Reply to the Interactive Comment Atmos. Meas. Tech. Discuss., 8, C5308–C5311, 2016

Major comments:

1) The methodology of the validation is not well described in the manuscript. Authors refer readers to another manuscript for the details [Sepúlveda et al. (2014)]. However, in order to apprehend results presented in this manuscript it is crucial to understand the methods used for the analysis. The methodology should be carefully described in Section 4 "Comparison strategy". Below I listed some specific comments related to that:

p. 13740, line 7: Please, explain your motivation for using a logarithmic scale versus linear (e.g. anomalies as % from the mean);

p. 13740, lines10-11: Please, include the equation for the statistical fit model used to remove "the trend and the intra-annual variations";

p. 13745, lines 1-5 (also Figure 7): In the text authors talk about "long-term trends" and correlations, but the method to compute these trends and correlations has never been explained. Did you compute those long-term time series by fitting linear trends in weekly averaged, de-trended, de-seasonalised residuals? Please, carefully describe your approach in Section 4.

Figure 5: The bottom panel shows median biases in %. How were the median biases computed? Do these biases account for differences in annual cycles between the sensors and overpasses? Please, explain your approach in Section 4.

A detailed explanation of the comparison methodology was not included in the submitted paper to avoid repeating results already published in previous works. But, we agree with the referee in that a complete description of the “Comparison Strategy” used in this work is key for understanding the comparison results. Therefore, we have extended the section 4 including all the Referee’s suggestions as follows (the new text is highlighted in italic):

The consistency and quality assessment of IASI-A and IASI-B products is addressed at different time scales: single measurements, daily, annual, and long-term trends. This temporal decomposition provides an added value for validating trace gases with a rather small variability, such as N_2O or CO_2. For such gases the uncertainty is often larger than the day-to-day concentration variations and thus a validation at longer temporal scales is more meaningful than a validation limited to a comparison of individual measurements. Moreover, this analysis allows us to quickly detect instrumental issues or inconsistencies.
For this purpose, we follow the procedure proposed by Sepúlveda et al. (2014) (and references therein) *and explained in detail in the following. Firstly, for analysing the time series on different time scales the measured total column time series of each target gas ([TC_{gas}]) is fitted to a time series model, which considers a mean [TC_{gas}] value and [TC_{gas}] variations on two different time scales (see equation 1): a linear trend and intra-annual variations:*

$$F(t) = f_o + f_{trend} + \sum_{i=1}^{p}[a_i \cos(w_i t) + b_i sen(w_i t)] \qquad (1)$$

where t is measured in years, f_o is a baseline constant and f_{trend} the linear trend in change per year. The annual cycle is modeled in terms of a Fourier series where a_i and b_i are the parameters of the Fourier series to be determined and $w_i = 2\pi i/T$ with T=365.25 days. Once the model fit is computed, the seasonal variations are obtained by subtracting the fitted linear trends from the measured time series.

The averaged annual cycle is then computed by averaging these de-trended time series on a monthly basis. It represents the de-trended multi-annual seasonal cycle of the target gas. In addition to the seasonal time scale we look on measurement-to-measurement and long-term time scales. For the separation into these two time scales we use the aforementioned time series model. The measurement-to-measurement time scale signal is calculated as the difference between the measured time series and the modeled time series (whereby all fitted time scales are considered: mean value, linear trend, and seasonal cycle). The so-calculated de-trended and de-seasonalised time series represents the very short-term variations, corresponding to the variations among individual observations. Finally, in order to calculate the long-term time scale signal (annual means) we reconstructed a time series that only considers the fit results obtained for the mean $[TC_{gas}]$ and the seasonal cycle. Then, by subtracting it from the measured time series, we get a de-seasonalised time series, for which we then calculate the annual mean values. Note that for IASI-A and IASI-B consistency study, the bias between both IASI sensors is also calculated. To do so, we directly compare the measured total columns of all the trace gases and compute the median difference of this difference time series.
This temporal decomposition has been done on a logarithmic scale, i.e., our measured time series correspond to the logarithm of the measured total columns of all the trace gases. This approach has two clear advantages in the subsequent IASI-FTS comparison: (i) the $[TC]_{gas}$ variations on this scale can be interpreted as variations relative to the reference mean values ($\Delta \ln[TC]_{gas} \cong \Delta[TC]_{gas}/[TC]_{gas}$), thereby we directly compare the anomalies observed by both remote sensing instruments, and (ii) the relative differences between IASI and FTS observations can directly be computed as the subtraction of the corresponding variabilities on the different time scales (note that the temporal decomposition produces values very close to zero, thereby computing the standard relative differences provides very extreme values in some cases).

