Anonymous Referee #1

The paper deals with an application of IASI data to the retrieval of N_2O. The authors analyzed IASI data for the Northern-Hemisphere summer season (June-July) in 2011 and claim that the relatively high concentration of N_2O over the Eastern Mediterranean basin is a result of pollution transport from Asia. I think that the paper shows potentially arguing and interesting results. However, an in-depth and accurate reading of the document shows, and I am sorry, that the paper presents a lot of weak points and many technical aspects, which need to be clarified and properly addressed before this study can be accepted for publication.

General Comment

It is known that the Mediterranean summer (June to September) is characterized by high pressure over the Mediterranean Europe and a low-pressure trough extending from the Persian Gulf through Iraq to the southeastern Mediterranean (see e.g., Y. Goldreich, Springer, 2003). It is now very well understood (e.g., Karnieli et al. JGR, 2009 and references therein) that this kind of weather pattern yields persistent northwesterly winds which causes long-range transport of air masses and pollutants from southeastern and southwestern Europe into the eastern Mediterranean basin. In agreement with this weather pattern, previous IASI studies (also cited in the present paper, e.g., doi:10.1364/OE.21.024753) have indeed evidenced higher concentrations of green-house gases over the Eastern Mediterranean basin. Conversely, the authors suggest that there could be another atmospheric pathway along which pollutants are transported to the Eastern Mediterranean basin. Because of the importance of this finding, the authors should be much more convincing in showing that their methodology has no weak points. In effect, their analysis is based on N_2O profiles retrieved with less than 1 degree of freedom, and they concentrate on N_2O layer average at ~309 hPa, but they fail to show that this layer average has been independently resolved of the rest of the profile. In view of the broad structure of N_2O AK they provide in the study, it is likely that they are mostly sensitive to the column amount of N_2O.

→ The transport process between the Asian surface and the eastern Mediterranean during summer monsoon has been already demonstrated in Kangah et al., (2017) using GOSAT N_2O retrievals, LMDz-Or-INCA chemical transport model and backtrajectories together with an extended literature. The aim of this paper is not to prove again this finding but to validate new IASI N_2O retrievals showing that these retrievals can capture this transport process at daily time scale (part 7 of the manuscript). Concerning the vertical resolution of the data refer to response #12 and #14.

I have detailed my points below, which I hope can help authors to improve the paper.

Major remarks

1. Page 4, line 17. NEDT depends on the scene temperature, which, because of atmospheric absorption, is wave number dependent. Was this dependence taken into account? By the way, I suggest that the comparison should be made in the radiance space, using NEDN, which is wave-number independent.

Spectral NEDT at 280 K is the reference value for IASI radiometric noise and is often used to check sensitivity of any retrieval to the measurement. Therefore, in our study, the NEDT is given at a reference temperature of 280 K. The incriminated sentence has been modified into:

> The IASI radiometric noise expressed as the Noise Equivalent Delta Temperature (NEDT) at the reference temperature of 280 K is superimposed to the $|\Delta BT|$ signals.

For NEDN, we do not agree with the statement of the reviewer saying that it is wavenumber independent. Considering Figure 1 from Amato et al. (1995), then it is obvious that NEDN is varying with wavenumber (see Figure R1).

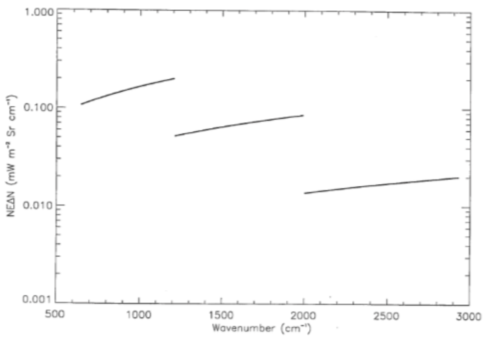

Figure R1: *IASI radiometric noise as a function of Wavenumber (taken from Amato et al., 1995).*

Amato, U., Cuomo, V., and Serio, C.: Assessing the impact of radiometric noise on IASI performances, Remote Sensing, 16(15), 2927-2938, 1995.

