Comparison between the assimilation of IASI Level 2 retrievals and Level 1 radiances for ozone reanalyses. Reply to referee # 1

Emanuele Emili[1], Brice Barret[2], Eric Le Flochmoën[2], and Daniel Cariolle[1]

[1]CECI, Université de Toulouse, Cerfacs, CNRS, Toulouse, France
[2]Laboratoire d'Aérologie, Université de Toulouse, CNRS, UPS, Toulouse, France

Correspondence: Emili (emili@cerfacs.fr)

1 Reply to general comments

We thank the anonymous reviewer for his comments, which helped to improve significantly the manuscript. Detailed replies to his comments follow:

1. *Although both SOFRID retrievals and assimilation of L1 use the RTTOV model, differences may be significant. Indeed the version of the RTTOV coefficients used in both cases is not given. I suspect that SOFRID uses coefficients on 43 levels which are the levels of the retrieval and that the L1 assimilation uses newer coefficients, on 54 or 101 levels (the authors state 104 vertical levels on P7 L20 which does not exist), which may have been build from a different line-by-line model (or different version of that model; or even from a different spectroscopic database). Such differences can have visible impacts on the radiance simulation by RTTOV. Could the authors give mode details about the coefficients? Is it possible to produce SOFRID L2 retrievals using the same version of RTTOV model and RTTOV coefficients as those used in CTM assimilation? And then run the L2 assimilation trials in CTM? That would be a significant improvement to the comparison proposed in this paper!.*

   Answer:

   A verification of the differences between the versions of RTTOV used for this study (v9 for SOFRID and v11.3 for L1 assimilation) confirmed the concerns of the reviewer: SOFRID retrievals were based on a mixture of HITRAN 2000 and 2004 spectroscopic databases, LBLRTM v11.1 radiative transfer and predictors computed for 43 levels, whereas L1 assimilation uses HITRAN 2008, LBLRTM v12.2 and predictors on 101 vertical levels. Therefore, we switched to an updated version of SOFRID and recomputed L2 retrievals with RTTOV 11.1 and the same predictors used for L1 assimilation (101 levels). Other minor differences between RTTOV 11.1 and 11.3 do not impact the radiance computations. New L2 retrievals using RTTOV 11 are named v3.0, as opposed to v1.6 used for the original manuscript. A short summary of the different versions of SOFRID used for this study is given in Tab. 1. The further assimilation of v3.0 retrievals confirmed the results observed previously at tropical latitudes but reduced significantly the differences between L1a and L2a in the Southern Hemisphere mid-latitudes (Fig. 1 and 2). We conclude that the interpretation of the results in the

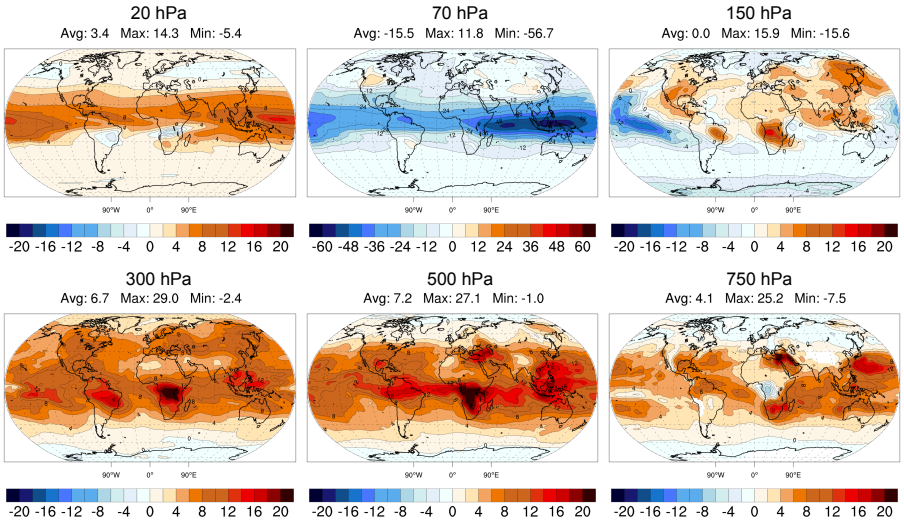

Figure 1. Relative differences (%) between radiances and Level 2 assimilation (L1a minus L2a divided by the correspondent O_3 values of the control simulation) averaged on July 2010. From left to right different pressure levels are displayed covering the stratosphere (top) and the free troposphere (bottom). Average, maximum and minimum values of the displayed fields are given on top of each map. The same pixels selection of the original manuscript is used to produce this figure. The only difference with the original manuscript is that SOFRID v3.0 retrievals are used instead of v1.6 (Tab. 1)

SH mid-latitudes given in the original manuscript was not correct: the L1 assimilation does not reduce the biases when the instrument's sensitivity is low thanks to a better prior. The better performances of L1a in the SH mid-latitudes were mostly due to improved radiative transfer computations.

These findings made us revise the discussion of the results (Sec. 4.1 and 4.3 of the revised manuscript) and, partly, the conclusions. Even if the positive results of the original manuscript are somehow mitigated in the SH, the main conclusions remain valid elsewhere. Moreover, we can now provide a more satisfactory explanation of the differences between L1a and L2a: large differences arise only where the model departures from the SOFRID prior are very large ($> 100\%$), i.e. at low latitudes ($< 40°$). As the second referee also pointed out, differences between L1a and L2a do not really depend on the sensitivity of the instrument but on the accuracy of the prior and on the consequent linearization of the RT. After switching to the same version of RTTOV the results better support this explanation. Please refer to the replies n. 2 and 32 to the second referee for a more detailed discussion on this point.

2. *ECMWF NWP forecasts are used in both SOFRID and CTM assimilation. In the CTM runs, forecasts are taken from the latest available analysis (00 or 12 UTC) as said in sec. 3.1, supposedly every hour, and scaled on the CTM grid. In the SOFRID retrieval process, are the forecasts from IFS used the same way? Before being fed to RTTOV, the meteorological forecasts have to be interpolated to the location of IASI pixels. I would appreciate that the authors describe how this*

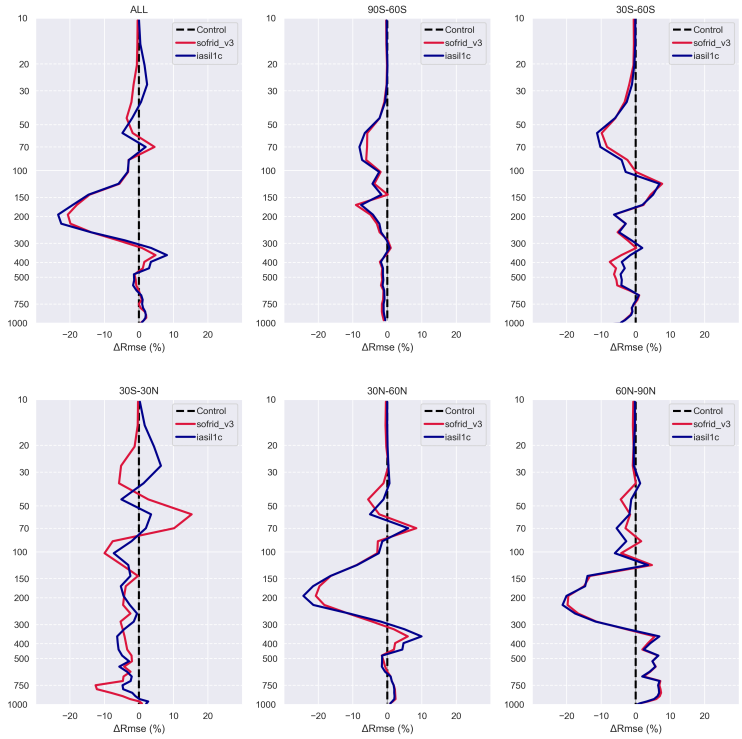

Figure 2. Relative difference of RMSE (ΔRMSE) with respect to radiosoundings for L1a (blue) and L2a (red). The difference is computed by subtracting the RMSE of L1a (L2a) against radiosoundings from the RMSE of the control simulation. Negative values mean that the assimilation improved (decreased) the RMSE of the control simulation, positive values indicate degradation (increase) of the RMSE. The same pixels selection of the original manuscript is used to produce this figure. The only difference with the original manuscript is that SOFRID v3.0 retrievals are used instead of v1.6 (Tab. 1).

Table 1. Versions of SOFRID used for the original and revised manuscript.

SOFRID Version	T and H2O	Cloud factor	RTTOV version	Usage
1.5 (Barret et al., 2011)	EUMETSAT L2	EUMETSAT plus L1	9.0	Original manuscript (only cloud factor)
1.6	ECMWF NWP	L1	9.0	Original manuscript (except cloud factor)
3.0	ECMWF NWP	EUMETSAT plus L1	11.1	Revised manuscript

interpolation is done in SOFRID and in the CTM. ECMWF 4DVAR analyses have ozone in the control variable and assimilate ozone-sensitive information (such as some IASI channels). I would not be surprised that the subsequent ECMWF forecasts are more consistent with the L1 assimilation than with the L2 products. Can the authors elaborate on that point?

Answer:

The meteorological forcing of MOCAGE is retrieved from the ECMWF MARS servers with a 3 hours stepping (available steps for the forecast type "fc"), further regridded to the CTM resolution (2°x2°) and stored as input files. During the MOCAGE execution the meteorological fields are interpolated linearly at hourly sub-steps, which corresponds to the advection time step of the CTM, and vertically (91 to 60 levels, linear interpolation). The observation operator performs an additional bi-linear interpolation at the position of each observation and a linear interpolation to the observation's time. The obtained profiles (temperature, water vapor and ozone on the CTM levels) are used to feed RTTOV, which is in charge of the final vertical interpolation to the coefficients levels. As a consequence, both the spatial and temporal resolution of the RTM vertical profiles are degraded with respect to the original NWP forecasts but are coherent with the resolution of the CTM model. On the other hand, surface properties such as surface skin temperature, which are only needed for the RT and might display a larger variability at smaller scales than the CTM resolution are taken from higher resolution IFS fields (0.125°x0.125°) and interpolated at the IASI pixel using nearest neighbor approach.

SOFRID preprocessor retrieves the IFS operational analysis (type "an") at 00-06-12-18 UTC, regridded to a resolution of 0.25° x 0.25°. All the fields are then interpolated at the closest hour to the IASI pixel and a nearest neighbor interpolation is done to extract the corresponding profiles and surface properties. These information have been added to the revised manuscript (Sec. 2.1.2 and 3.1) and Table 1 (now Table 2 in the revised manuscript) was upgraded accordingly.

Hence, differences between L1a and L2a due to the different origin, resolution and interpolation of the temperature and water vapor profiles might contribute to differences observed in our results. However, we assimilated the IASI main ozone window in the study (980-1100 cm^{-1}) and channels with strong sensitivity to water vapor were excluded both in L1a and L2a (Sec. 2.4). Therefore, we expect the impact of the meteorological profiles on our results to be minor. To confirm this we rerun all the experiments of the manuscript using ERA interim instead of the operational NWP forecasts to force the CTM. ERA interim not only differs in the model configuration with respect to the NWP operational model (e.g. 60 vertical levels for ERA interim versus 91 for the NWP model in 2010), but also for the assimilation: for example no IASI data are assimilated within ERA interim. This introduces some differences in the RTM computations for L1a but also in the control O_3 fields through the CTM forcing, thus requiring to recompute L1a, L2a and the control simulation. Since L2 products are kept the same, potential differences between L1a and L2a due to the meteorological profiles are now amplified. We show in Fig. 3 the same plots as in Fig. 6 (and revised manuscript) but computed using ERA interim forcing. The differences between L1a and L2a at the tropics show similar patterns to previous results and suggest that the main results of this study are not a consequence of the different meteorological profiles. We kept the original choice for the meteorological profiles in the revised manuscript and added a sentence to discuss this point (page 14, line 20).

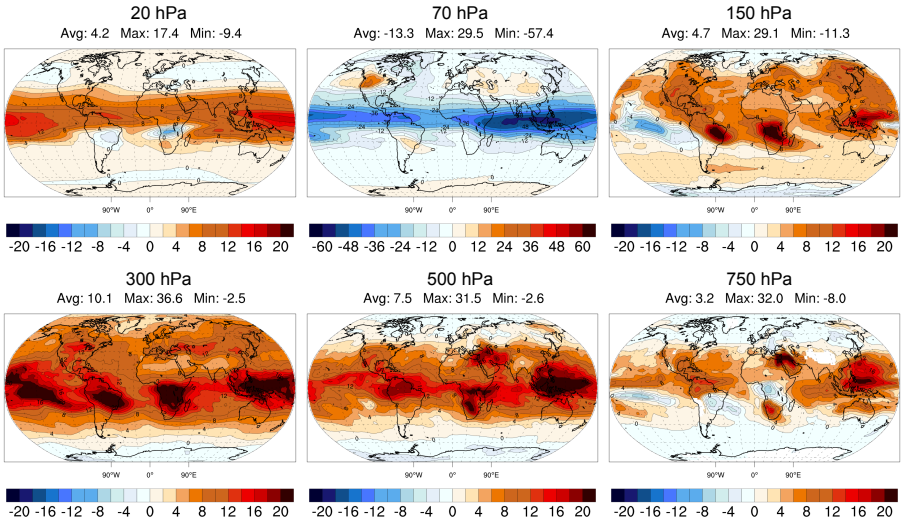

Figure 3. Relative differences (%) between radiances and Level 2 assimilation (same plots as in Fig. 6) but forcing the CTM with ERA-interim analyses instead of ECMWF operational forecasts.

The surface skin temperature has a strong signature in the IASI O_3 window and, even if it is included in the control vector, the different background values used in SOFRID and L1a might have an impact on our results. Hence, we replaced the surface skin temperature used originally in L1a (IFS 3-hourly forecasts at 0.125°x0.125°) with the one already used in SOFRID (IFS 6-hourly analysis at 0.25°x0.25°). Results in Fig. 4 show that the choice of the background skin temperature has not a significant impact on our results. We kept the original choice in the revised manuscript and added a sentence to discuss this point (page 14, line 18).

We do not fully understand the referee's comment about the "better consistency of L1 assimilation with ECMWF forecasts than L2 products": even if some IASI ozone channels are assimilated in IFS, we do not make any use of ozone fields from NWP forecasts in our study. Hence, we do not expect any particular advantage for L1 assimilation compared to L2 assimilation due to the meteorological forcing itself. Conversely, the fact that SOFRID (v1.6 and 3.0) uses the IFS analyses should in principle make SOFRID background radiances closer to L1 observations than the CTM ones, which uses instead forecast fields (also at a degraded spatial resolution).

3. *No description of the L1 and L2 innovation statistics is given. Figures on biases and standard deviations of L1 and L2 innovations would be of interest in this paper. How the value chosen for the observation error standard deviation (0.7 mWm-2sr-1) compare to those statistics? Cloud masks are not really described. Cloud fraction from AVHRR is mentioned but no threshold value is given. How clear cases are selected? A data thinning is applied. Which is the minimum distance between two pixels? No description of the spatial coverage of L1 and L2 is given. Would it be possible to have a typical daily coverage or an average density over the month?*

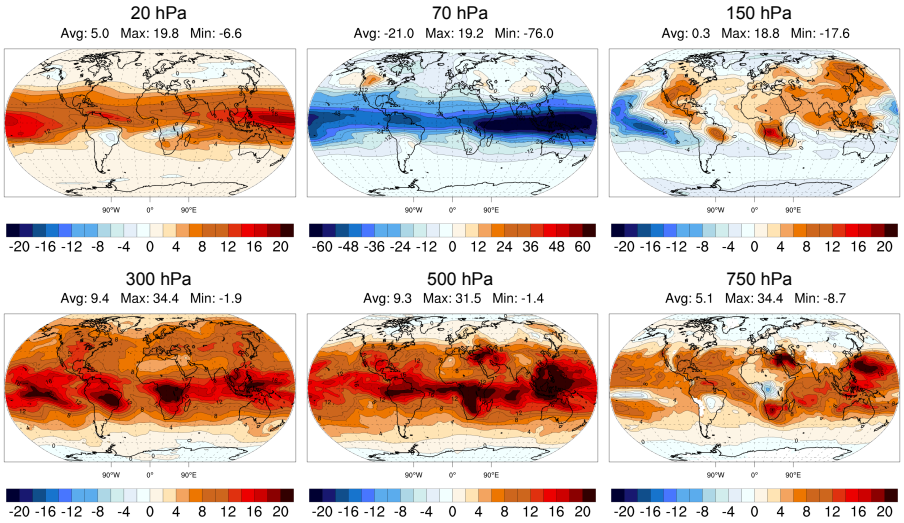

Figure 4. Relative differences (%) between radiances and Level 2 assimilation (same plots as in Fig. 6) but using exactly the same background surface skin temperature for L1 assimilation and L2 retrievals.

Answer:

The answer to this question is split in two parts.

Innovation statistics: The innovation statistics represent one of the main diagnostics of data assimilation experiments and have been carefully evaluated during the study. We report for example in Fig. 5 the average innovations of L1 assimilation (experiment L1a) for the entire month of July 2010. We remark for example that the biases of the control simulation are moderate (about 1 $\mathrm{mWm^{-2}sr^{-1}cm}$) and that the background (forecast) innovation in the middle of the spectral window is smaller than on the tails. The latter is likely due to the different spectral contributions of ozone and skin temperature and the fact that the skin temperature is not a prognostic variable of the CTM (i.e. the background SST is the same in the control and in the forecast). The value of the observation error was deliberately fixed equal to the one used for L2 retrievals to compare L1a and L2a for same settings. Improvements of the observation error covariance, potentially with the aid of more detailed innovation analysis, are left for a future study (page 15, line 22 of the original manuscript). Even though this type of plots contain highly valued information, we prefer not including them since they are not essential for the conclusions of this study and to avoid an excessive length of the manuscript. Moreover, L1 and L2 innovations are not directly comparable because of their different nature (radiances and profiles).

Preprocessing method: the results in Fig. 1 and 2 of this document have been obtained assimilating exactly the same satellite pixels as in the original manuscript, and were shown here to highlight differences due only to the RTTOV version. These pixels were selected based on a combination of cloud masks from an older version of SOFRID (v1.5, see Tab. 1 and description in Sec. 2.1, page 5, line 15) and AVHRR cloud mask available only in most recent L1c files.

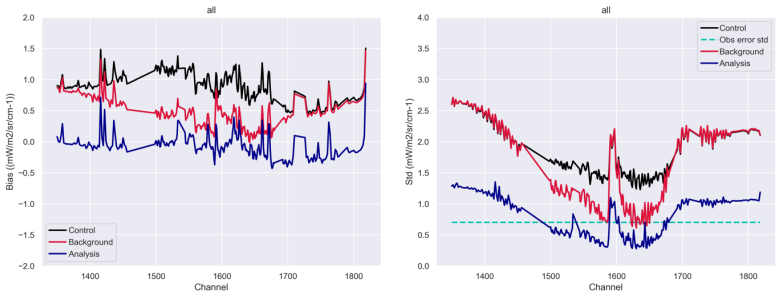

Figure 5. Average (left) and standard deviation (right) of the L1 innovations for the entire period of simulation. The control line represents the innovation with respect to the control simulation (no DA). The dashed turquoise line represents the observation error standard deviation used for the SOFRID retrievals and L1a experiments.

Due to the different data processors these cloud masks are not the same. For both data sources only pixels with a cloud factor smaller than 1% were first selected. The information was given at page 6, line 18 of the original manuscript. The resulting datasets were colocated to ensure that a valid SOFRID retrieval was available for each L1c pixels. Finally, a data thinning was performed hourly: we covered the Earth with a 1°x1° grid and within an hourly loop we retained only the first satellite pixel found within every two grid boxes. A minimum distance of 1° before assimilated pixels is therefore ensured. Overall, the selection resulted in about 3300 pixels per day for the assimilation, as mentioned in Sec. 2.4 of the original manuscript.

With version v1.6 (the retrievals assimilated in the original manuscript), SOFRID was upgraded to use water vapor and temperature profiles from IFS instead of EUMETSAT L2 retrievals (Tab. 1). This increased the number of retrieved pixels with respect to v1.5, since SOFRID was not subject anymore to the availability of the EUMETSAT Level 2. On the other hand, the original cloud mask of v1.5 based on both L1 spectra and EUMETSAT processor was replaced by the L1-only based mask (described in Sec 2.2 of Barret et al. (2011)). To avoid possible cloud contamination the best option was then to keep the original pixel selection done initially with SOFRID v1.5 but using the retrievals from v1.6. Therefore, all results presented in the original manuscript were based on about 3300 assimilated observations per day (page 6, line 22).