2) <u>Another major comment</u> is related to the analysis of the theoretical errors for the FTS retrievals presented in Appendix A. Authors followed Rodgers formalism to estimate these errors. Authors defined a covariance matrix Sa (see p. 13750, line 10), used to estimate the smoothing error, as "the assumed a priori covariance matrix". However, according to Rodgers (see p. 49 in his book, [Rodgers, 2000]) in order "to estimate the smoothing error, the covariance matrix of a real ensemble of states must be known". He further emphasizes: "To estimate it correctly, the actual statistics of the fine structure must be known. It is not enough to simply use some ad hoc matrix that has been constructed as a reasonable a priori constraint in the retrieval. If the real covariance is not available, it may be better to abandon the estimation of the smoothing error...". First of all, authors have to change a definition for the covariance matrix Sa and use Rodgers definition. Secondly, authors need to justify the use of WACCM model outputs for constructing the covariance matrices for considered atmospheric species. It is not clear for the referee how well the WACCM simulations represent the real atmospheric states: fine vertical structures, inter-level correlations. It would be nice if authors can provide references on works that show ability of the WACCM model to reasonably simulate vertical distribution of gases in comparison with sonde, lidar or any other high-resolution measurements. At the very least, authors have to clearly identify in the text that they use "assumed" covariance matrices due to lack of real observations. In this case obtained error estimates should be also considered and treated as "assumed" smoothing errors. Finally, authors stated that the smoothing error have only statistical component, which is incorrect. The smoothing error represents the error caused by a limited vertical resolution of the observing system, thus it has very pronounced features defined by the instrumental averaging kernels and natural variability of the considered atmospheric gas. Moreover, any purely statistical errors will be cancelled out by averaging a large number of observations, which is not a case for the smoothing error.

Following the referee's suggestions, we have changed the description of the smoothing error calculation, in the first part of Appendix A, as follows (the new text is highlighted in italic):

Theoretically, the error of the different FTS products can be estimated by following the formalism detailed by Rodgers (2000), where the difference between the retrieved state, $\hat{x}$, and the real state, x, can be written as a linear combination of the a priori state, x_a, the real and estimated model parameters, b and $\hat{b}$, respectively, and the measurement noise $\in$:

$$(\hat{x} - x) = (A-I)(x - x_a) + GK_b(b - \hat{b}) + G\in \quad (A1)$$

where G represents the gain matrix, Kb a sensitivity matrix to model parameters, *I the identify matrix*, and A the averaging kernel matrix. A relates the real variability to the measured variability of the considered atmospheric state and, thus, represents the way in which the remote sensing system smoothes the real vertical profiles (Rodgers, 2000). Thereby, Eq. (A1) defines three types of error: the first term is the smoothing error associated with the limited vertical sensitivity of the FTS instruments, the second one represents the errors due to uncertainties in the input/model parameters (instrumental characteristics, spectroscopy data, …) and the third one corresponds to the measurement noise.