In addition, the most important here is the use of the same unit to compare sensitivity of the forward model to the geophysical parameters relative to the IASI radiometric noises.

2. Page 5, Equation (1). This equation should be introduced this way... *We used optimal estimation based on the Levenberg-Marquardt (put reference) algorithm as modified by Fletcher (put reference) and adapted for Optimal Estimation by Rodgers.* By the way, the important aspect here is that Equation (1), as it is written, is wrong. The term multiplying the leftmost S_a^{-1} should be $(1 + \gamma)$ not simply, γ. In fact, Eq (1) should transform back to the OE estimator for $\gamma = 0$, which is not the case. Hope this is just a typo. Furthermore, how γ is chosen at each step? Do authors perform retrieval in the BT space or radiance space? Please, clarify.

→ We verified the equation (1) and the reviewer is right, it is a typo. It should be $(1 + \gamma)$. We modified the equation accordingly.

$$\hat{X}_{i+1} = X_a + \left(K_i^T S_y^{-1} K_i + (1+\gamma) S_a^{-1}\right)^{-1} \times \left\{K_i^T S_y^{-1}\left(\left[Y - F(\hat{X}_i)\right] + K_i\left[\hat{X}_i - X_a\right]\right) + \gamma S_a^{-1}\left[\hat{X}_i - X_a\right]\right\} \quad (1)$$

We have modified the introduction of the equation (1) into:

We used optimal estimation based on the Levenberg-Marquardt (Levenberg, 1944; Marquardt, 1963) algorithm and adapted for Optimal Estimation by Rodgers (Rodgers, 2000)…

Note that we did not used the Fletcher strategy.
We have inserted the 2 following references in the revised version:

Levenberg, K.: A method for the solution of certain nonlinear problems in least squares, Quart. Appl. Math., 2, 164, 1944.
Marquardt, D. W.: An algorithm for least-squares estimation of nonlinear parameters, SIAM J. Appl. Math., 11, 431, 1963.

→The initial γ is initialised to 10. At each step, γ is updated as follows:

- If the cost function χ^2 (cf. response #8) decreases: γ is divided by 5 for the next step
- If the cost function χ^2 increases: γ is multiplied by 5 and both the cost function and the estimated profile ($\hat{X}_{i+1}$) are recalculated.

3. Page 5, Equation (2) is not consistent with Eq. (1). Apparently, the authors use γ = 1 for the final iterate, but then Equation (1) is not the correct OE estimator, and the final iterate would depart from optimality.

→ With the equation (1) correctly written, we have clarified the value of γ (= 0) for the final iteration into:

The vertical sensitivity of the retrieval can be characterised using the averaging kernel matrix (A) defined as (with γ = 0 for the final iteration):

4. Page 5, Equation (3) applies just to one parameter. Considering that the authors claim to use a simultaneous approach, how is the a priori covariance of the whole state vector built up?

→ The parameters are independent to each other in building up S_a thus the extra "block-diagonal" elements of S_a are fixed to 0 (there are no apriori correlation errors between the different state vector parameters). We have clarified this point by inserting the following sentence:

…The a priori error covariance matrix S_a is built for all chemical species and by considering parameters independent to each other as follows (cf. Rodgers, 2000):

$$S_{aij} = \sigma_a^2 \times \exp(-|\ln(P_i) - \ln(P_j)|) \quad (3)$$

where σ_a^2 is an a priori variance error fixed for each parameter of the state vector and P_i the pressure level at the level i.

Diagonal matrices are used for temperature profile and surface emissivity…

5. Page 5, line 20. Please show the N2O profile. The retrieval approach is strongly depending on such a background.

→ Done (cf. Figure 3)

6. Page six, line 2. As before, show the CO_2 profile and those of other species used as a priori reference.

→ Done (cf. Figure 3)

7. Page 6, line 13, what is a sink parameter? Please, explain.

→ We removed part of the sentence linked to the term "sink parameter" that is too much confusing.

