump flow rate measurement for humidification of air. For a laboratory at 20°C and 50% RH this correction reduces F in equation 4 by around 1.5% (WMO, 2013), increasing ozone by the same amount – in other words equation 4 underestimates ozone by ~ 1.5%.

The overall effect of departures from the GAW recommendations is therefore small – much smaller than the error due to the background current uncertainty for tropical tropospheric ozone concentrations.

3. *Vömel and Diaz (2010) reported on difficulties using ozone destruction filters in tropical regions. The authors should therefore comment on the possibility of incomplete ozone destruction in the filters used during their experiment and possible impacts on their measurements.*

Please see response to point 2

4. *The authors use the outdated box temperature measurement instead of the pump temperature measurement, which is the current standard for ECC's. This may be of particular importance in the coldest parts of the atmosphere, i.e. the tropopause region and the authors should comment on this.*

   Please see response to point 2

5. *The ECC equation also contains a pump efficiency, which is not shown in equation 1. This factor largely plays a role at lower pressures than those studied here, but it would be good to know, which pump efficiency correction was used and which value was used in the upper troposphere.*

   The Vaisala software uses the pump correction from Komhyr et al (1995) for EnSci sondes. As the reviewer says, this correction is negligible in the troposphere. We have added after equation 1:

   A pump correction following Komhyr et al (1995) was also applied to the data but this is negligible for the altitude range considered in this paper.

6. *The authors need to point out that their empirical hybrid correction may only apply to their particular soundings. Since the source and mechanism of contamination was not clearly established it can only be stated, that this approach may work for this experiment and may not be a general result that applies to any other campaign. Furthermore, this empirical correction strongly impacts the uncertainty of the affected measurements.*

   We do point this out in the last paragraph of section 3, but then go on to examine whether there is any possible wider validity to the method. We have strengthened the caveat in the first sentence and added some text to the end of section 2 discussing the uncertainty of the contaminated sondes – indeed the empirical correction makes their uncertainty greater and we didn't discuss that properly. We were truly surprised to find such good agreement between sonde 6 and the aircraft profile however.

The final paragraph of section 2 now reads:

   The error in TTL ozone for the contaminated sondes cannot be assessed quantitatively but will certainly be greater than that for the uncontaminated sondes. We can only get an estimate of this error from a comparison with another technique, so we now turn to a comparison of ozonesonde profiles with aircraft measurements.

7. *Unfortunately only soundings #34 and #35 can serve as true comparisons with the aircraft measurements; therefore, the statistics of aircraft validations in not overwhelming. A better discussion of the uncertainties and their significance may help in the interpretation.*

   See 1 above. We argue that there is such a high degree of internal consistency between sonde profiles that the aircraft comparisons are transferrable to the dataset as a whole.

8. *The authors should elaborate more on the bell jar measurements. What is their source of air inside the bell jar? Are they just recycling air? Can they exclude any additional impacts from the bell jar?*

   Yes we are just recycling air. While we can envisage ozone being destroyed in the bell jar, it is very difficult to see how it can be created. We have modified the beginning of A3 to the following:

The effect of lowering the ambient pressure on the contamination was then investigated by placing the ozonesonde in a bell jar and lowering the pressure as the sonde continually sampled the air inside the bell jar. The bell jar was too small to admit the ozone destruction filter but ozone measurements inside the jar at ambient pressure were the same as in the laboratory with the filter attached; thus air in the bell jar was ozone-free.

9. *Figure A1 and Appendix A3: The authors clearly state that the background of the contaminated sondes decays with time. Therefore, a pressure dependence is somewhat misleading, even though it may be the more practical approach to apply the correction. The large scatter in the background measurements as function of temperature indicates that there is significant uncertainty in this correction. This should be described.*

   We do make the point that a pressure dependence in the hybrid formula is a convenient way of applying a time-dependent term. Fig A1 does indeed show that the background current is very uncertain for a contaminated sonde, which is why we discarded the data from the first two sondes in Manus.

   For the slightly contaminated sondes, fig A2 shows a general decrease in I_{bg} over time, again with uncertainties. This simply serves to show that the hybrid correction is consistent with the behaviour of the contaminant. We do not believe that a quantitative estimate of the error in I_{bg} is possible for these cases, and now say this explicitly in the text.

   We did not vary the temperature during these experiments – they were all conducted at room temperature.