We would like to thank the reviewer for his/her helpful comments which are shown in blue below. Our replies are in green.

This study provides a time series analysis by merging TES and IASI measurements over a ten-year period. The manuscript is well written and it is an interesting work which investigates a dedicated methodology for homogenizing two different datasets in order to study in the future long-term trends in tropospheric ozone. However, I do have one major concern related to the results presented in figures 7 and 8, which should be addressed before final publication.

Major comments:

My biggest concern is related to the time series you show in figures 7 and 8 (top panels for Eastern Asia). I do have doubts about the methodology you use when looking at the sharp decrease you obtain after 2011 which seems quite unrealistic and coincident with the change of TES observing strategy. Is it possible that the random distribution of IASI scenes in spring 2011 is biased due to reasons you evoke in p.31034, l.25? Did you over-sample the IASI data over a specific period of the month or part of the region, which could possibly explain such a change in the time series? I'm wandering to what extent the different sub-samples might affect the results. I would recommend to provide some more sample testing in order to get a feeling for the robustness of the time series.

Note that the fact that the change of TES observing strategy has no bearing on the sampling of IASI points in the regions of interest, since the IASI points used to construct the ROI time series are not co-located with TES points. The random distribution of IASI scenes in spring 2011 is not biased. We took care to avoid this situation. We have now provided additional clarification on this point in the manuscript, by adding/changing the following sentences:

Data gaps in the time series occur for several reasons. A data point for a whole month is removed if an instrument has missing data for more than a week for that given month or if the whole area of the ROI is not completely covered by an instrument due to cloud cover or missing orbits. Missing orbits can lead to biased sampling of the random number generator within the ROI of a given month. However, the data has been screened for this by looking at the number of satellite scenes per day used for calculating the monthly mean: If the number of satellite scenes for 3 or more consecutive days is twice as high as the average for the rest of the month, the monthly data points are removed as well. For all data points included in the time series, care was taken to ensure that the initial distribution is unbiased.

In terms of the realism of the sharp drop, we have now added a figure showing ozone from sonde measurements from Hilo, Hawaii. The Hilo ozone sondes are a completely independent dataset from the IASI satellite-based ozone estimates. We note that a sharp drop is also observed in the sonde observations at that particular location, which is strongly influenced by outflow of free tropospheric air from Asia.

p. 31035, l.1-3: You discuss the results considering the confidence limit, but not the standard deviation while this metric is independent of the sample size. I would suggest to represent the standard deviation associated with the monthly mean on Fig.7 and Fig.8 (error bars). If the std is larger than the sharp drop-off in 2011 or larger than ~5ppb which corresponds to the trend approximated over 2011-2014, it means that the decrease in 2011 or the increase since 2011 in Fig.8 might not be significant.

The standard deviation on the monthly means includes natural variability within a given month, within the domain of the ROI. The standard deviation on the monthly means is around 15 ppb. We agree that this number should be stated and have now added this information to the text.

15 ppb is indeed larger than the magnitude of the sharp drop off. However, we do not agree that the standard deviation is the most appropriate metric to assess the significance of the sharp drop. The significance of the calculated mean is clearly dependent on the sample size in this case, and therefore we maintain that the confidence limit is hence a more appropriate metric.

In addition, I'm also wandering in what way the offset values determined from the global measurement is suitable for the ROIs. If relevant, I would also recommend to provide IASI-TES frequency distribution panels for the ROIs.

Since the IASI and TES measurements for the ROIs presented in Figures 7 and 8 are not collocated, calculating a frequency distribution of the differences for the ROIs is not possible. The rationale for not co-locating the measurements for the ROIs is stated on p. 31033, l. 7-13.

We added to the manuscript a discussion about the width of the frequency distribution in Fig. 3. The width of the distribution can be explained with the precision of the instruments and the spatial and temporal colocation error only. The globally derived offset is valid for the ROIs since the most significant settings for the retrievals are identical (algorithm, a priori profiles, covariance matrices, see p. 31029, l. 21-23) and the offset is determined by the instrumental spectral resolution, i.e. weighting functions only. However, we added a table with Gaussian fit parameters for the global surveys split into individual seasons and latitude bands to the discussion in section 3.2. See more details in the reply to your specific comment below.