8. Page 6, Equation (4). The denominator is wrong. The degrees of freedom of the χ^2 form are dim(Y). This can be demonstrated by a trivial use of the Standard Theorem of Least Squares (e.g., Rao 1973, the authors should consider that $\hat{X}$ is estimated from the data, so that the remaining degrees of freedom of data are dim(Y)-dim($\hat{X}$) and dim($\hat{X}$)+[dim(Y)-dim($\hat{X}$)]=dim(Y)). Since the authors use a retrieval algorithm for which dim($\hat{X}$)~dim(Y), the χ^2 is artificially decreased by a factor of almost 2.

→ Equation (4) refers to the normalized cost function χ^2_{norm} and not to the χ^2 test. We have rewritten the sentence in order to define (1) the cost function $\chi^2$and (2) the normalized cost function χ^2_{norm}.

> Our retrieval process consists in the minimization of the cost function χ^2 defined as:
>
> $$\chi^2 = [\hat{X} - X_a]^T S_a^{-1} [\hat{X} - X_a] + [Y - F(\hat{X})]^T S_y^{-1} [Y - F(\hat{X})] \quad (4)$$
>
> We used the normalized cost function χ^2_{norm} to evaluate the quality of the retrieval by combining the calculated residuals relative to the observations error covariance matrix and the difference between the estimated and the a priori profiles relative to the a priori error covariance matrix:
>
> $$\chi^2_{\text{norm}} = \frac{[\hat{X} - X_a]^T S_a^{-1} [\hat{X} - X_a] + [Y - F(\hat{X})]^T S_y^{-1} [Y - F(\hat{X})]}{\dim(\hat{X}) + \dim(Y)} \quad (5)$$
>
> where $\dim(\hat{X})$ and $\dim(Y)$ are the dimensions of the state vector and of the radiances (number of channels), respectively. In theory, χ^2_{norm} should be close to unity.

Figure R2 shows histograms of χ^2_{norm} for converged pixels of N_2O_B1 and N_2O_B2 over the region corresponding to figures 14 and 15 of the manuscript. χ^2_{norm} is globally higher for

N_2O_B1 than for N_2O_B2 confirming the difficulties to minimize the cost function in B1 compared to B2.

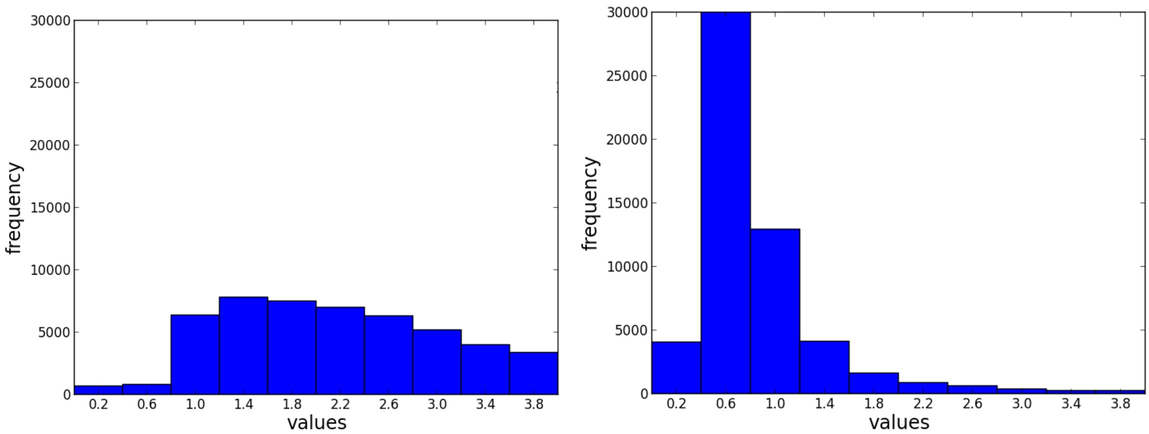

***Figure R2**:* χ^2_{norm} histograms for N_2O_B1 (left) and N_2O_B2(right).

9. Page 6, Equation (5) makes no sense in a Least Square retrieval approach, which seeks for a global minimum. Why one should look for a partial minimization, while using a simultaneous approach?