With SOFRID v3.0 (RTTOV 11) the EUMETSAT cloud mask was reintroduced in the L2 product, and allowed to apply the full preprocessing procedure described above but using only v3.0 files. At the end, this resulted in an increased number of pixels available for each day to about 5000 (Fig. 8). Differences between L1 and L2 assimilation are enhanced due to the higher number of assimilated observations (Fig. 6 and 7), but show the same patterns as in Fig. 1 and 2. We retained this configuration for the revised manuscript, we extended the description of data thinning (page 7, line 25) and we included a new plot showing the number of assimilated observations per grid point during the simulation period (Fig.

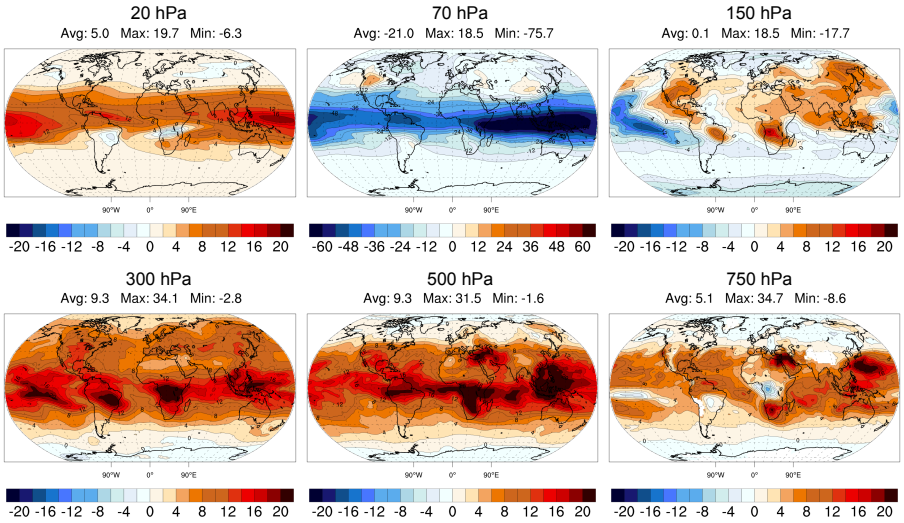

Figure 6. Relative differences (%) between radiances and Level 2 assimilation (same plots as in Fig. 1) but using the new set of colocated observations (right plot in Fig. 8) and SOFRID V3.0 (RTTOV 11). This figure replaces Fig. 2 of the original manuscript.

8, right plot). All figures in this document (except Fig. 1 and 2) and in the revised manuscript are based on this new set of experiments with increased number of observations.

4. *Background covariance error matrix: the values used in this study (2% / 10%) are barely supported by Figure 1. The authors state that the bias may be an important component of the RMSE in Figure 1, is it possible to provide profiles of bias and standard deviation in addition to RMSE? P10 L5, the vertical structure of the B matrix is described as "correlation length of 1 model grid point". Do you mean 1 model level? Please clarify.*

Answer:

We extended Figure 1 with the full validation statistics of the control simulation (bias, standard deviation and RMSE), which are reported here in Fig. 9,10 and 11 respectively. The validation values obtained against MLS are also added to these plots for completeness. We remark that biases can be as high as 30% close to the tropopause and that standard deviation and RMSE values relative to MLS are generally smaller due to the increased accuracy and number of MLS observations. Even if we consider MLS lines as reference for the stratosphere, the values chosen for the background standard deviation may still seem small with respect to those in Fig. 10. However, the assimilation background is more accurate than the control simulation (Fig. 12 for the MLSa forecast), with RMSE values that fell generally below 5% in the stratosphere. We also remind that we neglected the radiances error correlation in our study. This leads to a stronger weight of the assimilated observations, that we compensated by smaller values of the background error covariance. Values of 5%-25% for the standard deviation in the stratosphere and troposphere respectively lead typically to worse

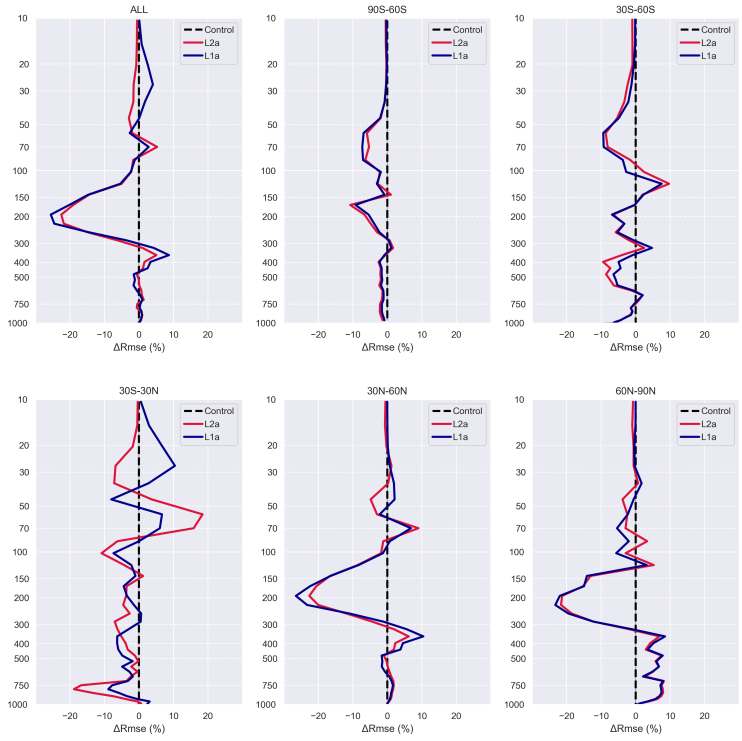

Figure 7. Relative difference of RMSE (ΔRMSE) with respect to radiosoundings for L1a (blue) and L2a (red) (same plots as in Fig. 2) but using the new set of colocated observations (right plot in Fig. 8) and SOFRID V3.0 (RTTOV 11). This figure replaces Fig. 3 of the original manuscript.

reanalyses (not shown). Using a relatively small error in the stratosphere (2%) mitigated the issues encountered with IASI assimilation (L1 and L2) and did not reduce significantly the positive impact of MLS.

This study is focused on comparing L1 and L2 assimilation with identical values for B: we think that the empirical choices for the B matrix are satisfactory for the objective of the study. Further optimization of B and R, which is often done simultaneously (Desroziers et al., 2005), is left for a future study, where non-diagonal terms of R should also be included in L1 assimilation. We extended the discussion of the background error covariance to include elements from this reply and the reply n 24 to the second reviewer (page 12, lines 6 and 12). We also precised that the scale of the vertical error correlation is expressed in number of model levels (page 13, line 1).

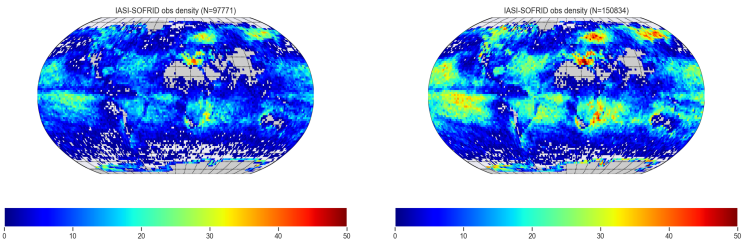

Figure 8. Number of assimilated observations for each model grid point (2°x2°) and for the entire simulation period (July 2010). Number of observations used in the original manuscript on the left (SOFRID 1.6) and for the revised manuscript on the right (SOFRID 3.0). The total number of observations is displayed on the top of each plot.

5. *Results L1 vs L2 Figure 2 shows the relative differences between L1 analyses and L2 analyses. As the values of background error variances are rather small, I would find interesting to show analysis increments difference statistics (average and/or standard deviation). All figures are given in relative difference and no ozone fields are plotted. Except from the value given P11 L6, the reader have no idea how these relative differences compare to the actual ozone concentration. I would appreciate the authors find a way to illustrate the 3D field they want to analyze in their study. Figure 3 (and similar figures) would be more useful if error bars were added. They would help understand whether the differences are statistically significant or not. The statistics are given over the whole month. How stable are they on a day to day basis? Would it be interesting to split the statistics between day and night? The paragraph about the computational cost and convergence issues is interesting but may be placed separately from the scientific results.*

Answer:

The answer is split in 3 parts.

Increments, like innovations, are the direct output of the variational minimization and are among the first diagnostics that we looked at. Examples of increments for the third assimilation window (2010-07-01 03 UTC) are shown in Fig. 13 and 14, which confirms that the absolute increment values are significant in term of typical ozone concentrations (Fig. 15). However, while this type of plots is very meaningful to verify the correct functioning of the DA system, we found not relevant to report average increments in the manuscript. With hourly DA windows the increments are equal to zero most of the time on the global grid due to the moving observation network. Hence, averaged increments do not give valuable information in terms of absolute or relative values. Weighting the average based on the satellite overpasses is not straightforward. On the other hand, the cumulative effect of all increments during the evaluation period is well represented by the analysis fields, which is also the only field that we validated against independent measurements. We think that presenting only the analysis statistics is the best choice for the objective of our study and to avoid an excessive length of the manuscript.

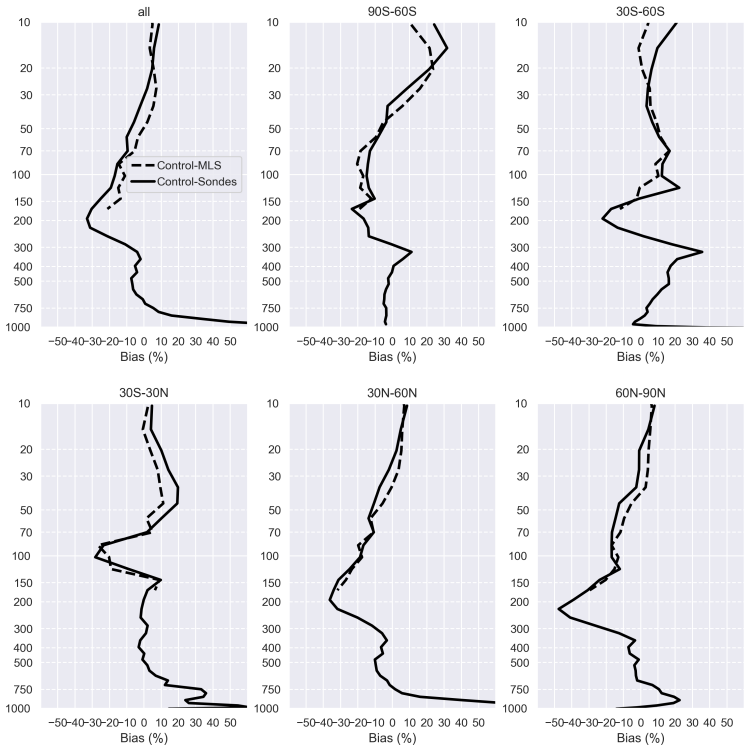

Figure 9. Relative bias of the control simulation with respect to radiosoundings (solid line) and MLS (dotted line) averaged globally (first plot) and for five latitude bands separately (90°S-60°S, 60°S-30°S, 30°S-30°N, 30°N-60°N, 60°N-90°N).

We chose to display only relative differences and relative improvements because ozone varies on an exponential scale. When showing absolute values it is often difficult to appreciate the impact of data assimilation on both the troposphere and the stratosphere, especially when examining differences between similar assimilation experiments. We report in Fig. 15 the average value of ozone of the control simulation, which are used to scale all the maps presented in the study. This figure has been included and commented in the revised manuscript (Sec. 4.1).

Figure 3, 4 and 6 of the original manuscript (and Fig. 2,7 in this document) represent differences of RMSE between the analyses and the control simulation. The RMSE for each simulation (e.g. Fig. 11 for the control simulation) is based on the differences between modeled and observed values for the ensemble of the observations, or for a selection based on latitude. It is not clear to us how to put error bars on such statistics. The statistical significance depends on the number of observations used to compute the various RMSEs, which are reported now in Table 2 and included in the

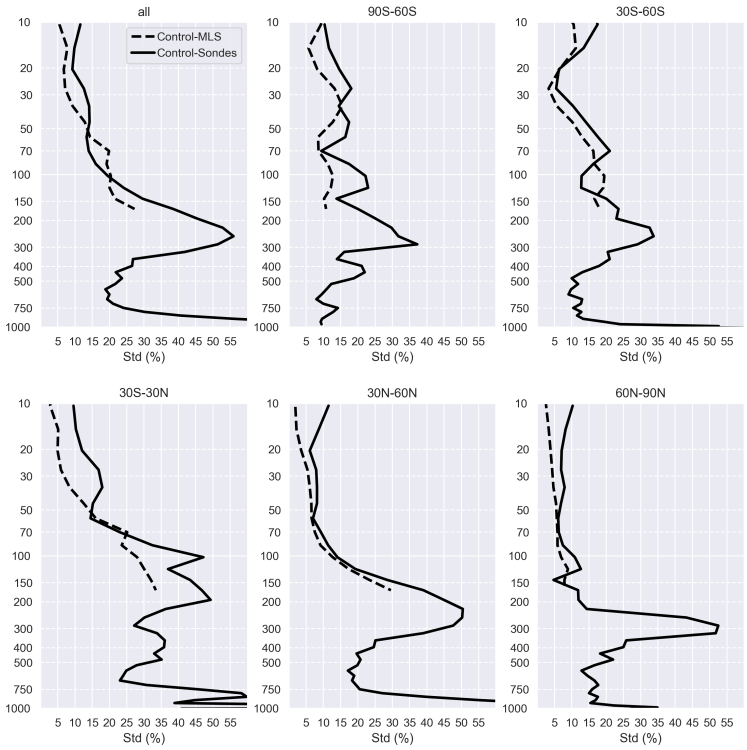

Figure 10. Relative standard deviation of the control simulation. Same plots as in Fig. 9.

revised manuscript for completeness. By looking at observation numbers we recognize that daily RMSE statistics would be difficult to compute for radiosoundings due to a too small number of observations. Similar issues arise if we try to separate between day and night, since radiosoundings are mostly launched at local noon. On the other hand, MLS allows to compute daily or night/day statistics.

We present in Fig. 16 the same plots as in Fig. 4 of the original manuscript but for five different days during the simulation period. We remark that the RMSE display a tendency during the period, with a slow degradation towards the end of the period. We suppose that, without MLS joint assimilation, some errors are continuously injected by IASI, especially in the case of L1a. This points to some unresolved issues with the inversion of radiances, which is probably exacerbated in L1a because of the propagation of the O_3 prior in time. An evaluation of the results over a longer period seem necessary to draw more robust conclusions on this issue. However, thanks to the MLS assimilation, this issue has a limited impact on the main results of the study.

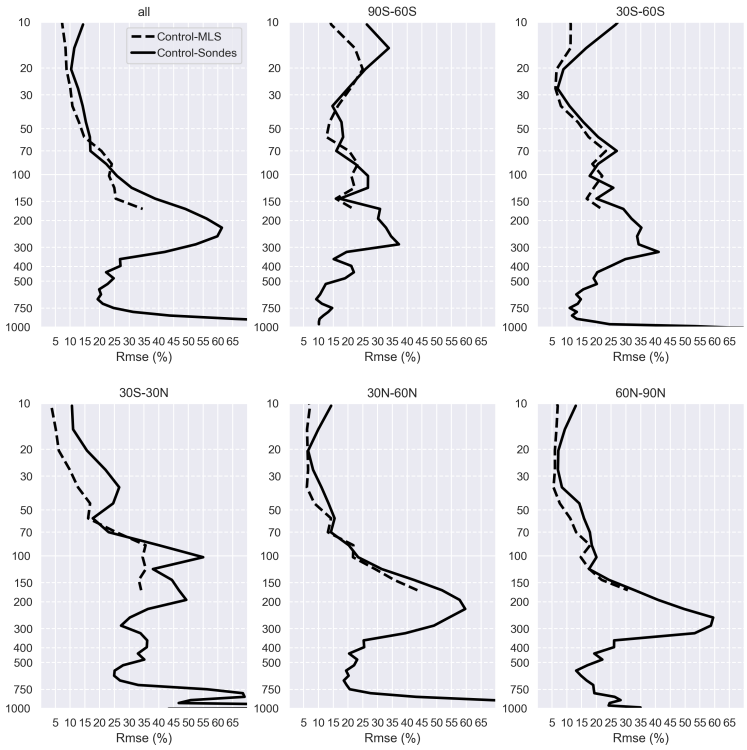

Figure 11. Relative Root Mean Square Error (RMSE) of the control simulation. Same plots as in Fig. 9.

Fig. 17 reports the RMSE statistics for the full period but split between day and night. The day-night separation is computed based on the local sun position at the time of the observation. We remark that some differences appear only at high latitudes (90°S-60°S and 60°N-90°N), but since the number of day-night observations changes dramatically in these regions (e.g. from 15755 to 1212 at 90°S-60°S), a robust interpretation of these differences looks problematic.

The paragraph on the computational cost has become an independent section in the revised manuscript (Sec. 4.2).

6. *Results when MLS is assimilated As in a real system, several sources of observation may be assimilated simultaneously, this section has a real added value. I regret that the results are not shown in a consistent way with the previous section. Figure 2 shows L1a - L2a; Figure 5 should show MLS+L1a - MLS+L2a because we want to compare these two settings.*

Answer:

We agree with the reviewer and the previous figure has been replaced in the revised manuscript with MLS+L1a - MLS+L2a (Fig. 18). The new plot shows that differences are largely reduced in the stratosphere, thanks to MLS, but

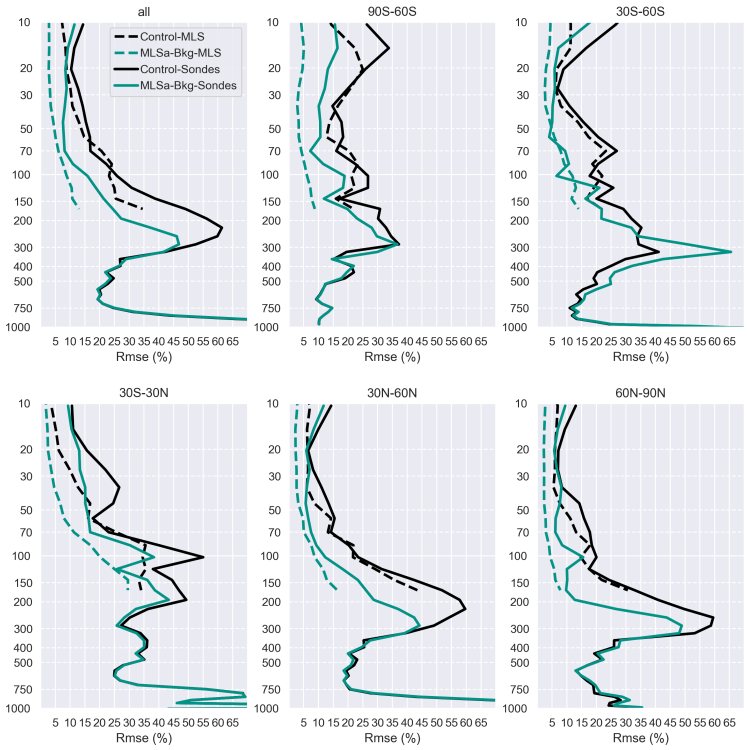

Figure 12. Relative Root Mean Square Error (RMSE) of the control simulation (black lines) and of the MLSa forecast (green lines). Same plots as in Fig. 11.

Table 2. Number of validation observations.

Latitudes	MLS	Radiosoundings
Global	100975	219
90°S-60°S	16967	19
60°S-30°S	17334	9
30°S-30°N	33046	38
30°N-60°N	16669	138
60°N-90°N	16959	15

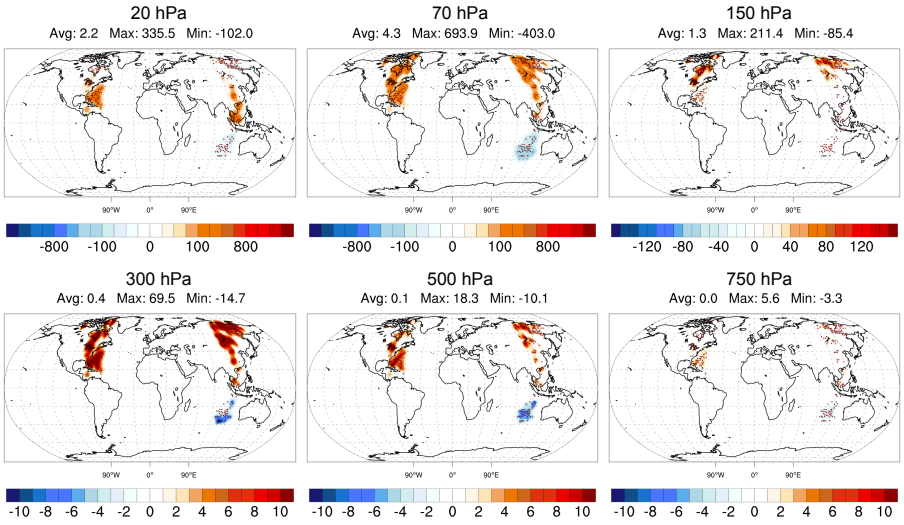

Figure 13. Absolute O_3 increments (ppb units) in L1a experiment for the 2010-07-01 03 UTC window and at different pressure levels in the stratosphere (top plots) and in the free troposphere (bottom plots).