Furthermore, the results presented in Fig.7 and 8 which are the most important results of the paper suffer from the lack of analysis and discussions with previous studies. It is quite frustrating to read here "Analysis and attribution of the ozone changes over Eastern Asia is under investigation in a follow-up study". Impact of drivers of tropospheric O3 variations, of East Asian monsoon,… could be discussed and numerous previous studies should be cited.

The focus of this paper is the demonstration of the creation of a joint TES and IASI data record. We have taken care to combine the data records from the two instruments in a way such that the combined record is meaningful. The time series are indeed intriguing, but we hope that the reviewer will appreciate that an in-depth analysis of the ozone record is beyond the scope of this paper.

In order to address the reviewer's comment, we have removed the sentence beginning *Analysis and attribution…* and have added the following text to the end of that same paragraph:

For any given region, long-term variations in free tropospheric ozone can be affected by changes in local emissions of ozone precursors, changes in long-range transport within the troposphere and downward transport from the stratosphere (see e.g. Lin et al., 2012, 2014). The combined TES/IASI time series presented here has the potential to be used to aid attribution of the relative contributions from these effects, although such a study is outside the scope of this paper.

Specific comments:

p. 31031, l.5-8: It is not clear for me from panel b that the TES ozone profile shape deviates more from the a priori than the IASI one. I could even suspect the contrary.

Instead of showing TES and IASI profiles for all global surveys which mask the differences, would it be more appropriate to show one example for one individual survey or for one of the selected regions of interest (e.g. eastern Asia)?

While the mean TES profile values are closer to the mean prior values, the mean TES profile shows a steeper s-shape than IASI or the a priori.

Following suggestions of reviewer 2 and to emphasise the point we are trying to bring across we added to the discussion of Fig. 1 the following text:

Despite the fact that the ozone profiles themselves vary significantly with latitude and season owing to variations in tropopause altitude (Fig. 1B), the differences between IASI and TES ozone (Fig. 1D) are relatively consistent in shape and magnitude across seasons and latitudes, with the IASI-TES differences showing a standard deviation of only ~20 % in the upper troposphere for the whole dataset, compared to the extremely large standard deviations in the ozone upper tropospheric volume mixing ratio.

We also added the August and the November global survey example overviews to the supplementary material:

Two examples for the individual GS overviews are given in the supplementary material in Fig. S1 and Fig. S2 for August and November, respectively.

p. 31031, l. 22-23: Could you here mention the DOFS you obtain for the TES and IASI retrievals in that altitude range?

We moved the sentence: *The degrees of freedom for signal for both TES and IASI for the considered altitude range are between 0.7 and 0.8*. from p. 31034, l. 11-12 to p. 31031, l. 22-23.

p. 31031, l.26-27: This is not clear to me. Do you mean here the sum of the "rows" of averaging kernel matrix in the relevant pressure range: : :? But then how do you obtain one single value (x axis in figure 2) and not a sensitivity profile with 5 levels corresponding to the 5 retrieval vertical grid points in the selected pressure range? I do not understand what represents the "IASI sum of AK matrix" values you plot in fig.2. Please clarify.

We extended the text to read as follows:

The sum of the averaging kernel matrix in the relevant pressure range is used as a measure of the sensitivity. For this we calculate the sum of the rows of the AK matrix for each of the retrieval levels in the specified range and then add those together. The peak values of the TES AKs are larger than IASI's; however, the FWHM of the TES AK is narrower. Since we are averaging ozone over a range of several retrieval levels, the total area under the AKs is more representative of the information content than just the sum of the peak values.

p. 31032, l. 3-4: Figures 2 is quite hard to visualize and looking at that figure, it even seems that the IASI-TES differences are dependent on the ozone values with larger differences corresponding to large amount, while I agree that the differences seem independent on what you called here the IASI sensitivity.

We acknowledge that the choice of size-dependant markers was not ideal since it draws the eye to the outliers. We changed the figure to have the ozone amount marked with different colours instead.

p.31032, l.8: Some words of caution related to the impact of using different external parameters for TES and IASI retrievals on the IASI-TES differences should be added here.