→ Equation (5) is not a partial minimization but another variable to assess the quality of the N_2O retrievals. In order to suppress the ambiguity with the cost function, we have redefined this variable Q_{N2O} for "quality of the N2O retrievals". We have modified the incriminated sentences accordingly. This parameter is a kind of normalized difference between The retrieval and the apriori N2O profile and is used to reject unrealistic N2O retrievals.

In addition to χ^2_{norm}, we computed another variable, $Q_{\mathrm{N_2O}}$, to assess the quality of the retrieved tropospheric N_2O profile which is our target species defined as the difference between the a priori and the retrieved N_2O relative to the N_2O a priori errors σ_a:

$$Q_{\mathrm{N_2O}} = \sum_{P_j<1000\,\mathrm{hPa}}^{P_j>200\,\mathrm{hPa}} \left[\hat{X}_j - X_{aj}\right]^2 \beta_{aj}/n_p \qquad (6)$$

where $\hat{X}_j$ and X_{aj} are the retrieved parameter and the a priori N_2O at the pressure level P_j, respectively. β_{aj} is the diagonal element of the a priori error precision matrix (the inverse of the a priori error covariance matrix) at the pressure level P_j and n_p is the number of levels used for the calculation.

An upper limit $\chi^2_{\mathrm{threshold}}$ for the χ^2_{norm} parameter is generally used to select good quality pixels. For instance, $\chi^2_{\mathrm{threshold}} = 3$ on a χ^2_{norm} calculated in the radiances space was used to select good quality pixels for CH_4 retrievals from IASI measurements (Xiong et al., 2013). Following the same methodology, we applied an upper limit $Q^{\mathrm{threshold}}_{\mathrm{N_2O}}$ on $Q_{\mathrm{N_2O}}$ to select good quality pixels. After performing sensitivity studies for both N_2O_B1 and N_2O_B2, we selected all the IASI pixels with $\chi^2_{\mathrm{norm}} \leq$

$\chi^2_{\text{threshold}}$ and $Q_{N_2O} \leq Q^{\text{threshold}}_{N_2O}$ with $\chi^2_{\text{threshold}} = 4$ and $Q^{\text{threshold}}_{N_2O} = 4$.

10. Page 7, lines 2 to 6. A χ^2 variable with $n = dim(Y)$ degrees of freedom has mean n and variance $2n$. Because for n large, the χ^2 distribution is approximately Gaussian, to compute a χ^2_{th} –tolerance limit, say within 3 standard deviations (or 3σ), we just need to calculate $\chi^2_{th} = n + 3\sqrt{2n}$. As an example, for n = 103, the number of channels the authors use in their B2 band, we have χ^2_{th} ~145, or $\frac{\chi^2_{th}}{n}$ ~ 1.42. Conversely, the authors use $\frac{\chi^2_{th}}{n}$ = 4, which in view of the factor 2 above (see point 8) increase to 8, which corresponds to a tolerance interval of 50 (fifty) standard deviation (sic!). With this convergence criterion, almost all retrievals are not converged!

→ As we previously explained (cf. #8 and #9), the parameter χ^2 is the cost function to be minimised and χ^2_{norm} is used as a quality parameter and not as a convergence criterion. However, the convergence criterion is performed by computing another parameter (d_2) which is roughly the distance between the next and the previous values of the forward model relative to the measurement errors covariance matrix:

$$d_2 = [Y_{i+1} - Y_i]^T S_y^{-1} [Y_{i+1} - Y_i]$$

Thus, we converge when d_2 is lower than the dimension of Y (namely the number of channels).

11. Page 7, Equation (6). I do not like the use of this empirical *Contamination Factor*. Why the final solution should be contaminated? They use a simultaneous retrieval. Averaging Kernels are good to assess vertical resolution. To check the interdependency of the retrieved state vector the authors have the a-posteriori covariance matrix. Please, use this matrix and compute the correlation matrix. In case the N_2O profile has not been independently resolved, the authors will see a relatively large correlation with other parameters, e.g. H_2O. If so, they have only one way to go, change or improve the retrieval algorithm, e.g., by using more IASI channels, which are sensitive to H_2O but not to N_2O. You have a lot in IASI.