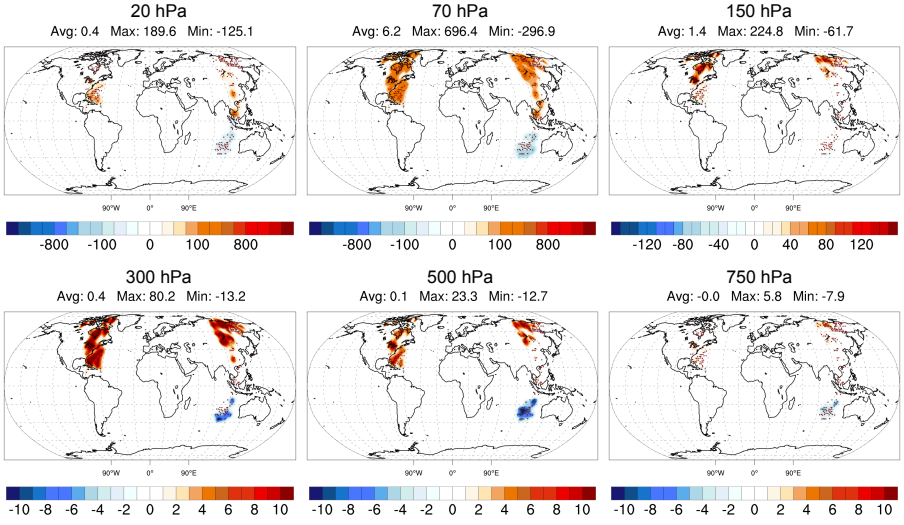

Figure 14. Absolute O_3 increments (ppb units) in L2a experiment for the 2010-07-01 03 UTC window and at different pressure levels in the stratosphere (top plots) and in the free troposphere (bottom plots).

are still significant in the free troposphere, although to a lesser extent than for L1a-L2a (Fig. 6). As a consequence, the discussion of the previous figure has also been removed from the revised manuscript.

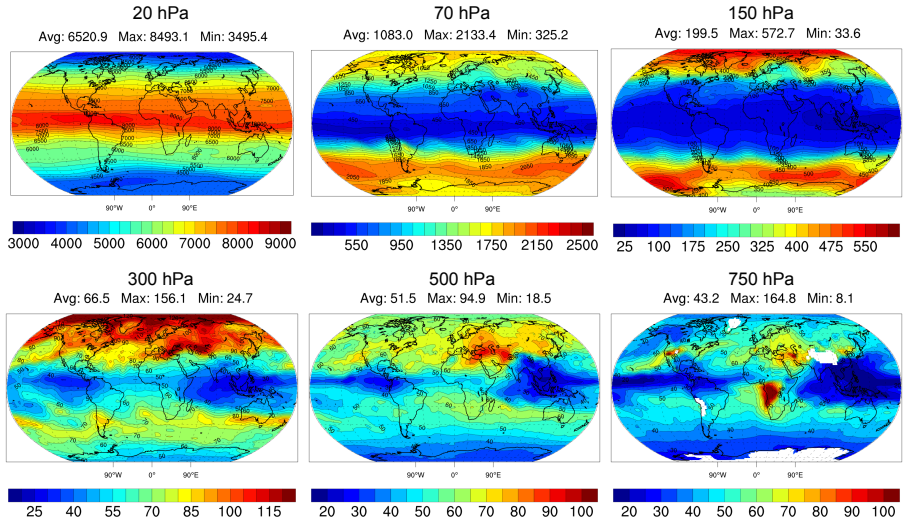

Figure 15. O_3 fields (ppb units) issued from the control simulation averaged on July 2010. From left to right different pressure levels are displayed covering the stratosphere (top) and the free troposphere (bottom). Average, maximum and minimum values of the displayed fields are given on top of each map.

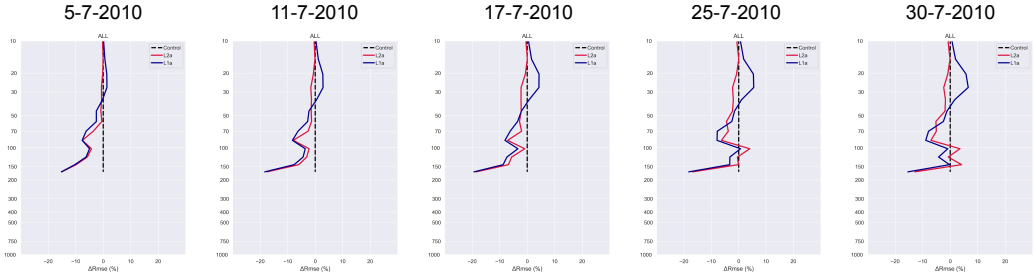

Figure 16. Gain of RMSE (ΔRMSE) computed with respect to MLS for L1a (blue) and L2a (red). Same plots as in Fig. 4 of the original manuscript but shown only for the global average and for five different dates.

2 Reply to specific comments

Answer:

All specific comments have been integrated in the revised manuscript.

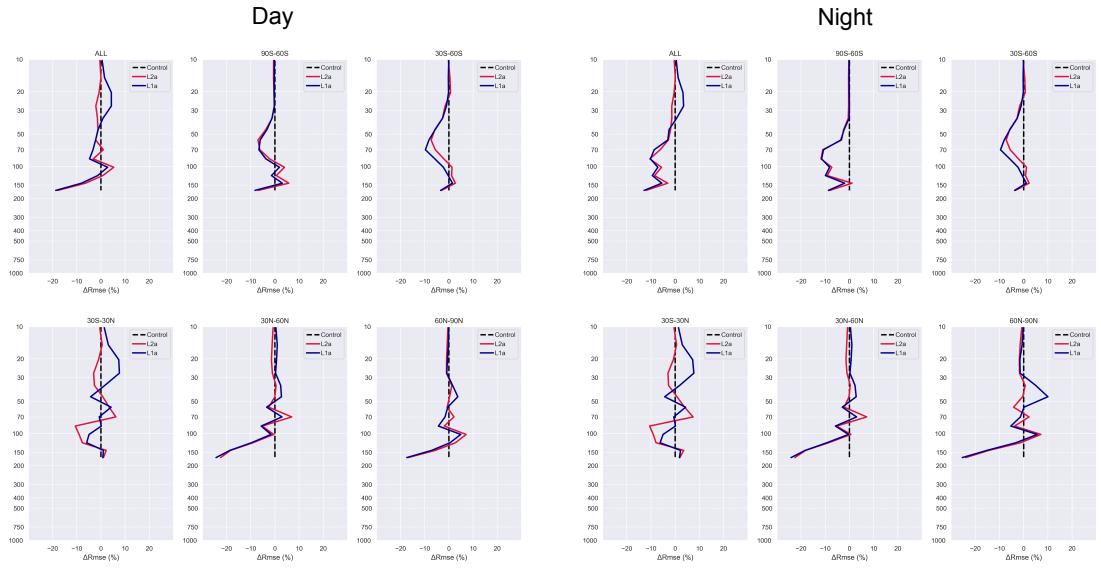

Figure 17. Gain of RMSE (ΔRMSE) computed with respect to MLS for L1a (blue) and L2a (red). Same plots as in Fig. 4 of the original manuscript but computed using only observations during daylight (left panel) and night (right panel).

References

Barret, B., Le Flochmoen, E., Sauvage, B., Pavelin, E., Matricardi, M., and Cammas, J. P.: The detection of post-monsoon tropospheric ozone variability over south Asia using IASI data, Atmospheric Chemistry and Physics, 11, 9533–9548, https://doi.org/10.5194/acp-11-9533-2011, http://www.atmos-chem-phys.net/11/9533/2011/, 2011.

Desroziers, G., Berre, L., Chapnik, B., and Poli, P.: Diagnosis of observation, background and analysis-error statistics in observation space, Quarterly Journal of the Royal Meteorological Society, 131, 3385–3396, https://doi.org/10.1256/qj.05.108, http://doi.wiley.com/10.1256/qj.05.108, 2005.

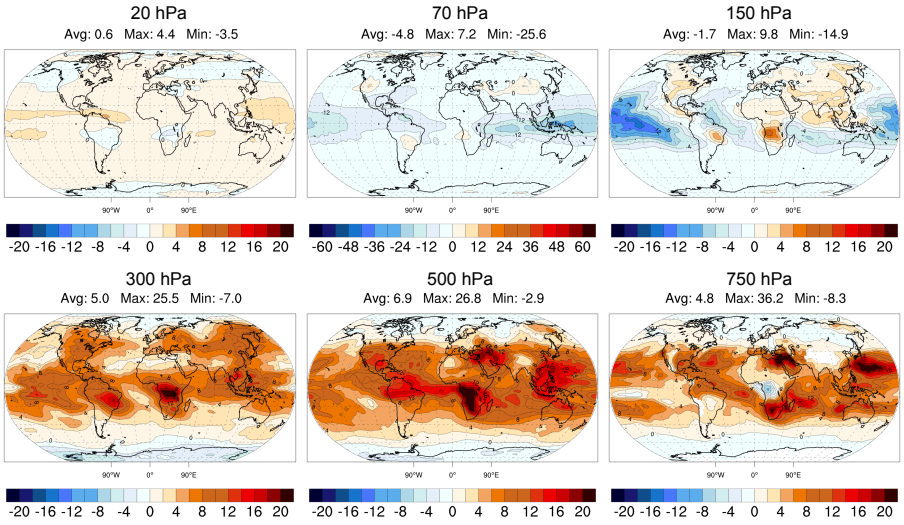

Figure 18. Relative differences (%) between L1a+MLS minus L2a+MLS divided by the correspondent O_3 values of the control simulation averaged on July 2010. From left to right different pressure levels are displayed covering the stratosphere (top) and the free troposphere (bottom). Average, maximum and minimum values of the displayed fields are given on top of each map. This figure replaces Fig. 5 of the original manuscript.

Comparison between the assimilation of IASI Level 2 retrievals and Level 1 radiances for ozone reanalyses. Reply to referee # 2

Emanuele Emili[1], Brice Barret[2], Eric Le Flochmoën[2], and Daniel Cariolle[1]

[1]CECI, Université de Toulouse, Cerfacs, CNRS, Toulouse, France
[2]Laboratoire d'Aérologie, Université de Toulouse, CNRS, UPS, Toulouse, France

Correspondence: Emili (emili@cerfacs.fr)

1 Reply to general comments

We thank the anonymous reviewer for his comments that helped to improve significantly the original manuscript. Detailed replies to his comments follow:

1. *The authors document clear differences between L2 and L1 assimilation, and seem to suggest that L1 assimilation should be considered. However, it seems to me that L2 with an improved a-priori may be a possible alternative to reach a similar performance of the analysis. Would running the IASI L2 retrievals with a varying a-priori, for instance taken from the Copernicus Atmosphere Monitoring Service daily analyses, be a feasilble option? Would that solve part of the problem with L2 compared to L1? Could the authors discuss this in a more balanced way in the conclusion section (and maybe in the abstract as well)?*

   Answer:

   The reviewer is right about the fact that L2 O_3 retrievals can be improved through a better a-priori (see reply n 3 for more details), and using O_3 forecast fields from Copernicus Services might represent a particularly valuable option for L2 production itself. We think that such option could reduce the differences between L1a and L2a in our experiments. However, more generally, the same question raised by our study concerns also models within the Copernicus services themselves (e.g. C-IFS). To assimilate L2 O_3 profiles in C-IFS we would then need to: i) run a first analysis/forecast excluding all IASI O_3 channels ii) run the L2 processor iii) run a second analysis/forecast cycle including only IASI L2 retrievals. On top of the extra numerical cost of such a system and practical difficulties when many instruments are assimilated, this could introduce error correlations between assimilated observations and model forecast that are not yet considered in DA algorithms. Another issue arises for spectral channels that are sensitive to multiple model variables (e.g. T and O_3): splitting the DA problem in different steps (e.g. 4D-Var for T plus 1D-Var for O_3) would result in the same observation to be used twice and might lead to different solution than solving the full problem at once (4D-Var). We have not investigated these aspects in our study and we cannot give final words on the best choice between L1 and L2 assimilation based only on our results in such a context. However, the ensemble of our results plus all the previous

arguments suggest that the L1 assimilation should deserve higher consideration, especially in the context of coupled systems such as the Copernicus Monitoring Services. Therefore, in general we prefer not to suggest an upgrade of the L2 processor with modeled a-priori for the scope of further assimilation. We included these elements in the revised conclusions (page 19, line 11).

2. *Although the L1 and L2 experiments are set up with as much as possible equal inputs and RTM (but different a-priori), there are still subtle differences as discussed in the text. I am wondering how much those differences may also result in differences in performance as documented in the paper? Especially since the differences documented between L1 and L2 are quite small. This leaves me with a bit an uneasy feeling that the results are maybe not fully understood.*

Answer:

Following the comments of the 1st referee we repeated all the experiments using exactly the same version of the radiative transfer for L1a and L2 retrievals (RTTOV 11, see reply n 1 to the 1st referee). We also verified if the differences on the meteorological profiles and surface skin temperature between L1a and L2 retrievals did not impact our results (reply n 2 to the 1st referee). This reduced further the possible sources of differences between the two approaches, which are now limited to: the a-priori, the vertical resolution and the minimization (3D-Var for L1a versus 1D-Var+3D-Var for L2a). Although the L1a-L2a differences are now much reduced in the SH mid-latitudes (original differences were due to the radiative transfer), the new results still show significant differences between L1a and L2a at low latitudes (as high as 30%). We report in Fig. 1 the average difference between the control simulation and the SOFRID a-priori, which show that departures are very large (> 50% and as high as 700%) only at low latitudes. Differences in percent are very large close to the tropical tropapause (150 hPa) because the SOFRID a-priori is representative of mid-latitudes. However, we remark that differences larger than 100% exist also in the free troposphere (300 to 750 hPa). Hence, Fig. 1 confirms that differences within data assimilation arise only when the L2 a-priori is strongly biased (i.e. at low latitudes, see also reply n 32) and strengthens the interpretation of the results given in the original manuscript. We included Fig. 1 in the discussion section (Sec. 4.1) and updated the conclusions of the revised manuscript (page 19, lines 5-8). We think that with these new elements the interpretation of the results is now more robust.

3. *The relative differences in Fig. 2 seem to indicate persistent biases. Maybe it is a lot of extra work, but I wonder how the difference plot of L2 retrievals for the climatological a-priori (presented in the paper), compared to L2 retrievals with MOCAGE profiles would look? Such a plot would be a valuable addition. Would that show similar features as in Fig.2, at around 300-500 hPa ?*

Answer:

An evaluation of SOFRID retrievals using an a-priori issued from a model was performed prior to this study and was actually the main motivation for this work. Indeed, results showed significant differences in the L2 tropospheric columns with the modeled a-priori and a better agreement with independent data (Fig. 2). Differences between the two SOFRID datasets seem qualitatively coherent with the L1a-L2a plot in the revised manuscript (increased O_3 at 300-500-750

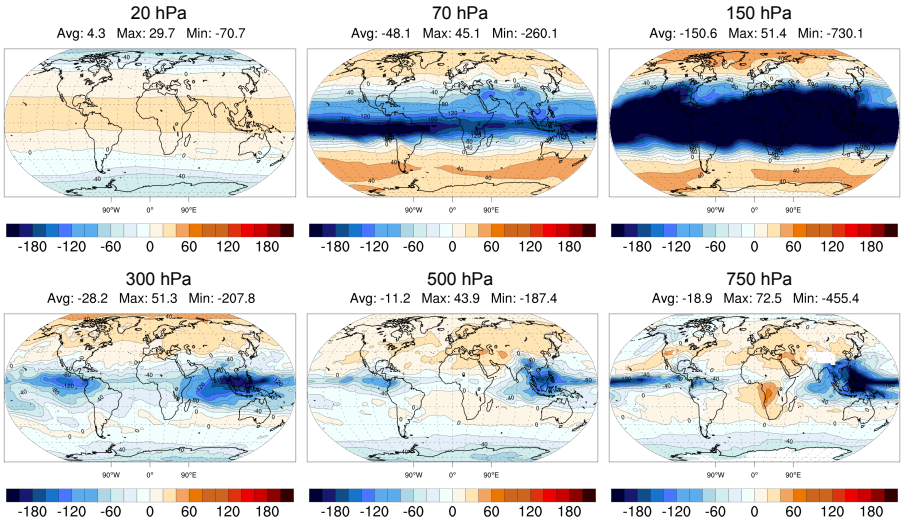

Figure 1. Relative differences (%) between control simulation and SOFRID a-priori (divided by the correspondent O_3 values of the control simulation) averaged on July 2010. From left to right different pressure levels are displayed covering the stratosphere (top) and the free troposphere (bottom). Average, maximum and minimum values of the displayed fields are given on top of each map.

hPa). However, these experiments were based on a different model configuration (linearized chemistry) with degraded resolution (10°x20°). This preliminary analysis was also limited to tropical latitudes and integrated O_3 columns were evaluated without considering averaging kernels. Differently from our manuscript, the above analysis is focused on the L2 retrievals themselves (without further assimilation) and is still under finalization (comparison with radiosoundings). It will be presented in a separate paper once it is finalized and we prefer not to include partial results in our manuscript. In particular, the analysis of averaging kernels cannot be neglected when the assimilation is concerned, which limits the interest of Fig. 2 for our manuscript. Even so, we added a sentence in the conclusion (page 19, lines 12-13) to link these preliminary results to our study.

4. *It would be helpful if the authors could add an image of the IASI averaging kernels, typical examples or averages, for NH, tropics and SH. In this way the reader can better understand at which pressures one may expect an impact of IASI.*

   Answer:

   We report in Fig. 3 the average kernels of the SOFRID retrievals for the month of July 2010, averaged globally and by latitude band. We included this figure in the revised manuscript and added the relative discussion (Sec. 2.1.2).

5. *The impact of IASI in both the L1 and L2 experiments seems to be relatively small, with also negative impacts. Especially when MLS is included as well, which already removes most of the bias around 200 hPa. This baseline ...*

   Answer:

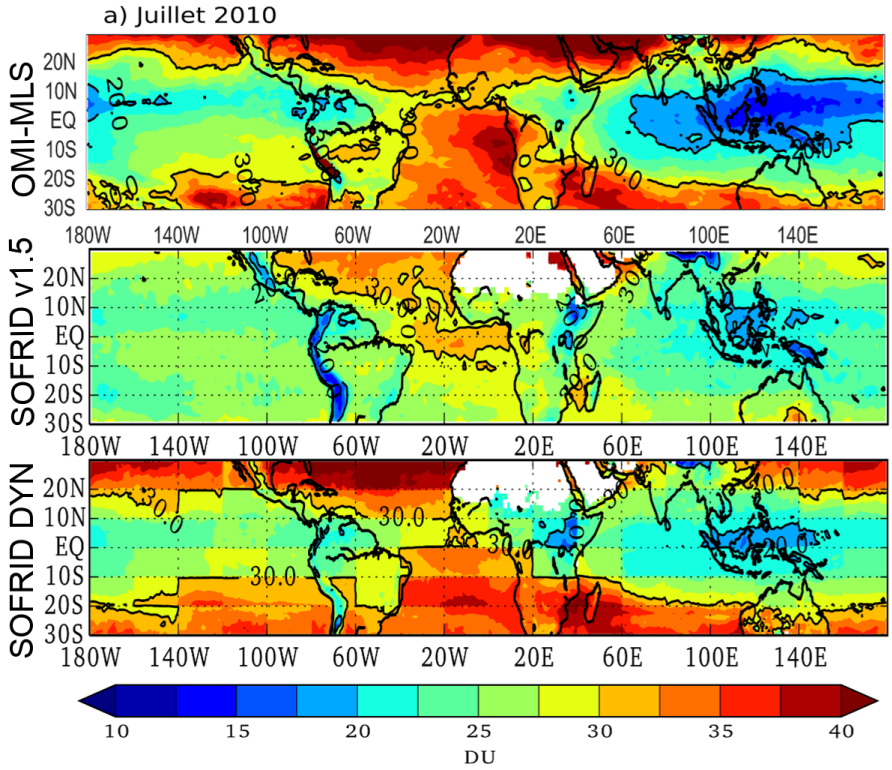

Figure 2. Average O_3 tropospheric columns (1000-100 hPa) from OMI-MLS residual method of Ziemke et al. (2011) (top), SOFRID v1.5 standard retrievals (middle) and SOFRID retrievals issued from a modeled a-priori (bottom).