The theoretical error estimation strongly depends on the assumed uncertainties. In our case, we consider the error sources and values listed in Table A1 for the input parameters, which are very realistic estimations coming from our experience and the literature (Schneider and Hase, 2008; Sepúlveda et al., 2014, and references therein), while the smoothing error is calculated as $(A-I)S_a(A-I)^T$, where *S_a is the covariance matrix of the target gas. Strictly, to estimate the smoothing error contribution, the covariance matrix of a real ensemble of atmospheric states must be known (Rodgers, 2000). However, due to lack of real observations of the vertical profiles of all the trace gas considered at IZO, the S_a for each target gas is assumed and calculated from the WACCM-V6 model estimates. WACCM is a global chemistry model of well-recognized prestigious that has widely demonstrated its ability to provide reliable estimations of the vertical profiles of trace gases and their expected concentration variations (Pan et al., 2006; SPARC CCMVal, 2010; Smith et al., 2011; Brakebusch et al., 2013). Therefore, here the Sa is calculated considering the variance of the corresponding gas concentrations at each altitude from the WACCM-V6 climatological data and a Gaussian distribution of strength 5km for the inter-layer correlation.* Note that the total error values are calculated as the root sum-squares of all the error sources considered, where the contribution of each error source has been split into statistical and systematic contributions. *The exceptions are the spectroscopic parameters and the measurement noise, which are considered as purely systematic and statistical, respectively.* This error estimation has been applied to the IZO FTS observations between 2010–2014 (period studied in the current work).

References

Brakebusch, M., C. E. Randall, D. E. Kinnison, S. Tilmes, M. L. Santee, and G. L. Manney (2013), Evaluation of Whole Atmosphere Community Climate Model simulations of ozone during Arctic winter 2004–2005, J. Geophys. Res. Atmos., 118, 2673–2688, doi:10.1002/jgrd.50226.

Pan, L.L., J.C. Wei, D. Kinnison, R.R. Garcia, D.J. Wuebbles, and G. Brasseur, 2007: A set of diagnostics for evaluating chemistry-climate models in the extratropical tropopause region, J. Geophys. Res. Atmos., 112, D09316, DOI: 10.1029/2006JD007792, 2006.

Smith, K.; Atlas, E. L.; Zhu, X.; Pope, L.; Lueb, R.; Moore, F. L.; Miller, B. R.; Montzka, S. A.; Elkins, J. W.; Nance, D.; Sweeney, C.; Pan, L.; Kinnison, D. E.; Hendershot, R.; Romashkin, P.; Wofsy, S. C.; Daube, B.; Kort, E. A.; Jimenez, R.; Pittman, J. V., Selected trace gas distributions and relationships: a comparison of HIAPER Pole to Pole Observations (HIPPO) and Whole Atmosphere Community Climate Model (WACCM), American Geophysical Union, Fall Meeting 2011, available at http://adsabs.harvard.edu/abs/2011AGUFM.A11K..02S, 2011.

SPARC CCMVal (2010), SPARC Report on the Evaluation of Chemistry-Climate Models, V. Eyring, T. G. Shepherd, D. W. Waugh (Eds.), SPARC Report No. 5, WCRP-132, WMO/TD-No. 1526. Available at http://www.atmosp.physics.utoronto.ca/SPARC, 2010.

Minor Comments:

p. 13732, line 13: It would be nice to specify in numbers (km) what is "excellent horizontal resolution".

The IASI horizontal resolution is of 12 km diameter at nadir. We will include this information in the introduction of the revised manuscript.

p. 13736, lines 7-14: Please, explain what does "a global a priori" and "single unique covariance matrix" mean. Does it mean that a priori information is independent of season? Do the a priori depend on latitude?

The optimal estimation method (OEM) used to generate the IASI operational temperature, humidity and ozone products in version 5 does not consider a priori information in terms of seasonality, trend, or geopraphical pattherns. This approach uses an unique global a priori covariance matrix, computed from a collection of ECMWF analysis records, independent on seasonal or latitudinal variations. Therey all the observed variability comes from IASI measurements rather than the a priori information.
We will include this claryfication in the section 2.2 of the revised manuscript.

p. 13741, line 20: It might be better to replace "the position of the spectrometer relative to the atmosphere" with, for example, "the geometry of observations".