→ First, we have to clarify the fact that the Contamination Factor is mainly derived from the averaging kernel of the whole state vector. According to the definition of the averaging kernel matrix, the block-diagonal matrix of A represents the averaging kernel matrix of each retrieved parameters and the extra block matrices represent the interference matrices between the different co-retrieved parameters. As it was demonstrated by Rodgers and Connors (2003), these interference matrices are sources of "interference errors" on the target species. Thus, we modified the text accordingly:

> the Contamination Factor (called hereafter *CF*) defined as follows:
>
> the Contamination Factor (called hereafter CF) defined as follows:

$$CF(i) = \sum_j \left| A_{xc(ij)} \right| \frac{\Delta c_j}{x_i} \times 100 \quad (7)$$

Where $A_{xc(ij)} = \frac{\partial \hat{x}_i}{\partial c_j}$ is the submatrix of A corresponding to the interference between the co-retrieve parameter c and the target retrieved species x (Rodgers and Connor, 2003);

Δc_j is the uncertainty on the parameter c at the level j and;

$CF(i)$ is the contamination of the parameter c on the retrieved N_2O $\hat{x}_i$ at the level i.

We have also inserted the following reference in the revised version:

Rodgers, C. D., and B. J. Connor (2003), Intercomparison of remote sounding instruments, J. Geophys. Res., 108, 4116, doi: 10.1029/2002JD002299

In addition, according to the definition of the averaging kernel matrix (cf. eq (2) of the manuscript) there is the following link between *A* and the a-posteriori errors covariance matrix S_x:

A= $S_x K^T S_y^{-1} K = I - S_x S_a^{-1}$

Thus, the interference matrices of *A* give almost the same information as the extra-diagonal submatrices of the a-posteriori errors covariance matrix S_x.

Concerning the possibility of using more channels sensitive to H_2O, the characterisation of the H_2O itself is not the problem since we have as apriori knowledge the operational IASI level 2 and the corresponding error variance. The difficulty here is to characterise the different spectral line comb of the isotopic component of H_2O (especially HDO) and to remove as far as possible the induced contamination of the N_2O profile. So, adding more channels is not the solution. Figure R2 show the isotopic ratio multiply by the concentration and multiply by the cross section of PAN, HDO and HNO3 in B1. The H2O vmr is from IASI operational level 2 product and the concentration of PAN and HNO3 are respectively from Fischer et al. (2014) and Wespes et al., (2007) respectively. The spectroscopic database used here is from HITRAN. This figure shows an important impact of these component in B1. However, the most important component between these three is HDO as it was shown in Clerbaux et al., (2009) since PAN is highly variable and HNO3 has a low impact on the transmittance. In addition, Liuzzi et al. (2016) shows the impact of HDO in B1 by analysing the IASI spectral residuals (obs-calc) with and without retrieving HDO. Thus, we modify the manuscript as follows:

…Clerbaux et al. (2009) and Liuzzi et al. (2016) show the presence of a relatively strong absorption band of the deuterium hydrogen oxide (HDO) also called semiheavy water in the band B1…

We inserted the following reference in the revised version:

Liuzzi, G., Masiello, G., Serio, C., Venafra, S., & Camy-Peyret, C.: Physical inversion of the full IASI spectra: Assessment of atmospheric parameters retrievals, consistency of spectroscopy and forward modelling. Journal of Quantitative Spectroscopy and Radiative Transfer, 182, 128-157, https://doi.org/10.1016/j.jqsrt.2016.05.022, 2016.