The question of the reviewer being incomplete, we suppose that he/she raises some doubts about the practical benefits of assimilating IASI on top of MLS for O_3. MLS assimilation is able to well correct the upper troposphere O_3 but, as Fig. 4 and 5 show, there is a significant positive correction of IASI in the tropics that MLS cannot perform. Also, MLS is on-board Aura satellite, which is already well beyond its mission's lifetime, whereas IASI and its successors will be flying for the next decades. We demonstrated in this and previous studies (Emili et al., 2014; Peiro et al., 2018) that the family of IASI sensors is valuable for data assimilation of tropospheric and lower stratosphere O_3. This study provides further elements that we believe are important before implementing IASI O_3 assimilation in operational systems. We updated the conclusions at page 19 lines 20-22 to remind the importance of assimilating MLS in the stratosphere.

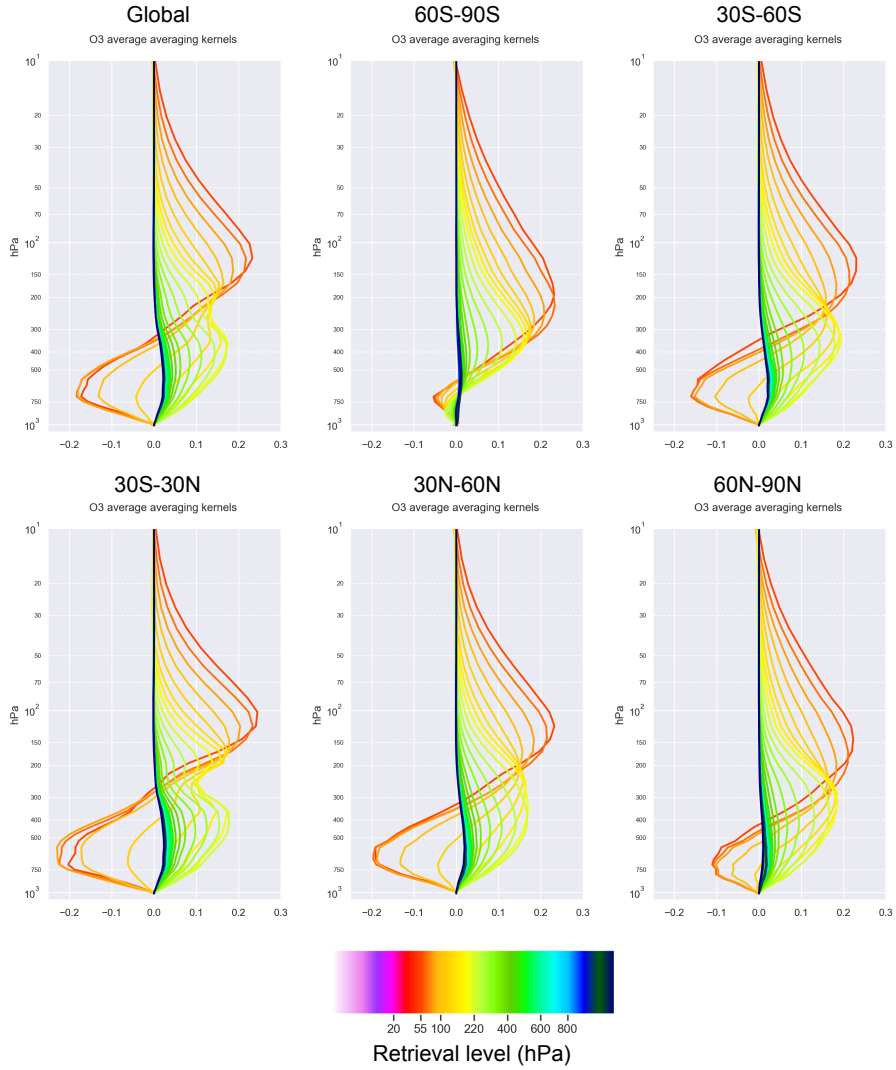

Figure 3. SOFRID O_3 averaging kernels for the month of July 2010 averaged globally (first plot) and for five latitude bands separately (90°S-60°S, 60°S-30°S, 30°S-30°N, 30°N-60°N, 60°N-90°N). Each coloured line corresponds to a retrieval's level, the corresponding pressure is indicated in the colorbar. Only SOFRID levels with a pressure > 50 hPa are displayed for better clarity.

2 Reply to specific comments

1. *Title: "for ozone reanalyses": upon first glance this seems to suggest that the paper presents results of a multi-year reanalysis, which is not the case. Is it necessary to include the word "reanalysis" in the title?*

   Answer:

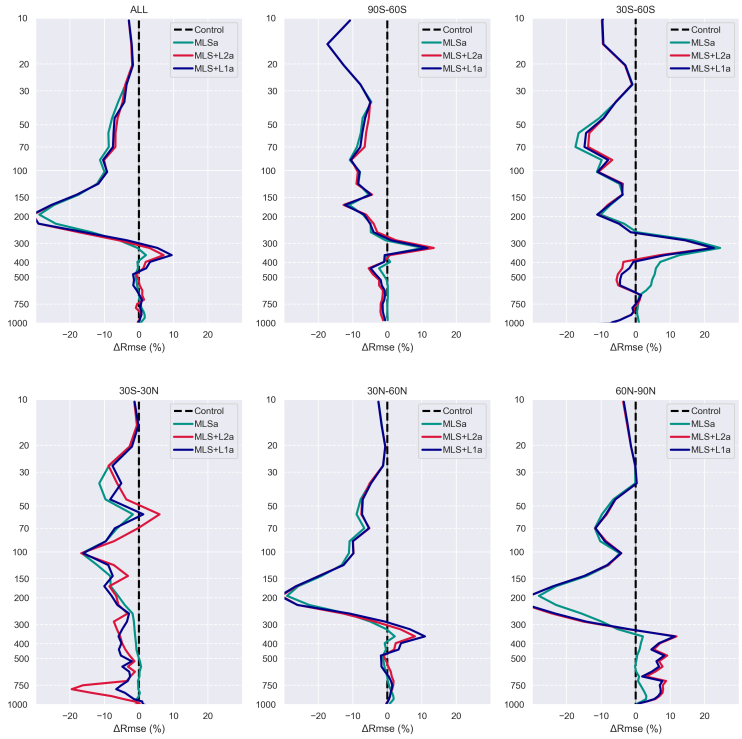

Figure 4. Relative difference of RMSE with respect to radiosoundings for MLS-a (teal), MLS+L1a (dark blue) and MLS+L2a (red). This figure replaced Fig. 6 of the original manuscript.

We used the word “reanalyses” because we presented only DA analyses in this study (instead of forecasts). However, we recognize that the study is mostly methodological and does not present results from long simulations. Therefore, we changed the title to “Comparison between the assimilation of IASI Level 2 ozone retrievals and Level 1 radiances in a chemical transport model” to avoid wrong expectations.

2. *Abstract, l9: "significant differences". The abstract does not give a very firm conclusion. Does the work presented justify the stronger statement that the non-linearity in the retrievals in combination with unrealistic a-priori profiles are the cause of the L1-L2 differences?*

Answer:

Although the results suggest that this seem the case we prefer not to give such a stronger statement in the abstract (and conclusions). The reason is that we did not evaluate explicitly the linearization error of the RTM in our study. Since

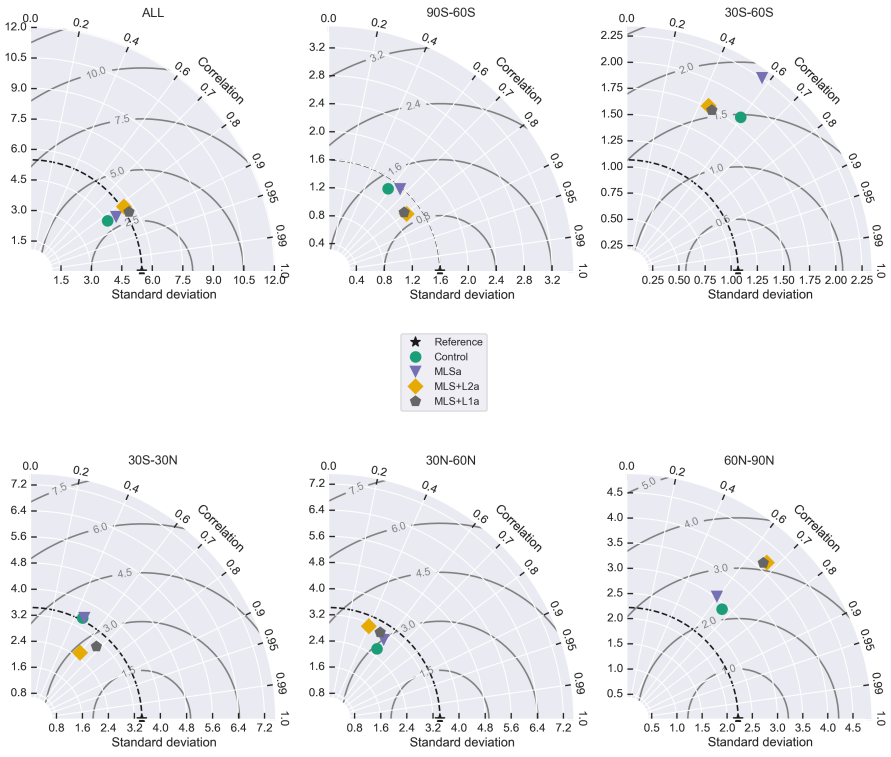

Figure 5. Taylor diagrams of modeled tropospheric ozone columns (340-750 hPa) for the Control simulation (green), MLS-a (violet), MLS+L1a (grey) and MLS+L2a (yellow) averaged globally and for five latitude bands separately. The Taylor statistics are computed against radiosoundings. This figure replaced Fig. 7 of the original manuscript.

RTTOV is already based on the linearization of a full line-by-line RTM (Saunders et al., 2018), doing this evaluation properly would require implementing the original RTM used by RTTOV in the CTM, which was out of the scope of this study. The main objective of our study was instead to provide some practical answers that can guide future developments for IASI assimilation.

3. *Abstract: Is there a clear recommendation from this work? Would L1 assimilation be preferred? A more clear statement would be helpful.*

Answer:

Our results indicate a slightly better variability of the tropospheric O_3 column when assimilating L1 data (Fig. 5). We included this element in the abstract. We also gave different arguments that promote L1 assimilation in the introduction

and in the conclusions (see also Reply n 1), but those are not a direct outcome of our simulations and we prefer addressing the reader to the conclusions to avoid a too lengthy abstract.

4. *Page 2, l5: Useful to mention the averaging kernels as well: .. and DOF linked to the averaging kernels ...*

   Answer:

   The sentence has been changed according to the suggestion of the reviewer.

5. *Page 2, l9: "First atmospheric composition models": Reformulate*

   Answer:

   The sentence has been reformulated.

6. *Page 2, l19: "However, some aspects of the Data Assimilation (DA) approach differ between the chemistry and meteorology communities." Please be more specific, or remove the sentence.*

   Answer:

   The sentence was removed.

7. *Page 2, l23: "This resulted necessary to avoid"; please reformulate.*

   Answer:

   The sentence has been reformulated.

8. *Page 2, l24: About the historical background: I was wondering about the problems encountered when assimilating L2 retrievals in NWP (paper of Eyre)? Is it the non-linearity and a-priori dependence (as suggested by the text), or is it a simplification of the retrieval results? The latter could arise when e.g. kernels and full covariance matrices are not used in the assimilation, or not provided by the retrieval teams, which would clearly lead to strong a-priori dependence of the analyses.*

   Answer:

   No mention to the averaging kernels was found in Eyre et al. (1993), which suggests that their difficulties arise from the simplification of the retrievals (missing use of averaging kernels). The sentence has been updated to make this point clearer (page 2, line 29).

9. *Page 2, l27: I would suggest to refer to the book of Rodgers as well. Also the paper of Migliorini 2012 is relevant here.*

   Answer:

   The references have been updated.

10. *Page 3, l29: "Both are based" It is not so clear what "both" refers to: the two studies, or SOFRID and MOCAGE.*

    Answer:

    We refer to SOFRID and MOCAGE DA system, the sentence has been corrected.

11. *Page 4, Section 2.1: METOP also has the GOME2 instrument. Has the synergy IASI-GOME2 been considered? Why the choise to use MLS?*

Answer:

MLS retrievals are very accurate and provide vertically resolved information that are inaccessible to UV sounders like GOME2. We used MLS to ensure an accurate stratospheric profile and evaluate the impact on IASI TIR assimilation, which was the focus of the study. The assimilation of GOME2 L2 profiles requires some particular care for correcting observation biases (Van Peet et al., 2018) and was not considered for this study. However, we agree with the reviewer's about the interest of performing IASI and GOME2 joint assimilation in future. We included this perspective in the conclusions of the revised manuscript (page 20, lines 1-2).

12. *Page 5, l4: "increased biases": what does "increased" refer to?*

Answer:

The word increased was removed.

13. *Page 5, l10: "LA" ?*

Answer:

LA standed for Laboratoire d'Aérologie, it has been replaced by B. Barret.

14. *Page 7, l11: "ECMWF NWP model" Please replace the word "model" by e.g. "NWP model and assimilation system".*

Answer:

Done

15. *Page 7, l19: The RTTOV versions for the L1 and L2 experiments are different, see table 1. Can the authors be sure that this does not significantly influence the results/conclusions?*

Answer:

The RTTOV version is now the same for both L1a and L2a and the revised manuscript has been updated based on new results. See also reply n 1 to the first reviewer.

16. *Page 7, l29: "... and was extended ..."*

Answer:

Correction included.

17. *Page 8, l19: "we assimilate here directly the full L2 profiles (43 levels)". Migliorini wrote a paper (2007) to discuss an efficient interface between L2 retrievals and data assimilation which is relevant in this context. Because the DOF is quite low, this impllies that a lot of noise (43-DOF) is presented to the assimilation when all 43 levels are included. In principle*

I agree that this avoids any loss of information, but in practice I wonder if the full information may introduce numerical issues (randomness) in the system, especially when this is combined with vertical interpolations? Please comment.

Answer:

We agree with the reviewer about the pertinence of assimilating transformed SOFRID retrievals instead of the full profiles. This would reduce the cost of the L2 assimilation. However, we remind that in this study we used the full L2 error covariance matrices for the assimilation. Hence, the intrinsic noise of each observation level is somehow dampened within the computation of the cost function and its gradient. Source of randomness could result from inaccurate inversion of the observations error covariance matrices. As a matter of precaution, retrievals with inaccurate inverse were already excluded from the assimilation, but they represented less than 0.5% of all the available retrievals (before cloud filtering). Finally, the minimization always showed expected convergence behavior and we did not experience any particular randomness that could be related to numerical issues: repeating the same simulation twice gave same results within the precision of the output format (32 bit floating point). We added a sentence and the appropriate references in the revised manuscript to mention the possibility of using transformed retrievals (page 10, lines 8-11).

18. *Page 8, l22: "The steps for the computation of modeled radiances are equal to the profiles ones until the vertical interpolation." Please reformulate.*

    Answer:

    The sentence has been reformulated.

19. *Page 8, l27: "climatological profile" -> "climatological profiles"*

    Answer:

    Done.

20. *Page 8, l29: "as it is done within SOFRID retrieval scheme" please reformulate*

    Answer:

    Done.

21. *Page 8, l29: Does this mean that the SST is treated differently in the L1 vs L2 assimilation experiments?*

    Answer:

    No. Since SOFRID is a 1D-Var retrieval it does not propagate information to further retrievals as well, it also does not include any SST spatial error covariance. We better underlined the similarities between the two approaches in the revised text (page 10, lines 24-25).

22. *Page 9, l8: "initialized on 1st June 2010"; replace by "initialized on 1 June 2010"*

    Answer:

    Done.

23. *Page 9, l11: "a diagonal matrix (i.e. with no inter-channel correlation) is used". Is the same diagonal matrix used in the retrievals that produce the L2 dataset?*

Answer:

Yes. It is specified in the sentence before.

24. *Page 9, bottom to p10, top: I got a bit lost with the numbers provided for the background standard deviation, also in comparison with Fig. 1. I understand that a background standard deviation during assimilation is often smaller than the std of a free model run, but I do not manage to connect the numbers with e.g. Fig.1 in combination with Fig.3?! What is the motivation to go from 5% to 2% in the stratosphere, which seems like a big step and does not seem justified given Fig.1? Does this choice lead to very small stratospheric increments? What is the justification for a step between stratosphere and troposphere? The standard deviation should depend on the data assimilated. Normally these kind of numbers are optimised with e.g. a chi-square test.*

Answer:

The intent of Fig. 1 in the original manuscript was not to provide quantitatively values for the background standard deviation but: i) to display the main features of the error profiles ii) provide the reference values for the following figures that compare different experiments (Fig. 3 and 6 of the original manuscript). The empirical choice of values reported at page 9 resulted from a large number of experiments and we address the reviewer to the reply n 4 to the 1st reviewer for additional arguments that support the standard deviation values used in the end. In the reply n 5 we also show some examples of stratospheric increments that remain significant in terms of O_3 concentration.

The justification of using a step between the stratosphere and the troposphere follows from the vertical features of the errors. The modeled profile is generally more accurate in the stratosphere, especially when MLS is also assimilated (see Fig. 12 of the replies to the 1st reviewer). We are aware of the fact that the background standard deviation depends on the assimilated observations. However, the objective of the study being to compare L1 and L2 assimilation, we did not want to introduce additional differences between the experiments due to different background error covariances. Hence, the most pragmatic option was to find a sort of compromise that fits reasonably well for all the presented experiments. The chi-square test is a useful diagnostics in DA but we did not consider it appropriate in our study for the following reasons: i) we use a very simplified observation error matrix and optimizing only the background error but keeping R fixed does not seem relevant ii) it is generally not possible to keep an optimum chi-square when using the same B but changing the assimilated instruments.

We included elements from these replies in the revised manuscript (page 12 and top of page 13) to better explain the reasoning behind the choices for B.

25. *Page 10, l15: One would expect that features in the boundary layer, and, to a lesser extent, the free troposphere show vertical correlations because of e.g. vertical mixing and convection. This in contrast to the stratosphere.*

Answer:

We agree in principle with the reviewer, and we did test a configuration with larger vertical correlations in the troposphere (1.5 model levels) than in the stratosphere (0.5 model levels), but results were not significantly better. We updated the text to include this element (page 12, lines 23-25, page 13, lines 1-2).

26. *Page 10, first part: I think the B matrix discussion can be shortened somewhat, because optimising it is not so important for the topic of the paper.*

Answer:

We removed some sentences corresponding to the settings employed in previous studies since they were not strictly necessary.

27. *Page 11, l5: "are in generally ", remove "in"*

Answer:

Done.

28. *Page 11, l7: "are found in correspondence of tropical latitudes". Please reformulate.*

Answer:

Done.

29. *Page 11, l14: "equivalence between L1 and L2 assimilation is not verified for O3". I suggest to explicitly add "for O3 retrievals in the thermal infrared".*

Answer:

Done.

30. *Page 12, l2: "The assimilation increases the RMSE of the tropospheric profile at northern latitudes (60N-90N)." I guess you mean in the range 350-1000 hPa.*

Answer:

The range has been added in parenthesis.

31. *Page 12, l8: "the other way round". Replace by "around".*

Answer:

Done.

32. *Page 12, l17: "Hence, we expect a stronger impact of the prior in the retrieval results,". I do not understand this. It means that the DOF is smaller, which is clearly observed at altitudes around 200hPa, where the improvement with IASI data is much more limited in the SH. But a-priori plays only a role through non-linear effects. Why would these non-linear effects be larger in the SH?*

Answer:

The reviewer's comment was very pertinent: our conclusion was not supported by the data (Fig. 1) but was a wrong interpretation of the original results (see reply n 1 to the 1st reviewer): the new experiments show indeed similar results also in the SH mid-latitudes. The discussion has been revised according to the new results.

33. *Page 12, l27: "The only exceptions are a lower RMSE degradation at 50 hPa". Should we believe the sondes or MLS here? How many sonde launches are included, and confirm the 60 hPa bias?*

Answer:

The number of radiosoundings has been reported in the revised manuscript, and we believe that higher confidence should be given to the MLS validation, especially with respect to standard deviation values. This aspect has been better highlighted in the revised manuscript (page 12, line 10 and page 16, lines 10-12).