We added to the text: *Apart from instrument specifications, the sensitivity of infrared instruments towards ozone depends on the atmospheric and surface temperatures, water vapour amount, residual cloud contamination, surface emissivity, and the amount of ozone itself. The uncertainties in the IASI ozone profile from the water vapour uncertainty has been* *estimated to be less than 2 % and from the temperature profile to be less than 5 % (Oetjen et al., 2014). The uncertainty from the surface temperature is negligible. Collocated retrievals, as considered here, should all be affected in a similar way by these external parameters.*

p.31032, l.11-12: Do you obtain the same offset (3.9 ppb) when looking at different regions or at different time periods? What is the offset for the measurements over the three regions of interest? I would suggest to show the frequency distribution for the regions of interest in addition to the one you present for the global surveys, to get a feeling of the impact of the differences on the homogenized time series.

We did look at different latitude bands and seasons in the global survey dataset and we added a table with the parameters for the frequency distributions for those to the revised manuscript. Overall, we greatly expanded the discussion in section 3.2 for Fig. 3:

The normalised frequency distribution of the offset of the data of Fig. 2 is shown in Fig. 3. The distribution of the difference between TES and IASI-TOE follows roughly a Gaussian function with the maximum at (−3.9 ± 0.1) ppb (see Tab. 1, last row). For merging the TES and IASI-TOE data series, only the location of the peak value is important. The width of the frequency distribution is 17.6 ppb and is determined by the precision of the measurements and the colocation error: The precision for the IASI-TOE retrieval was estimated to be better than 20 % (see Sect 3.1). Those 20 % were calculated from comparison with ozone sondes with a coincidence criterion of also 55km (Oetjen et al., 2014). Hence a possible spatial colocation error is included in this estimate. This translates into 10.4 ppb for a mean IASI-TOE ozone of 51.9 ppb for the IASI precision plus spatial colocation. The TES precision of 15 % translates to 8.3 ppb for 55.4 ppb mean ozone. To estimate the temporal colocation error we compared model fields (GEOS-Chem version 10.1 (Bey et al., 2001; Eastham et al., 2014)) for the dates of the 4 global surveys for the overpass times of IASI and TES and calculated the standard deviation of the ozone difference which is 1.7 ppb. Adding the IASI combined precision and spatial colocation estimate, the TES precision and the temporal colocation estimate in quadrature gives ~14 ppb, which is slightly smaller than, but not dramatically different from the FWHM in Fig. 3.

Table 1 gives an overview for the Gaussian fit parameters for frequency distributions of selected sub-sets of the GSs separated by season or by latitude regions. Included are the location of the maxima and the FWHM of the Gaussian fit as well as the correlation coefficient R^2 for the quality of the fit to the data. In general the results are variable because of the large differences in sample size. However, when only considering distributions with an R^2 larger than 0.95 (sub-samples for summer, winter, northern midlatitudes, and tropics), the peak values fall in the range of -3.4 to -4.9 ppb. This gives confidence in using a global offset *of -3.9 ppb to combine TES and IASI average ozone mixing ratios in the chosen pressure range.*

For the time series in the three regions of interests, we did not attempt to co-locate TES and IASI points (for the reasons discussed in the text). Therefore we do not have the frequency distribution for the ROIs.

p. 31034, l.15-16: Is the offset value of 3.9 ppb which has been determined for some specific short period at a global scale suitable for the three regions of interest?

This was determined using one global survey for each season over the course of 12 months, so the number is more representative than it would be for a single specific short period.

p. 31034, l. 21: Did you use a cloud fraction of 13% or 6% as previously mentioned in p. 1030, l.20?

We changed the sentence on p. 31034 to:

In this part of the study, we relax the cloud screening thresholds to 2.0 for the TES average cloud optical depth and to 13 % cloud fraction for IASI scenes which are more widely used thresholds (e.g. Clerbaux et al. 2009).

Technical corrections:

p.31032, l.16: I guess you mean “IASI-TES” differences, not IASI-TOE.

IASI-TOE: IASI - TES Optimal Estimation (see p. 31030, l. 2-3)

p.31033, l.26: “area the size” ! “area of the size”, same in the figure 5 caption.

Changed.