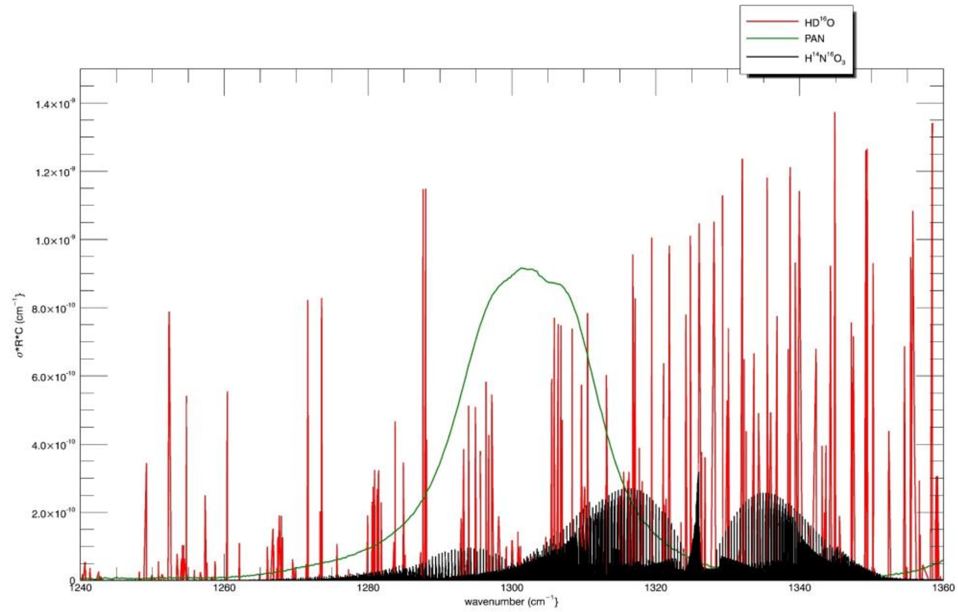

Figure R3*:* Spectral variations of the isotopic ratio multiply by the concentration and multiply by the cross section of PAN, HDO and HNO3 in B1.

12. Page 7 to 10, from Section Validation. With broad AK and a peak of 0.2 at most, the information retrieval comes from the background. There is no point in assessing this layer-by-layer accuracy. All AK broadly peaks around 400 mbar, which means that the N_2O at this layer also receives contribution from the rest of profile. The authors need to check for correlation. In case, as I suspect, the N_2O layer- retrieval at about 309 is strongly correlated with the rest of the profile, the accuracy alone (root mean square error of the a-posteriori covariance matrix) is not a good quality index. I suggest that the author should also compute the N_2O total column. Because of integration along the vertical, this parameter will be more depending on the true state than the background.

→ We agree with the reviewer concerning the fact that the N_2O retrieval at 309 hPa is impacted by the apriori background since the peak is around 0.2. However, the apriori profile is fixed so the variations observed in the retrievals at 309 hPa are not due to the apriori background.
Figure 7 and 8 show that the peak of the averaged *A* is around 0.22 at 309 hPa and around 0.12 at 309 hPa for N_2O_B1 and N_2O_B2 respectively.   These results are consistent with the previous studies from Kangah et al. (2017) and Garcia et al. (2018). The only way to figure out the relevance of using the dominant layer (309 hPa) is through validation studies including

large and representative reference in-situ datasets and scientific analyses of the strengths and weaknesses of these one-layer retrievals. Thus, a retrieval can be considered either as an estimation of the reality or as an estimation of a smoothed reality (Rodgers, 2000). Here we chose to consider the retrievals as estimations of the reality and then analysed the smoothing effects.
So, our work is also to show that despite the smoothing effect due to the shape of the averaging kernel, the retrieval at 309 hPa is sufficiently representative of this layer to study transport processes.
Kangah et al. (2017) studies upper tropospheric transport processes between Asia and the eastern Mediterranean using GOSAT/TANSO N_2O retrieval at 314 hPa. GOSAT averaging kernels have almost the same shape as our IASI N_2O averaging kernel, peaking at around 300 hPa with a full width at half maximum from ~500 hPa to 100 hPa. Despite the broad shape of *A* and a peak hardly higher than 0.1, this study shows that GOSAT N_2O retrievals at 314 hPa allow to study upper tropospheric N_2O transport processes between Asia and the eastern Mediterranean at monthly timescale. In the present paper, we show that our retrieval at 309 hPa capture these upper tropospheric transport processes at daily timescale.
In addition, refer to response #I to the referee #2 about the added value of this work compared with the previous studies.