34. *Page 12, l32: "total computing time is 3.9 CPU hours". Is this on a single core/node ??*

Answer:

The simulations have been performed on one Xeon E5-2680 node with 24 cores. The values given in hours are expressed in CPU time, which depends on the CPU type / frequency and give an approximate idea of the relative computational cost of L1a versus L2a. The run time (or elapsed time), given in minutes, depends on the parallel implementation and the number of cores/nodes used for the simulations. We replaced "total computing time" by "total CPU time" to avoid confusion.

35. *Page 13, Fig.5: The figure seems to prove that the analysis of MLS and of IASI are more consistent in the case of L1, while the L2 plot indicates biases between the instruments, especially in the tropics. This could be discussed a bit more explicitly.*

Answer:

Following the comment n 6 of the 1st reviewer, we replaced Fig. 5 with the MLS+L1a - MLS+L2a averages. The discussion has been revised accordingly.

36. *Page 14, l16: "mixed elsewhere"? Do you mean to say "mixed results are obtained elsewhere"?*

Answer:

The sentence has been changed.

37. *Page 14, l22: "which are differences" please correct the English.*

Answer:

The phrase has been corrected.

38. *Page 14, l23: "reanalyses". I suggest to broaden this to e.g. "analyses and reanalyses".*

Answer:

Done.

39. *Page 14, l24: "between the L2 retrieval and the assimilation algorithm ". I suggest to change to " between the L2 retrieval and the L1 observation operator" or something like that*

Answer:

Since we also extended the control vector of the assimilation, we rephrased with: “between the L2 retrieval and the L1 assimilation”.

40. *Page 14, l24: "using the same RTM". But the version of the RTM is different ?!*

Answer:

The version of the RTM is now the same in the revised manuscript.

41. *Page 14, l28: "between each other". I suggest "against each other".*

Answer:

Corrected.

42. *Page 14, l30: "Main findings suggest". I suggest "The results suggest ..."*

Answer:

Corrected.

43. *Page 15, l6: "We could imagine". I could as well, but is this a recommendation?*

44. *Page 15, l6: The non-linearity of the retrieval may be very different for different species and spectral ranges. Which ones would be candidates to show significant differneces between L1 and L2?*

45. *Page 15, l9: I was wondering if we may expect positive synergies between IASI and GOME2? They are both on the same platform. Please discuss.*

Answer:

This paragraph groups the the answers to the above 3 comments.

The original text at page 15, lines 3-8 was replaced since it was not appropriate anymore with the new results (reply n 1 to the first reviewer). We don’t have enough experience with other retrievals than O_3 to give detailed recommendations on which species and L2 product might be affected by similar issues. Preliminary analyses indicate that O_3 profile retrieval in the UV region might a display similar behavior than in the TIR. However, the degree of non-linearity depends significantly on the retrieval’s a-priori: a case-by-case analysis would be needed in this sense. We included the perspective

of GOME2 assimilation in the revised manuscript and recommended to analyze potential dependence of the results to the a-priori (page 19, line 23).

46. *Page 15, l15: "Level 2 products can be aggregated". This useful remark could be phrased more generally by refering to "Use of the Information Content in Satellite Measurements for an Efficient Interface to Data Assimilation" by Migliorini et al, 2007. Through L2 the number of useful observations presented to the assimilation may be optimised (to ultimately match the DOF).*

Answer:

The methodology of Migliorini et al. (2008) is now referenced and briefly discussed in Sec. 3.3 (reply n 17). At page 15, line 15 we were specifically addressing models that does not cover the full atmosphere and the case of vertical selection/aggregation of measurements based on user needs. To our understanding, the method proposed by Migliorini et al. (2008) does not seem to be adapted this type of needs because it compresses the full retrieval's information.

References

Emili, E., Barret, B., Massart, S., Le Flochmoen, E., Piacentini, a., El Amraoui, L., Pannekoucke, O., and Cariolle, D.: Combined assimilation of IASI and MLS observations to constrain tropospheric and stratospheric ozone in a global chemical transport model, Atmospheric Chemistry and Physics, 14, 177–198, https://doi.org/10.5194/acp-14-177-2014, http://www.atmos-chem-phys.net/14/177/2014/, 2014.

Eyre, J. R., Kelly, G. A., McNally, A. P., Andersson, E., and Persson, A.: Assimilation of TOVS radiance information through one-dimensional variational analysis, Quarterly Journal of the Royal Meteorological Society, 119, 1427–1463, https://doi.org/10.1002/qj.49711951411, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49711951411, 1993.

Migliorini, S., Piccolo, C., and Rodgers, C. D.: Use of the Information Content in Satellite Measurements for an Efficient Interface to Data Assimilation, Monthly Weather Review, 136, 2633–2650, https://doi.org/10.1175/2007MWR2236.1, http://journals.ametsoc.org/doi/abs/10.1175/2007MWR2236.1, 2008.

Peiro, H., Emili, E., Cariolle, D., Barret, B., and Le Flochmoën, E.: Multi-year assimilation of IASI and MLS ozone retrievals: variability of tropospheric ozone over the tropics in response to ENSO, Atmospheric Chemistry and Physics, 18, 6939–6958, https://doi.org/10.5194/acp-18-6939-2018, https://www.atmos-chem-phys.net/18/6939/2018/, 2018.

Saunders, R., Hocking, J., Turner, E., Rayer, P., Rundle, D., Brunel, P., Vidot, J., Roquet, P., Matricardi, M., Geer, A., Bormann, N., and Lupu, C.: An update on the RTTOV fast radiative transfer model (currently at version 12), Geoscientific Model Development, 11, 2717–2737, https://doi.org/10.5194/gmd-11-2717-2018, https://www.geosci-model-dev.net/11/2717/2018/, 2018.

Van Peet, J. C., Van Der, R. J., Kelder, H. M., and Levelt, P. F.: Simultaneous assimilation of ozone profiles from multiple UV-VIS satellite instruments, Atmospheric Chemistry and Physics, 18, 1685–1704, https://doi.org/10.5194/acp-18-1685-2018, 2018.

Ziemke, J. R., Chandra, S., Labow, G. J., Bhartia, P. K., Froidevaux, L., and Witte, J. C.: A global climatology of tropospheric and stratospheric ozone derived from Aura OMI and MLS measurements, Atmospheric Chemistry and Physics, 11, 9237–9251, https://doi.org/10.5194/acp-11-9237-2011, http://www.atmos-chem-phys.net/11/9237/2011/, 2011.

Comparison between the assimilation of IASI Level 2 ozone retrievals and Level 1 radiances [..*]in a chemical transport model

Emanuele Emili[1], Brice Barret[2], Eric Le Flochmoën[2], and Daniel Cariolle[1]

[1]CECI, Université de Toulouse, Cerfacs, CNRS, Toulouse, France
[2]Laboratoire d'Aérologie, Université de Toulouse, CNRS, UPS, Toulouse, France

Correspondence: E. Emili (emili@cerfacs.fr)

Abstract. The prior information used for Level 2 (L2) retrievals in the thermal infrared can influence the quality of the retrievals themselves and, therefore, their further assimilation in atmospheric composition models. In this study we evaluate the differences between assimilating L2 ozone profiles and Level 1 (L1) radiances from the Infrared Atmospheric Sounding Interferometer (IASI). We minimized potential differences between the two approaches by employing the same radiative transfer code (RTTOV) and a very similar setup for both the L2 retrievals (1D-Var) and the L1 assimilation (3D-Var). We computed hourly 3D-Var [..[2]]analyses assimilating respectively L1 and L2 data in the chemical transport model MOCAGE and compared the resulting O_3 fields among each other and against ozonesondes. We also evaluated the joint assimilation of limb measurements from the Microwave Limb Sounder (MLS) on top of IASI to assess the impact of stratospheric O_3 on tropospheric [..[3]]analyses. Results indicate that significant differences can arise between L2 and L1 assimilation, especially in regions where the L2 prior is strongly biased (at [..[4]]low latitudes in this study). In these regions the L1 assimilation provides a better variability of the free-troposphere ozone column. L1 and L2 assimilation give instead very similar results [..[5]]at high latitudes, especially when MLS measurements are used to constrain the stratospheric O_3 column. [..[6]]A critical analysis of the potential benefits and drawbacks of L1 assimilation is given in the conclusions. We also list remaining issues that are common to both the L1 and L2 approaches and that deserve further research.

1 Introduction

The global monitoring of the atmospheric composition relies on a large number of dedicated satellite missions and on the sustained improvement of numerical forecast models. Research and operational centers provide today both satellite based reanalyses and forecasts of atmospheric composition for a large number of applications, spanning from stratospheric ozone monitoring (van der A et al., 2010) to climate change (Flemming et al., 2017) and air-quality (Zhang et al., 2012; Marécal et al., 2015).

*removed: for ozone reanalyses
[2]removed: reanalyses
[3]removed: reanalyses
[4]removed: the tropics and in the southern hemisphere in
[5]removed: in the northern hemisphere
[6]removed: We conclude this study by listing

Satellite sensors measure the spectral signature of gases and aerosols on the radiation field that traverse the atmosphere. Retrieving the concentration of a given gas from the radiation measured at the satellite position represents an inverse problem that is in most cases ill-posed and under-determined, i.e. finding the solution requires some type of mathematical regularization or prior information (Rodgers, 2000). The accuracy of the solution depends in general on the intensity of the spectral signature of the retrieved compound, the source of radiation (e.g. the Earth or the Sun), the observation geometry and the accuracy of the Radiative Transfer Modeling (RTM). The latter means also correctly accounting for all the atmospheric constituents or surface properties that affect the radiation field but are not retrieved themselves (auxiliary RTM inputs).

When the retrieval is done within a Bayesian framework, like the optimal estimation method (Rodgers, 2000), the measurements errors, the RTM errors and the uncertainty in the prior information (also named background or a-priori) are prescribed. The procedure provides then an estimation of the error covariance for the retrieved quantity and the Averaging Kernels (AK), which quantify the sensitivity of the retrieval to the true state and are linked to the Degrees Of Freedom (DOF) of the solution. The retrieval errors and [..[7]]the AK (or the DOF) can be used first to diagnose the quality and the relevance of the atmospheric retrieval. They become even more important when retrievals are further assimilated in numerical forecast models, because they weight the impact of the observations in the system.

[..[8]]Chemical Transport Models (CTM) [..[9]]solve the chemical and physical processes within the atmosphere but are based on meteorological fields [..[10]]from a Numerical Weather Prediction (NWP) model to advect the chemical species. Coupled Chemistry Meteorology Models (CCMM) that simulate both meteorology and chemistry online became available later but are today quite common in operational centers (Zhang et al., 2008; Flemming et al., 2015). There are currently growing efforts to introduce even stronger coupling of the atmosphere with both ocean and surface models, which gives so-called Earth System Models (ESM) (Brown et al., 2012; Hurrell et al., 2013). ESMs provide a comprehensive tool for climate predictions and reanalyses, but they are also considered for state-of-the-art air quality modeling (Neal et al., 2017).

Following closely the historical advances in modeling, the assimilation of satellite data has been introduced first in CTMs (Geer et al., 2006; Lahoz et al., 2007), and it is now well integrated also in operational CCMMs (Flemming et al., 2017). Today, numerous satellite retrievals of trace gases (e.g. O_3, CO, NO_2, CH_4, CO_2) and aerosols (AOD) are assimilated daily within operational CTMs and CCMMs (Inness et al., 2015; Bocquet et al., 2015). [..[11]]

Since long time, meteorological variables such as temperature and water vapor profiles are corrected by means of assimilating directly satellites radiances (Level 1 data) in NWP models. Therefore, the RTM became part of the observation operator of the assimilation system (Andersson et al., 1994). This [..[12]]avoided the introduction of biases in NWP that [..[13]]arised from poor prior information used in satellite retrievals at that time and neglecting the AK (Eyre et al., 1993). On the other hand, chemical species and aerosols are mostly corrected by means of assimilating geophysical retrievals (Level 2 or L2 data) that are

[7] removed: DOF
[8] removed: First atmospheric composition models were named
[9] removed: because they
[10] removed: computed separately with
[11] removed: However, some aspects of the Data Assimilation (DA) approach differ between the chemistry and meteorology communities.
[12] removed: resulted necessary to avoid
[13] removed: arise

made available by satellite data providers. To remove the impact of the prior information when assimilating L2 retrievals, the [..[14]]AK of the retrieval must be multiplied by the modeled profiles before computing the innovation vectors [..[15]](Rodgers, 2000; Eskes and Boersma, 2003; Migliorini, 2012). However, within standard methods based on the linearization of the RTM, like the optimal estimation, issues might still arise when the prior information used in the retrieval sits far from the true atmospheric state: this might challenge the linearization of the observation operator and result in sub-optimal retrievals. Since the AK themselves are also a result of the retrieval (and depend upon its prior information), we [..[16]]expect that a perfect removal of the prior information within DA cannot always be ensured.

The precise conditions that provide an equivalence between assimilating retrievals (using some kind of weighting functions) and radiances have been formalized by Migliorini (2012) and further tested by Prates et al. (2016) on synthetic satellite observations. These authors conclude that the equivalence holds under the hypothesis of an almost linear RT regime and with a careful selection of the prior error covariances in a way to maximize the measurements information in the retrieval step. Nonetheless, testing the two approaches within an operational system and with real observations remains crucial to verify if these conditions are met in practice. Moreover, the perfect equivalence holds only when all the auxiliary inputs of the RTM are exactly the same in both the retrieval and the radiance assimilation. It is clear that a climatological option for some RTM inputs will always be a more practical choice when computing L2 retrievals. On the other hand, the evolution towards strongly integrated ESMs will allow in principle to dispose of the most accurate prior information for all RTM inputs and favors the radiances assimilation approach. In this context, it appears important to introduce and evaluate the assimilation of radiances for chemical applications as well.

To the knowledge of the authors, the existent literature on this topic only concerned meteorological applications. Han and McNally (2010) explored the possibility of assimilating O_3 sensitive radiances within a NWP model but without comparing the two approaches. Similarly, Weaver et al. (2007) examined the assimilation of satellite radiances for aerosols but the focus was on the impact of using modeled aerosols microphysical properties as auxiliary input for the RTM and no comparison was provided. No other studies could be found concerning the assimilation of chemical compounds.

The objective of this study is to perform a first strict comparison between radiances and retrievals assimilation, with respect to O_3 estimation in the Thermal Infrared (TIR). To this end, systematic differences between the retrieval and radiances assimilation have been minimized as much as possible, for example by means of employing the same RTM within the two approaches.

We consider the case of O_3 assimilation using the Infrared Atmospheric Sounding Interferometer (IASI) onboard the European [..[17]]Metop satellites (Clerbaux et al., 2009). Several IASI O_3 retrievals have been already well validated (Dufour et al., 2012) and used directly to provide multi-annual time-series of the global O_3 budget (Wespes et al., 2016) or successfully assimilated within global (Peiro et al., 2018) and regional CTMs (Coman et al., 2012). However, an empirical correction of the retrievals has been found necessary to ensure globally unbiased reanalyses and slightly degraded assimilation results are still

[14] removed: Averaging Kernels (AK)
[15] removed: (Eskes and Boersma, 2003)
[16] removed: suppose
[17] removed: MetOP

found at mid and high latitudes (Emili et al., 2014). Since the tropospheric O_3 signature in the selected IASI spectral window decreases over colder surfaces, the impact of the retrieval's prior might become more relevant at high latitudes. In addition, the majority of IASI O_3 retrievals use a single a-priori profile globally (Barret et al., 2011; Boynard et al., 2016), which might present very large local departures from the true O_3 profile. Hence, IASI O_3 assimilation represents a good benchmark to evaluate the differences between retrievals and radiances assimilation.

The IASI-SOFRID O_3 product (Barret et al., 2011) and MOCAGE CTM have been used here to benefit from the experience of previous studies (Emili et al., 2014; Peiro et al., 2018). [..[18]]SOFRID and MOCAGE DA are based on a variational algorithm and, since SOFRID employs RTTOV (Saunders et al., 1999), which is a community RTM developed originally for NWP applications, the same RTM has been implemented in the MOCAGE [..[19]]system. Global O_3 [..[20]]analyses are computed for July 2010 and the results are compared against all available radio-soundings to evaluate their accuracy. Since the sensitivity of IASI TIR measurements to O_3 is not uniform along the atmospheric column, we also investigate the impact of assimilating more accurate stratospheric profiles from the Microwave Limb Sounder (MLS) in combination with IASI radiances. This might reveal possible synergies when assimilating multiple instruments that sense different layers of the atmospheres.

The paper is organized as follows. The satellite measurements, the Level 2 retrievals and the validation measurements used for this study are described in Sec. 2, as well as main steps concerning the preprocessing for some of the datasets. The chemical transport model, the radiative transfer model, the assimilation algorithm and the setup of the experiments are described in Sec. 3. The assimilation of IASI retrievals and radiances is compared in Sec. 4.1 and the impact of MLS assimilation on top of IASI is discussed in Sec. 4.3. The conclusions are summarized in the last section, where some recommendations are also given.

2 Measurements

2.1 IASI

[..[21]]IASI flies onboard the series of polar-orbiting satellites [..[22]]Metop operated by the EUropean organization for the exploitation of METeorological SATellites (EUMETSAT). It provides hyper-spectral measurements of the Earth's thermal radiation in the 3.62 - 15.5 μm (2760 - 645 cm^{-1}) window and serves meteorological and atmospheric chemistry applications [..[23]](Clerbaux et al., 2009; Hilton et al., 2012). IASI is an operational mission meant to provide long-term (> 20 years) time series of accurate TIR spectra at high spatial resolution. A total of three IASI instruments will be flying simultaneously at the end of 2019, providing nearly global coverage [..[24]]six times per day (morning and evening overpasses). Hence, they

[18] removed: Both
[19] removed: DA
[20] removed: reanalyses
[21] removed: The Infrared Atmospheric Sounding Interferometer (IASI)
[22] removed: MetOP
[23] removed: (Clerbaux et al., 2009)
[24] removed: three

represent a great opportunity for both NWP and climate-chemistry reanalyses. Only [..[25]]Metop-A data, available from 2008 to present, have been employed for this study.

2.1.1 L1 radiances

IASI L1c data contain calibrated and geolocalized spectra at 0.5 cm^{-1} spectral resolution (after apodization), i.e. [..[26]]8461 radiance values for each ground-pixel, with a footprint of 12 km for nadir observations. For this study, historical L1c data granules have been downloaded from the EUMETSAT Earth Observation data portal (https://eoportal.eumetsat.int) in NETCDF format. Data files contain also the observation geometry (sun and satellite angles) for each ground-pixel and the collocated land mask and cloud fraction values, obtained from the Advanced Very High Resolution Radiometer (AVHRR) measurements, also onboard [..[27]]Metop.

2.1.2 SOFRID L2 retrievals

The Software for a Fast Retrieval of IASI Data (SOFRID) was developed at the Laboratoire d'Aérologie to retrieve O_3 (Barret et al., 2011) and CO (De Wachter et al., 2012) profiles from IASI[..[28]]. It is based on the Radiative Transfer for TOVS (RTTOV) RTM (Saunders et al., 1999) and the 1D-VAR scheme developed within the Numerical Weather Prediction Satellite Application Facilities (NWP-SAF) program. SOFRID retrieves the O_3 profile in volume mixing ratio (vmr) units at 43 pressure levels between the surface and 0.1 hPa using 469 spectral channels within the main IASI O_3 window (980 - 1100 cm^{-1}). The choices that are made in SOFRID and are relevant for this study are summarized in Tab. 2. Note that a single a-priori profile and error covariance matrix are used globally and that the Surface Skin Temperature (SST) is estimated within the retrieval.

The number of DOF of the SOFRID retrieval has been evaluated between 2 and 3 for the full atmospheric column, with about one DOF for the tropospheric column (Dufour et al., 2012). SOFRID's averaging kernels corresponding to the retrievals used within this study (see Sec. 2.4) are shown in Fig. 1. We remark that the largest sensitivities are found for the lower stratosphere (50-100 hPa) and upper troposphere (200-300 hPa) levels. The sensitivity in the free troposphere (400-600 hPa) is maximum at tropical latitudes and decreases towards the poles due to the decreasing thermal contrast. Very low sensitivities are in general found for levels below 700 hPa at all latitudes.

The accuracy of the retrieved O_3 depends on the latitude and the vertical level, but sits generally within 10-20 % of the corresponding radiosoundings values, once the averaging kernels are applied. However, [..[29]]biases are found in the troposphere with SOFRID (+10%) and positive biases of about 15% are found in the Upper Troposphere - Lower Stratosphere (UTLS) region with all current IASI O_3 products (Dufour et al., 2012). The reasons for such biases are not yet fully understood and can impact negatively data assimilation (Emili et al., 2014) or trends analysis (Gaudel et al., 2018). This study will provide further insights about the impact of the constant a-priori on IASI retrievals.