This sentence will be changed following the Referee's suggestion.

p. 13743, lines 14-15: Please, consider to rephrase the sentence "Figure 3 shows ..."

The sentence "Figure 3 shows, for O3 for example, these analysed time series as well as the TC time series as observed by IASI-A and IASI-B." will changed by "Figure 3 shows the time series of O_3 TC as observed by IASI-A and IASI-B and the corresponding differences".

p. 13743, line 23: Do you see significant differences for the O3 distributions between evening and morning overpasses? If so, please specify that in the text.

When analysing the IASI-A and IASI-B consistency, effectively, we observe significant differences for the O_3 distributions between evening and morning overpasses. Thereby, we conclude that the IASI sensors are able to distinguish the O3 intra-day concentration variations. We will include this claryfication in the section 5 of the revised manuscript.

p. 13746, lines 9-13: Authors stated here that "IASI-FTS comparison also confirms the results observed for the consistency study of IASI-A and IASI-B sensors". It would be nice if you can be more specific and list similarities. In the following sentence it is stated that inter-comparison of IASI sensors could replace a validation against groundbased instruments. Two IASI sensors have similar systematic errors (instrumental, smoothing error etc.), and inter-comparisons will never reveal these types of errors. Only validation against independent observations can help to asses all type of errors. I would suggest to re-write this text.

The consistency analysis between both IASI sensors and the IASI-FTS comparison similarly reveal that IASI is able to capture the long-term trends of the all trace gas considered and the annual cycles of all of them, expect for CO_2. While at shorter time scale (single or daily observations), only O_3 and CO are

moderately measured. Indeed, the correlations and the scatter of the differences observed for both analyses are very consistent (recall Figure 5 and 7 of the paper). For example, the correlations among sensors and overpasses, on a measurement-to-measurement basis, are ~0.8 for O_3 and significantly lower for the rest of the trace gases (~0.4). For IASI-FTS comparison we find similar values (Section 6 of the paper): ~0.8 for O_3 and lower than 0.3 for the rest of the trace gases. Similar conclusions can be obtained for the annual cycles, with the exception of CO_2. For this trace gas, the consistency study reveals a moderate agreement between both IASI sensors (correlations between 0.6-0.8), but we do not document any agreement for the IASI-FTS comparison (correlation less than 0.2). This is likely due to the degree of maturity of the IASI CO_2 products.
Following the referee's suggestion, we will change the following sentence "Note that IASI-FTS comparison also confirms the results observed for the consistency study of IASI-A and IASI-B sensors" by "Note that IASI-FTS comparison also confirms the results observed for the consistency study of IASI-A and IASI-B sensors. The correlations and the scatter of the differences observed for both analyses are very similar both at short-term and intra-annual time scales (recall figure 5 and 7) with the exception of CO_2. For this trace gas the consistency study reveals a moderate agreement between both IASI sensors (correlations between 0.6-0.8 for the annual cycles), but we do not document any agreement for the IASI-FTS comparison (correlation less than 0.2). This is likely due to the degree of maturity of the IASI CO_2 products as aforementioned."

We totally agree with the referee that only validation using high quality ground-based observations can be used to evaluate the quality and consistency of the satellite data. Indeed, this is the main objective of this paper. Nonetheless, here, we mean that the inter-comparison between both IASI sensors could help to identify, for example, instrumental issues presented in only one sensor (e.g., drift in the calibration, …), and only in lack of reference ground-based observations. Therefore, we consider that this sentence could be kept as it is in the revised manuscript.

p. 13754, line 11: What does "Gr" mean here? Please, define it.

"Gr" does not have a special meaning; It was just used as a generic name to explain how the standard deviation of a linear function of slope "Gr" is calculated.

p. 13765, Table 1: Does "Daily" mean +/- 12 hours or 24 hours?

Daily means 24-hour means. This will be clarified in the Table.

Figure 3, upper panel: Please, consider to change colors. It is difficult to see differences between light grey and dark grey symbols.

The colors of the different time series will be changed to order to make the differences clearer.