13. Page 11, Section Troposphere variations... Since the authors failed to show that the N_2O layer- retrieval at 309 hPa is independently resolved, the results in this section could be seriously flawed. Once again, by looking at AK in Fig. 6 and 7 the 309 hPa layer retrieval of N_2O gets contribution from any other layer along the profile. It is important here that the authors show maps of the correlation matrix to get insight into a better understanding of the retrieval quality and accuracy at 309 hPa.

→ Cf. #12 and #11

Minor Comments

14. Page 1, line 11. The authors say, ... *Over the mid-latitude regions, both variations of N2O_B1 and N2O_B2 at 309 hPa are influenced by the stratospheric N2O-depleted air* ***because*** *of the relative coarse shape of the averaging kernel...* I have found clumsy sentences like this throughout the paper. It seems that the cause of a physical phenomenon is the mathematical structure of IASI AK. I think that authors here want to say that... *Because of the relative coarse shape of the averaging kernel, IASI is sensitive to variations of N2O_B1 and N2O_B2 at 309 hPa influenced by the stratospheric N2O-depleted air.* Here the authors seem to suggest that the best retrieval function capable of assessing this phenomenon could be a proper average over the broader part of AK, but then they go completely another way and focus on a single layer.

→ N2O_B1 and N2O_B2 variations are not only due to physical phenomena.
They contain both physical structures from transport processes and other structures partly due to the smoothing effect of the averaging kernel matrix at this level. The physical structure has been clearly demonstrated throughout the paper through validation with in-situ datasets and by showing transport processes at 309 hPa consistent with an extended literature (see

Kangah et al., 2017). We also want to show here the weaknesses of the retrievals at 309 hPa by analysing the impact of these mathematical smoothing effects on N2O_B1 and N2O_B2 variations.

15. Page 3, section 2 IASI. Please put the reference to IASI at the beginning of the section... *IASI (Hilton et al., 2011)...* Furthermore, on line 11 remove the reference to Clerbaux et al. (2009), it is not appropriate once you have cited Hilton et al. at very beginning. Furthermore, on line 13 the reference again to Hilton et al is redundant. Please, remove it.

→ Done

16. Page 3, line 21, Change *RRTOV is a fast...* in *RTTOV is a polychromatic fast...*

→ Done

17. Page 4, Beginning of section. That the IASI spectral coverage includes the N_2O $\nu 1$,$\nu 3$ fundamental absorption band is a well assessed result from molecular spectroscopy, please rephrase the sentence on line 2, the reference to Clerbaux et al. (2009) is not appropriate and unnecessary.

→ We have rephrased the sentence into:

> From molecular spectroscopy (Rothman et al., 2009), it is known that three absorption bands of N_2O are present in the IASI spectral range centred at ~1280 cm^{-1}, ~2220 cm^{-1} and ~2550 cm^{-1}.
>
> Rothman, L. S., Gordon, I. E., Barbe, A., Benner, D. C., Bernath, P. F., Birk, M., et al., The HITRAN 2008 molecular spectroscopic database, Journal of Quantitative Spectroscopy and Radiative Transfer, 110(9-10), 533-572, 2009.

18. Table 1 is never called in the text. Please make proper reference to this table in the text and show also the number of IASI channels used for retrievals.

→ We have inserted the following sentence in the section 5.1.

> Table 1 synthetises the a priori standard deviation errors (σ_a) used for each retrieved parameter in B1 and B2.

→ Although the number of IASI channels selected for the retrievals is written in section 4 and Figure 2, we have clarified this point by inserting a new sentence at the end of section 4.

> To sum up, the number of IASI channels used for the retrievals in the bands B1, B2 and B3 is N = 126, 103 and 0, respectively.