[25] removed: MetOP-A
[26] removed: 8700
[27] removed: MetOP
[28] removed: in near-real time
[29] removed: increased

The SOFRID V1.5 retrievals described in Barret et al. (2011) are available for the full [..[30]]Metop-A period at http://thredds.sedoo.fr/iasi-sofrid-o3-co . The [..[31]]V3.0 version of SOFRID retrievals has been used for this study and have been obtained from [..[32]]B. Barret (personal communication). The [..[33]]main difference with version 1.5 concerns the temperature and water vapor profiles employed in the radiative transfer computations, which are taken from the ECMWF NWP [..[34]]model instead of EUMETSAT L2 retrievals[..[35]]. SOFRID v3.0 preprocessor retrieves the operational analysis (type "an") at 00-06-12-18 UTC from the ECMWF NWP model and assimilation system (IFS), regridded to a resolution of 0.25°x0.25° degrees. All the fields are then interpolated at the closest hour to the IASI pixel and a nearest neighbor interpolation is done to extract the corresponding profiles and surface properties. Since the CTM is also based on ECMWF NWP forcing fields (Sec. 3.1), this choice minimizes possible systematic differences between L2 retrievals and L1 assimilation. Also, SOFRID v3.0 is based on a more recent version of RTTOV and newer IASI coefficients (v11.1, coefficients on 101 levels) than the original L2 product (v9.0, coefficients on 43 levels). In addition to the O_3 retrieval and its error covariance, SOFRID files contain a number of auxiliary and diagnostic fields. [..[36]]The cloud fraction is based on a combination of EUMETSAT L2 product (AVHRR) and a Brightness Temperature (BT) analysis at 11 and 12 μm to fill pixels with missing AVHRR data (Barret et al., 2011)[..[37]]. An index based on V-shaped sand signature computed as $\Delta BT = (BT_{829cm^{-1}} - BT_{972.5cm^{-1}}) + (BT_{1202.5cm^{-1}} - BT_{1096cm^{-1}})$ is used to detect pixels affected by large aerosols load. Usage of these products will be detailed in the data preprocessing section (2.4).

2.2 MLS L2 retrievals

Since 2004 The Microwave Limb Sounder (MLS) flies on-board the research mission AURA and measures thermal emission at the atmospheric limb (Waters et al., 2006). It provides about 3500 stratospheric profiles of multiple atmospheric constituents each day, including O_3 (Froidevaux et al., 2008). Since the version 3 of MLS products, O_3 profiles are retrieved on 55 pressure levels with a recommended range for scientific usage between 0.02 and 261 hPa for version 4.2 (Livesey, 2018). The biases of MLS O_3 profiles are typically within 5% with respect to ozonesondes and lidar measurements (Hubert et al., 2016), with slightly higher values below 200 hPa. Given its good accuracy, MLS O_3 has been widely used both for trend analysis (Froidevaux et al., 2015) and assimilation experiments (Massart et al., 2010; Miyazaki et al., 2012; Inness et al., 2015). Similarly to previous studies (Emili et al., 2014), we retain only the most accurate data using MLS, i.e. above 170 hPa. The MLS V4.2 product used in this study has been downloaded from the Goddard Earth Sciences Data and Information Services Center (GES DISC) web portal (https://disc.gsfc.nasa.gov).

[30] removed: MetOP-A
[31] removed: V1.6
[32] removed: LA
[33] removed: only
[34] removed: forecasts (V1.6)
[35] removed: (V1.5) .
[36] removed: We considered in particular the SOFRID cloud fraction , based on
[37] removed: , and an

2.3 Radiosoundings

Ozonesondes are launched on weekly bases by meteorological services and provide accurate profiles of O_3 up to 10 hPa with a vertical resolution of 150-200 m. ECC type sondes, which represent the largest percentage of the global network, have a precision of about 5% (Thompson et al., 2003). Radiosoundings are relatively sparse and their geographical distribution is much more representative of the northern mid-latitudes. However, they provide since several decades the most precise information on vertical ozone distribution in the troposphere. Therefore, they have been used to derive widely used tropospheric O_3 climatologies (McPeters et al., 2007) and validate both satellite products (Dufour et al., 2012) or models (Geer et al., 2006). They will be used in this study to validate all model simulations. Data are collected and distributed by the World Ozone and Ultraviolet Radiation Data Center (WOUDC, http://www.woudc.org).

2.4 Data preprocessing

Some further preprocessing has been applied to the original L1c and SOFRID datasets to ease the interpretation of the assimilation experiments presented later in Sec. 4. The objective was to ensure that exactly the same spectra are used for both L1 and L2 assimilation.

Only the spectral channels that are used in SOFRID are extracted from IASI L1c granules, i.e. channel n. 1350 (980 cm^{-1}) to 1818 (1100 cm^{-1}). Some further screening is applied to remove channels that are affected by strong H_2O absorption, as also done in SOFRID.

The spatial resolution of the CTM (2°x2° degrees, Sec. 3.1) is much coarser than IASI pixels size. Since it is preferable to avoid all kind of spatial averaging of the observations, a significant reduction of ground-pixels is needed. In return, we employ strict selection criteria to avoid as much as possible contamination from clouds and bright surfaces, which [..[38]]reduces the RT accuracy and increase retrieval or assimilation errors. The data selection is performed as follows.

First, only L1 pixels with both IASI and AVHRR highest quality flags are kept. Then, ground-pixels from IASI L1 and SOFRID products are filtered using their respective cloud masks (Sec. 2.1.1 and 2.1.2) and keeping only pixels with cloud fraction less or equal to 1%. SOFRID pixels with a sand signature greater than 0.5 and with a number of retrieved levels lower than 35 (mountains) are also filtered out. Resulting datasets are then matched, i.e. only common ground-pixels that remained available after the previous L1 and SOFRID independent selections are kept. Finally, a data thinning is performed [..[39]]using a regular grid of 1°x1° resolution and keeping only the first pixel that falls in every two grid boxes. This ensures a minimum distance of 1° among assimilated observations. After the completion of the data selection procedure the final number of retained ground-pixels for L1 and SOFRID is about [..[40]]5000 per day, compared to about 10^5 when only the cloud screening is applied. The total number of L1/L2 observations resulting from the above selection and further assimilated in this study is displayed in Fig. 2.

[38] removed: reduce
[39] removed: to retain a maximum of about two pixels for each model grid point.
[40] removed: 3300

3 Method

This section summarizes the main characteristics of the CTM (3.1), the RTM (3.2) and the assimilation algorithm (3.3) used in this study. Further details on the particular selection of the main parameters of the assimilation experiments (e.g. the error covariances) are given in Sec. 3.4.

3.1 Chemical transport model

The Chemical Transport Model (CTM) MOCAGE (Josse et al., 2004) is used in this study. A global configuration with an horizontal resolution of 2°x2° degrees and 60 hybrid sigma-pressure levels up to 0.1 hPa has been used. The vertical resolution varies from about 100 m in the planetary boundary layer to about 700 m in the upper troposphere, decreasing further to approximately 2 km in the upper stratosphere. Chemical mechanism, emissions and physical parameterizations follow the setup used for operational air-quality forecasts (Marécal et al., 2015), which includes about 100 species and 300 chemical reactions. A similar configuration has been employed by Barré et al. (2013) to assimilate IASI O_3 columns over Europe, but with a lower model top at 5 hPa. Other authors favored a simplified chemistry scheme but with a model top at 0.1 hPa to assimilate satellite O_3 products globally (Emili et al., 2014; Peiro et al., 2018).

We considered for this study the highest available model top because we need to simulate the full atmosphere to compute radiances. In addition, the 0.1 hPa top matches with the vertical grid used for SOFRID retrievals (Sec. 2.1.2), making the comparison of the two assimilation approaches (radiances vs L2) stricter. The full chemical scheme is chosen instead of a simplified chemistry to reduce as much as possible biases of the modeled O_3 in the troposphere. The main intent of this study is in fact to evaluate the impact of a dynamical and accurate O_3 prior on assimilation results.

The meteorological forcing comes from the ECMWF [..[41]]IFS, from which we [..[42]]retrieved the forecast (type “fc”) initialized with the [..[43]]analysis at 00 [..[44]]UTC of each day. The NWP fields are interpolated on the horizontal grid of MOCAGE (2°x2° degrees) during the retrieval process and stored with a timestep of three hours. During the integration of MOCAGE, the meteorological forcing is linearly interpolated at the advection time step of MOCAGE (hourly) and on the CTM’s vertical grid.

3.2 Radiative transfer model

RTTOV (Saunders et al., 1999) is a community RTM developed for operational NWP models. One of its main advantages is computational efficiency, which is achieved by running accurate but costly line-by-line RT simulations for a large number of satellite sensors, observation geometries and atmospheres and storing the corresponding coefficients in large look-up tables. RTTOV provides API interfaces for the direct RT computations plus the tangent linear and adjoint model, which are needed in variational assimilation systems.

[41] removed: NWP model (IFS)
[42] removed: selected the forecast steps
[43] removed: latest available analysis (
[44] removed: or 12 UTC).

Version 11.3 of RTTOV (Saunders et al., 2013) has been used in this study for the L1 assimilation. This version includes coefficients for the IASI TIR channels computed using a fine atmospheric grid ([..[45]]101 vertical levels). The SST, 2 m temperature, 2 m pressure and 2 m wind vector are taken from high resolution (0.125°x0.125° degrees) global IFS forecasts initialized [..[46]]with the analysis at 00 UTC of each day and collocated (nearest neighbor) with satellite ground-pixels prior to data assimilation. A linear interpolation from the 3-hours forecast steps to the closest hour of the IASI observations is also perfomed for these fields. The surface emissivity is based on the RTTOV monthly TIR emissivity atlas (Borbas and Ruston, 2010). Only clear-sky RT computations are performed for this study and no aerosols have been prescribed. The RTM configuration is summarized in Table 2. Due to the different processing chains, the auxiliary inputs of RTTOV could not be set exactly equal for L1 assimilation and L2 retrievals (Tab. 2). The potential impact of these residual differences is discussed in Sec. 4.1.

3.3 Assimilation algorithm

The assimilation suite for MOCAGE is based on a variational algorithm and was developed initially within the ASSET (Assimilation of Envisat data) project (Lahoz et al., 2007). The objective was to assimilate satellite products at a global scale and a 3D-FGAT implementation was chosen. It evolved later to provide air-quality reanalyses at the surface based on a 3D-Var implementation (Jaumouillé et al., 2012) and was extended to 4D-Var [..[47]]when employing linearized chemistry schemes (Massart et al., 2012; Emili et al., 2014). In all cases the minimization of the variational cost function is performed using the limited-memory BFGS algorithm (Liu and Nocedal, 1989). We used in this study a 3D-Var algorithm with hourly assimilation windows and with O_3 as control variable.

The 3D background error covariance is modeled through a diffusion operator (Weaver and Courtier, 2001) and allows the specification of heterogeneous correlation length scales. Compared to previous studies using MOCAGE assimilation suite, a new vertical correlation operator has been employed here: the vertical error correlation is now assigned by explicitly filling a positive definite matrix using the gaussian formulation of Paciorek and Schervish (2006) and by numerically computing its square root. This avoids difficulties encountered with diffusion based operators concerning the normalization in presence of boundaries (e.g. the surface) and heterogeneity (Mirouze and Weaver, 2010). Since the vertical dimension of the model grid is relatively small, this choice does not impact significantly the numerical cost and the memory requirements with respect to the previous implementation based on diffusion.

The observation operator of MOCAGE allows to assimilate a large number of measurements, spanning from columns of gases (Massart et al., 2009) to aerosol optical depth (Sič et al., 2016). Next, we give some details of the implementation used in this study to assimilate vertical profiles and radiances.

After the horizontal and temporal interpolation of the model fields at the satellite ground-pixel position, modeled profiles are linearly interpolated to the retrieval's vertical grid. When the averaging kernels are used (i.e. for SOFRID assimilation),

[45] removed: 104
[46] removed: from ECMWF 4D-Var
[47] removed: in case of

the linear estimation equation (Barret et al., 2011) is used to remove the impact of the prior from the innovation vector. The ensemble of these operations is stored as coefficients of a large sparse matrix and done through its multiplication by the model 3D field. This approach is practical since numerous application of the linearized and adjoint operator are needed during the minimization of the variational cost function. Differently from all previous studies involving IASI O_3 assimilation (Massart et al., 2009; Emili et al., 2014; Peiro et al., 2018), where L2 profiles were first reduced to total or partial columns prior to assimilation, we assimilate here directly the full L2 profiles (43 levels). This avoids any loss of information and allows a fairer comparison between L2 and radiances assimilation. The error covariance matrix of the profile-type observations is diagonal in the latitude/longitude dimensions but off-diagonal terms are allowed along the vertical dimension. Alternative approaches exist to optimally reduce the dimension of the L2 observation space based on the DOF of the retrievals (Migliorini et al., 2008; Mizzi et al., 2016), which are of interest to further reduce the numerical cost of SOFRID assimilation without loss of accuracy. However, this is left for future work.

[..[48]]To compute modeled radiances we employ the same horizontal and temporal interpolation as in the case of profile observations, except for the vertical interpolation. In fact, the RTTOV vertical interpolator is used for radiances computations instead of the MOCAGE one. All model levels (60) and corresponding levels pressure are given as input to RTTOV, which performs internally the vertical interpolation to the IASI coefficients levels. Since the model vertical resolution is lower than the one available in RTTOV for IASI coefficients ([..[49]]101 levels), we used the default option based on Rochon et al. (2007). Also, O_3 profiles above the CTM top (0.1 hPa) are completed using RTTOV climatological [..[50]]profiles. Auxiliary inputs for the radiances computation include the pressure, temperature and water vapor profiles, which are interpolated from the correspondent MOCAGE fields.

The MOCAGE control vector has been extended to include the SST, as [..[51]]in SOFRID retrieval scheme. This proved to be important since small errors in the SST translate in significant differences between [..[52]]modeled and measured radiances. Not accounting for this would produce wrong O_3 analyses. The SST does not belong to the MOCAGE prognostic fields nor it is prescribed on the MOCAGE grid. Hence, the SST analysis is not propagated in time and no spatial covariance model have been implemented so far. In this [..[53]]regard the treatment of the SST is equivalent to that done in L2 retrievals. In the context of MOCAGE 3D-Var, it can be interpreted as a variational bias correction term in the observation space (Dee and Uppala, 2009), with prior values given by the NWP model (IFS, see Sec. 3.2).

3.4 Setup of the experiments

We performed numerical experiments for the month of July 2010, which corresponds to the typical presence of summer O_3 maxima in the northern hemisphere linked to photochemical pollution. July 2010 is also interesting due to the development

[48] removed: The steps for the computation of modeled radiances are equal to the profiles ones until the
[49] removed: 104
[50] removed: profile
[51] removed: it is done within
[52] removed: modelled
[53] removed: sense

of a strong La Nina episode (Peiro et al., 2018). The main difference between assimilating L2 and L1 data consists in using a climatological (L2 assimilation) versus a dynamical a-priori (L1 assimilation) for the inversion of the radiative transfer problem. The chosen period presents large local deviations of the O_3 field from climatological values. Therefore, it provides an interesting benchmark period with respect to the objective of this study.

The CTM has been initialized on [..[54]]1 June 2010 with a zonal climatology and run for one month period (spinup) to provide chemically balanced initial condition on [..[55]]1 July 2010 for all simulations.

The observation error covariance matrix ($\mathbf{R}$) is prescribed according to the choices adopted in SOFRID [..[56]]V3.0. When the radiances are assimilated, a diagonal matrix (i.e. with no inter-channel correlation) is used with a constant standard deviation of to 0.7 mW m^{-2} sr^{-1} cm for all channels. This is a simplified although common setting for most IASI O_3 retrievals (Barret et al., 2011; Boynard et al., 2016). The SST, which is controlled as well within radiances assimilation, has a prescribed standard deviation of 4° C for all ground-pixels. When L2 profiles are assimilated we used the full non-diagonal error covariance matrix provided by SOFRID or MLS retrievals.

We considered a dynamical rejection of observations based on the relative differences between simulated and measured values with respect to simulated values. It avoids assimilating observations with too large departures from corresponding model background. The thresholds values are set to 12% for L1 radiances and 2000% for L2 profiles and trespassing the threshold for any particular channel or profile level rejects the entire spectrum or profile. The strong difference between the two thresholds is a consequence of the very different nature of assimilated observations: the exponential shape of O_3 profiles can produce very large departures where the gradient is the steepest (tropopause) and a small rejection threshold would filter out most of the profile observations. This is not the case for radiances, which vary on a linear scale. Thresholds values have been chosen based on misfit histograms in a way to remove abnormal tails. As a consequence, L1 and L2 pixels that pass the selection and are further assimilated could differ. However, the relative number of rejected observations for the entire month of July is quite limited in both cases (3% for L1, 6% for L2), thus not affecting statistically the results.

The setup of the background error covariance ($\mathbf{B}$) is a critical step both for L2 retrievals and data assimilation. [..[57]]We did benefit from past experiences using MOCAGE, IASI and MLS O_3 (Massart et al., 2012; Emili et al., 2014; Peiro et al., 2018) to define a first guess of $\mathbf{B}$ and we tried to further derive an optimal parameterization for [..[58]]this study. Note that the $\mathbf{B}$ matrix (3D) used in data assimilation is by definition different with respect to the one specified within SOFRID (1D), but the same 3D $\mathbf{B}$ is used for all data assimilation experiments (L1 and L2).

Concerning the standard deviation, Emili et al. (2014) and Peiro et al. (2018) employed vertically varying errors expressed as percentage of the background O_3 profile[..[59]], with larger relative errors in the troposphere [..[60]]and smaller in the

[54] removed: 1st
[55] removed: 1st
[56] removed: V1.6
[57] removed: For this study we could
[58] removed: $\mathbf{B}$ based on previous results
[59] removed: . Optimal results were found setting a value of 5% in the stratosphere and 30% in the troposphere, with the tropopause being arbitrary set at about 150 hPa. Peiro et al. (2018) kept the same error parameterization but reduced the errors to 15%
[60] removed: to analyze the tropical O_3 distribution

stratosphere. Since we use here a more detailed chemistry model (Sec. 3.1) we [..[61]]evaluated the Root Mean Square Error (RMSE) of [..[62]]the free model simulation (control) against [..[63]]ozonesondes and MLS profiles (Fig. [..[64]]3). We remark that the model's RMSE reproduces the vertical features observed in previous studies, with smaller errors in the stratosphere (between 20 and 50 hPa), larger errors in the free troposphere, and highest errors close to the tropopause and within the planetary boundary layer. Note also the zonal variability of the maxima, which appear linked to the variability of the tropopause height. [..[65]]Thanks to the detailed chemical mechanism, biases ([..[66]]Fig. 4) are generally smaller than in the studies cited previously [..[67]]but remain significant compared to standard deviation values (Fig. 5), especially around the tropopause. Interestingly, RMSE and standard deviation values computed against MLS are generally smaller than those evaluated against ozonesondes whereas biases are more consistent between the two datasets. We attribute this effect to the larger number of MLS observations (Tab. 1) which provides more robust standard deviation statistics.

The background standard deviation is prescribed through a smooth step function that takes values of 2% above 50 hPa and 10% below [..[68]]to reproduce roughly the patterns observed in Fig. 5. Values are smaller than those in Fig. 5 to account for the error reduction during the assimilation, which is particularly strong when MLS observations are used (see Sec. 4.3). Also, neglecting error correlations between IASI channels within $\mathbf{R}$ leads to a strong weight towards the observations: reducing the background standard deviation compensates in part this effect. All the choices made to define $\mathbf{B}$ are a result of a large number of assimilation evaluations, where different options were considered. For example, setting values of 5% and 25% lead to less accurate results for both L1 and L2 assimilation (not reported). The percent profile is multiplied by the hourly O_3 field of the control simulation once for the entire period and not at every forecast time step. Therefore, all assimilation experiments presented in this study are based on the same $\mathbf{B}$ matrix. This choice has been taken to permit a stricter comparison between L1 and L2 assimilation experiments.

The vertical error correlation diffuses the assimilation increments between model levels and has been found to significantly impact the quality of O_3 [..[69]]analyses with current model vertical resolutions (not shown). In general, small values of vertical correlation are favored [..[70]]in the stratosphere due to the stratification and to avoid injection of large stratospheric O_3 increments in the troposphere[..[71]], whereas larger values are expected within the troposphere due to vertical mixing. In

[61] removed: first
[62] removed: a
[63] removed: the ozonesondes
[64] removed: **??**
[65] removed: However, thanks
[66] removed: not shown
[67] removed: .
[68] removed: . These values are slightly smaller than in previous studies (Emili et al., 2014; Peiro et al., 2018) because of the smaller biases of the forecast model. Also, the
[69] removed: reanalyses
[70] removed: to
[71] removed: . For example, Emili et al. (2014) used a constant correlation length

this study a constant value of 1 model [..[72]]level defines the length scale of the Gaussian correlation (Sec. 3.3). Different choices for the stratosphere and troposphere did not lead to particular improvements (not shown).

Finally, the exponential scale of the horizontal error correlation is set equal to 200 km, with the zonal component that is reduced towards the poles to account for the increasing resolution of the model's grid (Emili et al., 2014).

The choice of the background and observation errors is relatively simplistic in this study. Further improvements of the $\mathbf{B}$ parameterization could be achieved by diagnosing the forecast errors hourly (Desroziers et al., 2005) or using ensembles of model forecasts. However, [..[73]]more complex and costly estimations do not always improve systematically and significantly the results of chemical assimilation (Massart et al., 2012). Moreover, a good estimation of $\mathbf{B}$ cannot be done independently from that of $\mathbf{R}$, which is kept fixed here on purpose. Additional research is needed in this regard, which is out of the scope of this study.

4 Results

A total of six simulations for the month of July 2010 have been performed (Tab. 3), starting on 1st July: a free model simulation (control) and five 3D-Var [..[74]]analyses assimilating respectively SOFRID L2 profiles (named L2a), IASI L1 radiances (L1a), MLS L2 profiles (MLSa), MLS plus SOFRID L2 profiles (MLS+L2a), MLS plus L1 radiances (MLS+L1a). The first three simulations (control, L2a and L1a) are discussed in Sec. 4.1. The control simulation and the three [..[75]]analyses that include MLS are discussed in Sec. 4.3. All simulations have been validated against ozonesondes profiles to elucidate the differences of the resulting O_3 vertical distribution. A total of [..[76]]219 radiosoundings are available globally in July [..[77]]2010 (Tab. 1). The colocation of ozonesondes profile with model fields in time and space is performed through the MOCAGE observation operator (Sec. 3).

4.1 IASI assimilation

[..[78]]The average O_3 values of the control simulation [..[79]]are displayed in Fig. 6. O_3 fields have been first interpolated vertically from the model grid to a selection of pressure levels, covering both the stratosphere and the troposphere, and averaged afterwards. The maps show well known properties of the O_3 distribution such as the strong zonal gradients in the

[72] removed: grid point; Peiro et al. (2018) found that switching off the vertical correlation provided even better results for MLS analyses. However, a non-zero correlation seems more appropriate for generic usage, because it allows to assimilate effectively also point measurements. Second, the SOFRID prior covariance is also non-diagonal (Barret et al., 2011) and it is better to preserve a certain consistency between the two approaches. Therefore, we used the value of 1 model grid point in this study.

[73] removed: such

[74] removed: reanalyses

[75] removed: reanalyses

[76] removed: 220

[77] removed: 2010.

[78] removed: We discuss the geographical differences between L1a and L2a reanalyses by looking at the monthly bias between the two experiments, divided by the

[79] removed: . To this end

stratosphere and the presence of local minima in the tropical free troposphere due to the deep convection. The average difference between the control simulation and the fixed a-priori used in SOFRID retrievals (Sec. 2.1.2) is displayed in Fig. 7. Large differences (>100%) are found at low latitudes both in the lower stratosphere and in the troposphere, with largest values close to the tropical tropopause (150 hPa). This is expected since SOFRID a-priori is based on a global ozonesondes climatology that is more representative of mid-latitudes O_3 profiles (Sec. 2.1.2).

We discuss the geographical differences between L1a and L2a analyses by looking at the monthly bias between the two experiments, divided by the average O_3 of the control simulation (Fig. 6). Relative differences are displayed in Fig. 8. First, we remark that differences are [..[80]]generally significant both in the stratosphere and in the troposphere, with absolute values that can exceed 50% of the O_3 field locally and global averages [..[81]]as high as to 20%. Largest differences in the stratosphere are found [..[82]]at tropical latitudes, L1a showing larger O_3 values than L2a at 20 hPa and lower at 70 hPa. In the troposphere the strongest positive differences are still found in the tropics, especially over central Africa[..[83]], Eastern Asia, [..[84]]South America and Middle-East regions. Differences become smaller when moving down to 750 hPa and tend to disappear at lower altitudes (not shown), which is normal considering the vertical sensitivity of IASI. [..[85]]At mid and high latitudes, relative differences are smaller than [..[86]]at the tropics. This behavior [..[87]]is consistent with the fact that the SOFRID prior is [..[88]]much less accurate for tropical [..[89]]latitudes than for mid and high latitudes (Fig. 7). Overall, these plots suggest that [..[90]]when the L2 a-priori is strongly biased (>100%), the equivalence between L1 and L2 assimilation in the thermal infrared is not verified for O_3, even when the averaging kernels are employed.

To confirm that the observed differences are not a consequence of the slightly different NWP inputs used in L1 assimilation and L2 retrievals (Tab. 2) we rerun the L1a simulation using exactly the same SST a-priori values used in SOFRID retrievals. Doing the same for the temperature and water vapor profiles being technically more complex, we followed a different path and repeated all assimilation experiments but using ERA interim (Dee et al., 2011) instead of the NWP forecasts as meteorological forcing for the CTM. This increases potential differences between the L1 and L2 assimilation, due to the different configurations of operational NWP and ERA interim (model resolution, assimilated instruments etc.). In all above cases we obtained very similar results to those presented in Fig. 8 (not shown), which suggests that differences between L1a and L2a discussed previously do not depend on the auxiliary RTM inputs.

To further verify which one, between the L1a and L2a experiments, reproduces better the measured O_3 profiles, we validated the three simulations against radiosoundings. Figure 9 reports the RMSE differences computed globally and for five

[80] removed: in
[81] removed: that are between 1%and 15%
[82] removed: in correspondence of
[83] removed: and
[84] removed: but significant negative differences appear in the Southern Hemisphere (SH) mid latitudes
[85] removed: More remarkably, in the Northern Hemisphere (NH)
[86] removed: elsewhere
[87] removed: seems coherent
[88] removed: more representative of the NH mid-latitudes (Sec. 2.1.2) and
[89] removed: and SH latitudes .
[90] removed: the

different latitude bands. The displayed values are the differences between the RMSE of the assimilation experiment and the corresponding value for the control simulation (Fig. [..[91]]3). Negative values in Fig. 9 indicate that the assimilation improved the O_3 field and decreased the relative RMSE with respect to ozonesondes by the amount displayed on the plot. Looking at the global averages we remark that below 70 hPa the gain is similar for both L1a and L2a experiments, and quite significant at 200 hPa (20%). Note, however, the strong similarity between the global and 30°N-60°N statistics, due to the over-representation of ozonesondes for NH mid-latitudes (63% of the total).

In the NH the RMSE of the control simulation is effectively reduced between 70 and 300 hPa (up to 20%). L1a shows a slightly better gain than L2a between 150 and 300 hPa. Interestingly, both L1a and L2a display increased RMSE between 300 and 400 hPa. This behavior is also confirmed when the vertical error correlation is switched off in the 3D-Var $\mathbf{B}$ and with different choices for the vertical interpolation of O_3 optical coefficients within RTTOV (log-linear or Rochon, not shown). Since large negative biases were present in the control simulation (as low as -30%, [..[92]]see Fig. 4), a possible explanation is that part of the strong positive correction of O_3 between 100 and 300 hPa is propagated downwards, where both absolute O_3 concentrations and relative biases are much lower. This can degrade the [..[93]]analysis accuracy below 300 hPa. Whether this propagation is carried out by the Jacobian matrix of the observation operator (either through the RTM or the retrieval's AK) or by vertical O_3 transport is not yet elucidated and would need further investigation. Also, other possible factors affecting the accuracy of the RTM exist, like inadequate vertical resolution close to the tropopause, uncertainties in meteorological profiles or impact of aerosols. Nonetheless, these errors impact both L1a and L2a in our study: [..[94]]further optimization of the L1 assimilation configuration with respect to the L2 retrievals is left for a future study. The RMSE is reduced again at about 500 hPa between 30°N-60°N, although not very significantly. The assimilation increases the RMSE of the tropospheric profile (350-1000 hPa) at northern latitudes (60°N-90°N). In general, the validation confirms that L1a and L2a have a very similar accuracy in NH at mid and high latitudes, as also suggested previously by Fig. 8. However, the strongest positive corrections are confined to the UTLS.

At the tropics (30°S-30°N) the results differ more significantly. In the troposphere (below 100 hPa), both L1a and L2a reduced the RMSE of the control simulation, although by a smaller amount than in NH (5%). Note also that L1a RMSE reduction is larger than L2a between [..[95]]400 and 600 hPa, whereas it is the other way [..[96]]around at about 250 and 800 hPa. Above 100 hPa we observe an increase of RMSE that peaks at 60 hPa with L2a and at 30 hPa with L1a, but smaller in magnitude for L1a. This behavior might be linked to the strong differences that exist between the SOFRID prior and the modeled O_3 at the tropical tropopause, to some other factor affecting the RT computations, to overestimation of the background error covariances or to a complex combination of all previous causes. A full satisfactory explanation has not been found yet.

[91] removed: **??**
[92] removed: not shown
[93] removed: reanalysis
[94] removed: a more profound revision
[95] removed: 300 and 500
[96] removed: round at about 600

Results in the SH (30°S-90°S) are [..[97]]again similar: lower RMSE than for the control simulation is found for both L1a and L2a in the upper and lower stratosphere (between 30 hPa and 100 hPa [..[98]]and between 150 and 300 hPa). An improvement is also found in the troposphere ([..[99]]400-600 hPa) at mid-latitudes [..[100]](30°S-60°S), but the low number of ozonesondes available in this band (Tab. 1) requires a more careful interpretation.

Since radiosoundings do not provide a uniform global coverage and vertical coverage also lacks in the vicinity of the O_3 maximum, we validated the three simulations against MLS measurements. The RMSE differences for stratospheric profiles can be found in Fig. 10. These statistics are based on more than 10^5 profiles for the global average and between 15000 and 30000 for zonal averages, depending on the latitude band [..[101]](Tab. 1). The patterns observed in the stratosphere with respect to ozonesondes are confirmed also with MLS. The only exceptions are a [..[102]]smaller RMSE degradation at 50 hPa for L2a in the tropics and for both L1a and L2a at 150 hPa in the 30°S-60°S band. [..[103]]Higher confidence should be given to the RMSE values provided by MLS than those obtained with radiosoundings (see also Fig. 3 and the relative discussion in Sec. 3.4). However, a similar RMSE behavior is observed overall and this bolsters the robustness of the conclusions derived [..[104]]with the radiosoundings in the troposphere.

4.2 Computational cost

The computational cost of L1 assimilation is necessarily higher than for L2 assimilation. Additional CPU time is due not only to online RTM computations but also to a higher number of iterations needed by the minimizer to converge. For a typical 24 hours long simulation performed on Intel Xeon E5-2680 V3 CPU the total [..[105]]CPU time is 3.9 CPU hours for L2a and 13.2 hours for L1a. Note that the L2a time does not include the cost of the L1 to L2 processor but only the cost of the 3D-Var assimilation plus the model forecasts. Most of the CPU time for L1a is spent in the linearized and adjoint calls of the RTM (50% of the total CPU time), whereas the corresponding time spent for the observation operator within the L2a experiment is about 1%. However, the total CPU time can be significantly decreased by reducing the maximum number of iterations of the minimizer. A simulation with halved number of iterations (75) showed very similar results to the ones that have been reported (150 iterations) and could be considered if computation time is a critical factor. Moreover, with standard high performance computers and thanks to the parallel nature of the observation operator and the RTM, we could obtain a speedup of about 24

[97] removed: in favor of L1a

[98] removed:), with L1a slightly better at polar latitudes (60°S-90°S). More noticeably, L2a is equal or worse than the control simulation

[99] removed: below 250 hPa) , whereas L1a improves the RMSE. The SOFRID prior is biased towards NH

[100] removed: , where tropospheric O_3 concentrations are generally the highest. The sensitivity of IASI TIR channels to tropospheric O_3 decreases over colder surfaces (e.g. in the SH during July). Hence, we expect a stronger impact of the prior in the retrieval results, which can be detrimental if the prior is biased. Indeed, we note that L1a remains close to the (already accurate) control profiles, whereas L2a adds a positive bias (not shown). Such behavior was already diagnosed by Emili et al. (2014) when assimilating SOFRID partial columns and we provide here a possible explanation. Using a more adapted prior in the SH could in principle also improve L2 retrievals themselves, which seems the case with a newer versions of SOFRID (B. Barret, personal communication).

[101] removed: .

[102] removed: lower

[103] removed: Overall, the validation against MLS

[104] removed: previously for the

[105] removed: computing

on the 24 CPU cores. This reduces the run time of L1a to about 36 minutes for the 24 hours-long simulation, versus 13 minutes for L2a. The extra cost of L1 assimilation seems therefore acceptable also for operational applications.

4.3 IASI and MLS assimilation

Some issues were identified in the previous section in the stratosphere, especially at tropical latitudes. Among possible reasons, one is that inversion of TIR measurements might be particular sensitive to the vertical distribution of O_3 in the tropical stratosphere. We consider here assimilating MLS L2 profiles on top of IASI to correct the model stratosphere and troposphere simultaneously, as done also in previous studies (Emili et al., 2014; Peiro et al., 2018). When the radiances are assimilated, the RT problem is solved for the entire atmospheric column within the iterations of the variational algorithm. Therefore, enhanced and better synergies could be observed than when only L2 products are assimilated.

We report in Fig. [..[106]]11 the impact of assimilating MLS [..[107]]on top of IASI L1/L2 by computing the average differences between MLS+L1a [..[108]]and MLS+L2a[..[109]]

[..[110]]. We remark that the differences in the stratosphere are highly reduced with respect to Fig. 8, which is expected due to the [..[111]]direct constraint of MLS observations. Significant differences (> 10%) [..[112]]remain below 150 hPa, with patterns and sign similar to those in Fig. 8. The amplitude of the differences is, however, slightly reduced also at 300 and 500 hPa.

We compared the RMSE of MLSa, MLS+L1a and MLS+L2a computed against ozonesondes (Fig. 12) to evaluate if [..[113]]the joint assimilation improves the overall O_3 distribution. MLSa provides particularly accurate results down to 200 or 300 hPa, depending on the latitude, with a robust reduction of the RMSE with respect to the control simulation. The only exception is in the SH mid-latitudes below 250 hPa, where the MLSa RMSE increases. We suspect that this might be linked again to the combination of strong O_3 gradients at the tropopause height and the negative bias of the control simulation above the

[106] removed: **??**

[107] removed: alone by computing the average difference between MLSa and the control (upper plots), and the impact of assimilating MLS

[108] removed: (

[109] removed:) and L1a (L2a) respectively. As expected, the impact of MLS assimilation is very significant at 100 hPa (relative differences as high as 60% of control O_3) but becomes minor at 500 hPa, where no direct constraint exists from the observations. Interestingly there are regions at mid-latitudes where the impact of MLS is not negligible (> 5%). Since the 3D-Var increments are confined to higher levels (above 200 hPa), we reckon that the impact of MLS assimilation at 500 hPa is due to the model dynamics at mid-latitudes, e. g. Stratosphere-Troposphere Exchanges (STE).

[110] removed: When comparing the MLS impact at 500 hPa with the bottom plots we remark that there is no sign of a strong spatial correlation in the NH and in the tropics. This suggests that the impact of MLS in the troposphere is supplanted by IASI assimilation (either L1 or L2)

[111] removed: strong sensitivity of IASI TIR measurements at 500 hPa. Traces of superposition of the MLS impact on IASI reanalyses appear in SH mid-latitudes, which is coherent with the fact that the IASI tropospheric impact is smaller over colder surfaces (Sec. 4.1). In case of no synergy between MLS and IASI we would expect to see in bottom plots either very small values or patterns similar to what observed in the SH mid-latitudes. Instead, significant differences (as high as

[112] removed: arise at tropical latitudes, which are also opposed in sign , i. e. a positive feedback of MLS isobserved within MLS+L2a, both negative and positive, but smaller in amplitude, within MLS+L1a. This confirms that constraining the model with MLS above 200 hPahas a significant impact on the free troposphere when assimilating IASI

[113] removed: some of the observed feedbacks improve the

tropopause (see Sec. 4.1). Overall MLSa confirm results found in past studies (Massart et al., 2012; Emili et al., 2014) and represents a much better prior for assimilation of radiances or retrievals.

We remark that MLS+L1a and MLS+L2a provide now closer results in the NH and in the tropics compared to Fig. 9. [..[114]]The stratospheric O_3 gain is much more significant with MLS+L1a/MLS+L2a than with L1a/L2a and remains very close to MLSa, demonstrating that assimilating accurate stratospheric profiles remains essential for O_3 reanalyses. [..[115]]The only region where IASI further improves the UTLS profile with respect to MLSa is in the NH: a positive, albeit small, effect of assimilating IASI on top of MLS is found between 150 and 300 hPa. [..[116]]On the other hand, below 300 hPa, the addition of MLS (Fig. 12) does not bring further improvements with respect to IASI alone [..[117]](Fig. 9). We can conclude that MLS corrects most of the errors introduced by IASI assimilation in the stratosphere (Fig. 9) but no particular synergy is observed in the case of MLS and L1 assimilation in the troposphere.

We report in Fig. 13 the Taylor plots concerning the free troposphere O_3 column (340-750 hPa), to further evaluate the skills of the assimilation experiments in terms of variability. We examine here the free troposphere since it is where the direct impact of IASI assimilation is the largest and the impact of MLS the smallest (except for the 30°S-60°S band). IASI assimilation improves the variability of the modeled O_3 field when looking at global averages, but this conclusion varies as a function of the latitude band. Robust and significant improvements are found only at the tropics and in the SH polar region, mixed results are obtained elsewhere. This confirms previous findings obtained with L2 assimilation (Emili et al., 2014) and adds the conclusion that a better prior does not necessarily solve all issues related to the assimilation of TIR measurements at mid and high latitudes. Nevertheless, the assimilation of radiances provides in general slightly better results at all latitudes and permits to extract more variability from IASI spectra especially at tropical latitudes.

5 Conclusions

In this study we addressed the following question: [..[118]]what are differences between the direct assimilation of IASI radiances (Level 1) and the assimilation of Level 2 products for O_3 analyses and reanalyses. We used an experimental setup where differences between the L2 retrieval and the [..[119]]L1 assimilation have been minimized as much as possible, for example by using the same RTM (RTTOV) and control vector (O_3 and SST) in both approaches. This allowed to delve into the impact of the O_3 prior and its error covariance on the quality of the [..[120]]analysis.

[114] removed: This suggests that the small differences found previously between L1a and L2a in the NH (Fig. 8) were mostly due to the impact of the stratospheric O_3 on the radiative transfer computations.

[115] removed: In the NH,

[116] removed: Below

[117] removed: . Significant differences persist in the tropical troposphere and in the 30°S-60°S band, where MLS+L1a shows improved RMSEs with respect to MLS+L2a. In particular , only the assimilation of radiances allows to partially mitigate the RMSE degradation due to MLS in the SH (30°S-60°S)

[118] removed: which

[119] removed: assimilation algorithm

[120] removed: reanalysis

We performed twins assimilation experiments with the MOCAGE CTM and the SOFRID O_3 retrievals, using the same IASI ground-pixels for both L1 and L2 assimilation, named L1a and L2a respectively. We compared the obtained [..[121]]analyses against each other and against ozonesondes and MLS profiles for the month of July 2010.

[..[122]]The results suggest that the accuracy of the O_3 prior information used in the L2 retrievals can influence [..[123]]the analysis, even when the averaging kernels are employed within the assimilation. When the O_3 prior is [..[124]]strongly biased (at low latitudes in this study), L1a and L2a differ significantly (> 10%) and the analysis shows a better variability when assimilating directly L1 radiances instead of L2 profiles. L1a and L2a are [..[125]]otherwise very similar at mid and high latitudes, where the SOFRID [..[126]]prior is closer to the true O_3 profile.

We conclude that particular care should be taken before assimilating satellite retrievals [..[127]]with prior information that can, in some circumstances[..[128]], differ significantly from the local ozone profile. Computing retrievals using an a-priori issued from a model could be relevant to improve current IASI O_3 L2 products, and might reduce the differences between L2a and L1a observed in our study. Preliminary results with SOFRID based on a modeled a-priori show also significant differences with the original product (B. Barret, personal discussion), with patterns similar to those presented in this study. However, when the final purpose is data assimilation, the L1 [..[129]]approach is more practical and statistically consistent, especially in case the observations need to be assimilated within the same forecast model that was used to compute L2 retrievals.

[..[130]]A positive impact has been found when assimilating simultaneously MLS profiles and IASI (either L1 or L2), which corrected stratospheric biases due to IASI assimilation alone. [..[131]]Differences between L1 [..[132]]and L2 assimilation are globally reduced by MLS in the stratosphere but remain significant (> 10%) in the tropical troposphere. Also, MLS assimilation strongly improves the model's accuracy down to 200 hPa and a clear added value of IASI assimilation (L1 or L2) can only be observed in the tropical troposphere. These results remind that the information brought by limb sounders like MLS in the DA system remain essential to improve upper stratosphere O_3. Interesting perspectives for future work are to: i) verify whether the assimilation of O_3 retrievals from UV spectrometers like GOME-2 or TROPOMI shows also

[121] removed: reanalyses between

[122] removed: Main findings

[123] removed: significantly the assimilation results

[124] removed: biased and the sensitivity of the retrieval is small (e.g. in the SH troposphere in winter)increased errors with respect to the control simulation are found assimilating L2 profiles (with the respective kernels). When the sensitivity is larger, but the retrieval's prior is still biased (in the tropical troposphere) , the reanalysis

[125] removed: instead very similar in the NH

[126] removed: O_3 prior is the closest to the truth

[127] removed: that

[128] removed: can have a low sensitivity to the true profile. A thorough analysis of the retrieval's DOF and averaging kernels represents the first step in this direction, but the dependence of these diagnostics to the prior itself can make this analysis troublesome. Assimilating directly

[129] removed: radiances represents a viable alternative to this. We could imagine extending this analysis to other chemical species (e. g. CO)and spectral regions (e.g. UV) that show a similar behavior to O_3 in the TIR spectrum in terms of information content of the measurements

[130] removed: Finally, a positive synergy

[131] removed: The addition of MLS was found to influence the results of IASI assimilation also in the free troposphere (500 hPa) , with L1 assimilation providing in general better results than L2 in the tropics and in the SH. This suggests that using

[132] removed: data might also be beneficial in a context of assimilating multiple instruments with different vertical sensitivities at the same time

issues related to the a-priori dependence, ii) examine if UV assimilation could replace MLS when assimilated jointly with IASI and provide similar performances in the stratosphere. This will be important to ensure the capacity to carry accurate O_3 reanalyses when the MLS instrument will be phased out.

We reckon that L1 assimilation requires modeling the full atmosphere, which may be not available to some models, those for example conceived exclusively for tropospheric applications. Moreover, Level 2 products can be aggregated vertically to correct selectively some model layers and averaged spatially to fit models with coarser resolution than the satellite ground-pixel size. This cannot be easily done with radiances and should be addressed in future research.

In this study the observations, their error covariance and the RTM auxiliary inputs were kept almost identical between L1 and L2 assimilation on purpose. Further research is needed to address issues that are common to L1 and L2 assimilation, e.g. increased errors close to the tropopause in the NH or in the tropical stratosphere. Improvements are expected for example by increasing the vertical resolution of the model, including modeled aerosols within the RT or using more realistic observation error covariances. Including more modeled variables among the RTM inputs is in particular of interest in the context of the evolution towards ESMs, where hyper-spectral sounders like IASI can provide very valuable constraint for multi-variate reanalyses (atmosphere plus surface). Including inter-channel and ground-pixel correlations in the observations error covariance matrix seems necessary to correctly weight IASI very dense observations within higher resolution models than the one used in this study. All these aspects deserve further research.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We acknowledge EUMETSAT for providing IASI L1C data, WOUDC for providing ozonesondes data and the NASA Jet Propulsion Laboratory for the availability of Aura MLS Level 2 O_3. We also thanks the MOCAGE team at Météo-France for providing the chemical transport model, the RTTOV team for the radiative transfer model, Andrea Piacentini and Gabriel Jonville for the help on technical developments of the assimilation code. This work has been possible thanks to the financial support from the Région Midi-Pyrénées, who sponsored the preliminary work of Hélène Peiro on the subject, and CNES (Centre National d'Études Spatiales), through the TOSCA program.

References

Andersson, E., Pailleux, J., Thépaut, J.-N., Eyre, J. R., McNally, A. P., Kelly, G. A., and Courtier, P.: Use of cloud-cleared radiances in three/four-dimensional variational data assimilation, Quarterly Journal of the Royal Meteorological Society, 120, 627–653, https://doi.org/10.1002/qj.49712051707, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49712051707, 1994.

Barré, J., Peuch, V.-H., Lahoz, W. a., Attié, J.-L., Josse, B., Piacentini, A., Eremenko, M., Dufour, G., Nedelec, P., von Clarmann, T., and El Amraoui, L.: Combined data assimilation of ozone tropospheric columns and stratospheric profiles in a high-resolution CTM, Quarterly Journal of the Royal Meteorological Society, https://doi.org/10.1002/qj.2176, http://doi.wiley.com/10.1002/qj.2176, 2013.

Barret, B., Le Flochmoen, E., Sauvage, B., Pavelin, E., Matricardi, M., and Cammas, J. P.: The detection of post-monsoon tropospheric ozone variability over south Asia using IASI data, Atmospheric Chemistry and Physics, 11, 9533–9548, https://doi.org/10.5194/acp-11-9533-2011, http://www.atmos-chem-phys.net/11/9533/2011/, 2011.

Bocquet, M., Elbern, H., Eskes, H., Hirtl, M., Žabkar, R., Carmichael, G. R., Flemming, J., Inness, A., Pagowski, M., Pérez Camaño, J. L., Saide, P. E., San Jose, R., Sofiev, M., Vira, J., Baklanov, A., Carnevale, C., Grell, G., and Seigneur, C.: Data assimilation in atmospheric chemistry models: current status and future prospects for coupled chemistry meteorology models, Atmospheric Chemistry and Physics, 15, 5325–5358, https://doi.org/10.5194/acp-15-5325-2015, http://www.atmos-chem-phys.net/15/5325/2015/, 2015.

Borbas, E. E. and Ruston, B. C.: The RTTOV UWiremis IR land surface emissivity module, Tech. rep., NP-SAF report, http://nwpsaf.eu/vs{_}reports/nwpsaf-mo-vs-042.pdf, 2010.

Boynard, A., Hurtmans, D., Koukouli, M. E., Goutail, F., Bureau, J., Safieddine, S., Lerot, C., Hadji-Lazaro, J., Wespes, C., Pommereau, J.-P., Pazmino, A., Zyrichidou, I., Balis, D., Barbe, A., Mikhailenko, S. N., Loyola, D., Valks, P., Van Roozendael, M., Coheur, P.-F., and Clerbaux, C.: Seven years of IASI ozone retrievals from FORLI: validation with independent total column and vertical profile measurements, Atmospheric Measurement Techniques, 9, 4327–4353, https://doi.org/10.5194/amt-9-4327-2016, https://www.atmos-meas-tech.net/9/4327/2016/, 2016.

Brown, A., Milton, S., Cullen, M., Golding, B., Mitchell, J., and Shelly, A.: Unified Modeling and Prediction of Weather and Climate: A 25-Year Journey, Bulletin of the American Meteorological Society, 93, 1865–1877, https://doi.org/10.1175/BAMS-D-12-00018.1, https://doi.org/10.1175/BAMS-D-12-00018.1, 2012.

Clerbaux, C., Boynard, A., Clarisse, L., George, M., Hadji-Lazaro, J., Herbin, H., Hurtmans, D., Pommier, M., Razavi, A., Turquety, S., Wespes, C., and Coheur, P.-F.: Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder, Atmospheric Chemistry and Physics, 9, 6041–6054, https://doi.org/10.5194/acp-9-6041-2009, http://www.atmos-chem-phys.net/9/6041/2009/, 2009.

Coman, A., Foret, G., Beekmann, M., Eremenko, M., Dufour, G., Gaubert, B., Ung, A., Schmechtig, C., Flaud, J.-M., and Bergametti, G.: Assimilation of IASI partial tropospheric columns with an Ensemble Kalman Filter over Europe, Atmospheric Chemistry and Physics, 12, 2513–2532, https://doi.org/10.5194/acp-12-2513-2012, 2012.

De Wachter, E., Barret, B., Le Flochmoën, E., Pavelin, E., Matricardi, M., Clerbaux, C., Hadji-Lazaro, J., George, M., Hurtmans, D., Coheur, P.-F., Nedelec, P., and Cammas, J. P.: Retrieval of MetOp-A/IASI CO profiles and validation with MOZAIC data, Atmospheric Measurement Techniques, 5, 2843–2857, https://doi.org/10.5194/amt-5-2843-2012, https://www.atmos-meas-tech.net/5/2843/2012/, 2012.

Dee, D. P. and Uppala, S.: Variational bias correction of satellite radiance data in the ERA-Interim reanalysis, Quarterly Journal of the Royal Meteorological Society, 1841, 1830–1841, https://doi.org/10.1002/qj, 2009.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haim-

berger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, https://doi.org/10.1002/qj.828, http://doi.wiley.com/10.1002/qj.828, 2011.

Desroziers, G., Berre, L., Chapnik, B., and Poli, P.: Diagnosis of observation, background and analysis-error statistics in observation space, Quarterly Journal of the Royal Meteorological Society, 131, 3385–3396, https://doi.org/10.1256/qj.05.108, http://doi.wiley.com/10.1256/qj.05.108, 2005.

Dufour, G., Eremenko, M., Griesfeller, A., Barret, B., Leflochmoën, E., Clerbaux, C., Hadji-Lazaro, J., Coheur, P. F., and Hurtmans, D.: Validation of three different scientific ozone products retrieved from IASI spectra using ozonesondes, Atmospheric Measurement Techniques, 5, 611–630, https://doi.org/10.5194/amt-5-611-2012, http://www.atmos-meas-tech.net/5/611/2012/, 2012.

Emili, E., Barret, B., Massart, S., Le Flochmoen, E., Piacentini, a., El Amraoui, L., Pannekoucke, O., and Cariolle, D.: Combined assimilation of IASI and MLS observations to constrain tropospheric and stratospheric ozone in a global chemical transport model, Atmospheric Chemistry and Physics, 14, 177–198, https://doi.org/10.5194/acp-14-177-2014, http://www.atmos-chem-phys.net/14/177/2014/, 2014.

Eskes, H. J. and Boersma, K. F.: Averaging kernels for DOAS total-column satellite retrievals, Atmospheric Chemistry and Physics, 3, 1285–1291, https://doi.org/10.5194/acp-3-1285-2003, https://www.atmos-chem-phys.net/3/1285/2003/, 2003.

Eyre, J. R., Kelly, G. A., McNally, A. P., Andersson, E., and Persson, A.: Assimilation of TOVS radiance information through one-dimensional variational analysis, Quarterly Journal of the Royal Meteorological Society, 119, 1427–1463, https://doi.org/10.1002/qj.49711951411, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49711951411, 1993.

Flemming, J., Huijnen, V., Arteta, J., Bechtold, P., Beljaars, A., Blechschmidt, A.-M., Diamantakis, M., Engelen, R. J., Gaudel, A., Inness, A., Jones, L., Josse, B., Katragkou, E., Marecal, V., Peuch, V.-H., Richter, A., Schultz, M. G., Stein, O., and Tsikerdekis, A.: Tropospheric chemistry in the Integrated Forecasting System of ECMWF, Geoscientific Model Development, 8, 975–1003, https://doi.org/10.5194/gmd-8-975-2015, http://www.geosci-model-dev.net/8/975/2015/gmd-8-975-2015.html, 2015.

Flemming, J., Benedetti, A., Inness, A., Engelen J, R., Jones, L., Huijnen, V., Remy, S., Parrington, M., Suttie, M., Bozzo, A., Peuch, V. H., Akritidis, D., and Katragkou, E.: The CAMS interim Reanalysis of Carbon Monoxide, Ozone and Aerosol for 2003-2015, Atmospheric Chemistry and Physics, 17, 1945–1983, https://doi.org/10.5194/acp-17-1945-2017, 2017.

Froidevaux, L., Jiang, Y. B., Lambert, A., Livesey, N. J., Read, W. G., Waters, J. W., Browell, E. V., Hair, J. W., Avery, M. A., McGee, T. J., Twigg, L. W., Sumnicht, G. K., Jucks, K. W., Margitan, J. J., Sen, B., Stachnik, R. A., Toon, G. C., Bernath, P. F., Boone, C. D., Walker, K. A., Filipiak, M. J., Harwood, R. S., Fuller, R. A., Manney, G. L., Schwartz, M. J., Daffer, W. H., Drouin, B. J., Cofield, R. E., Cuddy, D. T., Jarnot, R. F., Knosp, B. W., Perun, V. S., Snyder, W. V., Stek, P. C., Thurstans, R. P., and Wagner, P. A.: Validation of Aura Microwave Limb Sounder stratospheric ozone measurements, Journal of Geophysical Research, 113, D15S20, https://doi.org/10.1029/2007JD008771, http://www.agu.org/pubs/crossref/2008/2007JD008771.shtml, 2008.

Froidevaux, L., Anderson, J., Wang, H. J., Fuller, R. A., Schwartz, M. J., Santee, M. L., Livesey, N. J., Pumphrey, H. C., Bernath, P. F., Russell, J. M., and McCormick, M. P.: Global OZone Chemistry and Related trace gas Data records for the Stratosphere (GOZCARDS): Methodology and sample results with a focus on HCl, H2O, and O3, Atmospheric Chemistry and Physics, 15, 10 471–10 507, https://doi.org/10.5194/acp-15-10471-2015, 2015.

Gaudel, A., Cooper, O. R., Ancellet, G., Barret, B., Boynard, A., Burrows, J. P., Clerbaux, C., Coheur, P. F., Cuesta, J., Cuevas, E., Doniki, S., Dufour, G., Ebojie, F., Foret, G., Garcia, O., Granados Muños, M. J., Hannigan, J. W., Hase, F., Huang, G., Hassler, B., Hurtmans, D., Jaffe, D., Jones, N., Kalabokas, P., Kerridge, B., Kulawik, S. S., Latter, B., Leblanc, T., Le Flochmoën, E., Lin, W., Liu, J., Liu, X.,

Mahieu, E., McClure-Begley, A., Neu, J. L., Osman, M., Palm, M., Petetin, H., Petropavlovskikh, I., Querel, R., Rahpoe, N., Rozanov, A., Schultz, M. G., Schwab, J., Siddans, R., Smale, D., Steinbacher, M., Tanimoto, H., Tarasick, D. W., Thouret, V., Thompson, A. M., Trickl, T., Weatherhead, E., Wespes, C., Worden, H. M., Vigouroux, C., Xu, X., Zeng, G., and Ziemke, J.: Tropospheric Ozone Assessment Report: Present-day distribution and trends of tropospheric ozone relevant to climate and global atmospheric chemistry model evaluation, Elem Sci Anth, 6, 39, https://doi.org/10.1525/elementa.291, https://www.elementascience.org/article/10.1525/elementa.291/, 2018.

Geer, A. J., Lahoz, W. A., Bekki, S., Bormann, N., Errera, Q., Eskes, H. J., Fonteyn, D., Jackson, D. R., Juckes, M. N., Massart, S., Peuch, V.-H., Rharmili, S., and Segers, A.: The ASSET intercomparison of ozone analyses: method and first results, Atmospheric Chemistry and Physics, 6, 5445–5474, https://doi.org/10.5194/acp-6-5445-2006, http://www.atmos-chem-phys.net/6/5445/2006/, 2006.

Han, W. and McNally, A. P.: The 4D-Var assimilation of ozone-sensitive infrared radiances measured by IASI, Quarterly Journal of the Royal Meteorological Society, 136, 2025–2037, https://doi.org/10.1002/qj.708, http://doi.wiley.com/10.1002/qj.708, 2010.

Hilton, F., Armante, R., August, T., Barnet, C., Bouchard, A., Camy-Peyret, C., Capelle, V., Clarisse, L., Clerbaux, C., Coheur, P.-F., Collard, A., Crevoisier, C., Dufour, G., Edwards, D., Faijan, F., Fourrié, N., Gambacorta, A., Goldberg, M., Guidard, V., Hurtmans, D., Illingworth, S., Jacquinet-Husson, N., Kerzenmacher, T., Klaes, D., Lavanant, L., Masiello, G., Matricardi, M., McNally, A., Newman, S., Pavelin, E., Payan, S., Péquignot, E., Peyridieu, S., Phulpin, T., Remedios, J., Schlüssel, P., Serio, C., Strow, L., Stubenrauch, C., Taylor, J., Tobin, D., Wolf, W., and Zhou, D.: Hyperspectral Earth Observation from IASI: Five Years of Accomplishments, Bulletin of the American Meteorological Society, 93, 347–370, https://doi.org/10.1175/BAMS-D-11-00027.1, https://doi.org/10.1175/BAMS-D-11-00027.1, 2012.

Hubert, D., Lambert, J.-C., Verhoelst, T., Granville, J., Keppens, A., Baray, J.-L., Bourassa, A. E., Cortesi, U., Degenstein, D. A., Froidevaux, L., Godin-Beekmann, S., Hoppel, K. W., Johnson, B. J., Kyrölä, E., Leblanc, T., Lichtenberg, G., Marchand, M., McElroy, C. T., Murtagh, D., Nakane, H., Portafaix, T., Querel, R., Russell III, J. M., Salvador, J., Smit, H. G. J., Stebel, K., Steinbrecht, W., Strawbridge, K. B., Stübi, R., Swart, D. P. J., Taha, G., Tarasick, D. W., Thompson, A. M., Urban, J., van Gijsel, J. A. E., Van Malderen, R., von der Gathen, P., Walker, K. A., Wolfram, E., and Zawodny, J. M.: Ground-based assessment of the bias and long-term stability of 14 limb and occultation ozone profile data records, Atmospheric Measurement Techniques, 9, 2497–2534, https://doi.org/10.5194/amt-9-2497-2016, https://www.atmos-meas-tech.net/9/2497/2016/, 2016.

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., Lamarque, J.-F., Large, W. G., Lawrence, D., Lindsay, K., Lipscomb, W. H., Long, M. C., Mahowald, N., Marsh, D. R., Neale, R. B., Rasch, P., Vavrus, S., Vertenstein, M., Bader, D., Collins, W. D., Hack, J. J., Kiehl, J., and Marshall, S.: The Community Earth System Model: A Framework for Collaborative Research, Bulletin of the American Meteorological Society, 94, 1339–1360, https://doi.org/10.1175/BAMS-D-12-00121.1, https://doi.org/10.1175/BAMS-D-12-00121.1, 2013.

Inness, A., Blechschmidt, A.-M., Bouarar, I., Chabrillat, S., Crepulja, M., Engelen, R. J., Eskes, H., Flemming, J., Gaudel, A., Hendrick, F., Huijnen, V., Jones, L., Kapsomenakis, J., Katragkou, E., Keppens, A., Langerock, B., de Mazière, M., Melas, D., Parrington, M., Peuch, V. H., Razinger, M., Richter, A., Schultz, M. G., Suttie, M., Thouret, V., Vrekoussis, M., Wagner, A., and Zerefos, C.: Data assimilation of satellite-retrieved ozone, carbon monoxide and nitrogen dioxide with ECMWF's Composition-IFS, Atmospheric Chemistry and Physics, 15, 5275–5303, https://doi.org/10.5194/acp-15-5275-2015, http://www.atmos-chem-phys.net/15/5275/2015/, 2015.

Jaumouillé, E., Massart, S., Piacentini, A., Cariolle, D., and Peuch, V. H.: Impact of a time-dependent background error covariance matrix on air quality analysis, Geoscientific Model Development, 5, 1075–1090, https://doi.org/10.5194/gmd-5-1075-2012, http://www.geosci-model-dev.net/5/1075/2012/, 2012.

Josse, B., Simon, P., and Peuch, V.: Radon global simulations with the multiscale chemistry and transport model MOCAGE, Tellus B, 56, https://doi.org/10.1111/j.1600-0889.2004.00112.x, http://www.tellusb.net/index.php/tellusb/article/view/16448, 2004.

Lahoz, W. A., Geer, A. J., Bekki, S., Bormann, N., Ceccherini, S., Elbern, H., Errera, Q., Eskes, H. J., Fonteyn, D., Jackson, D. R., Khattatov, B., Marchand, M., Massart, S., Peuch, V.-H., Rharmili, S., Ridolfi, M., Segers, A., Talagrand, O., Thornton, H. E., Vik, A. F., and von Clarmann, T.: The Assimilation of Envisat data (ASSET) project, Atmospheric Chemistry and Physics, 7, 1773–1796, https://doi.org/10.5194/acp-7-1773-2007, http://www.atmos-chem-phys.net/7/1773/2007/, 2007.

Liu, D. C. and Nocedal, J.: On the limited memory BFGS method for large scale optimization, Mathematical Programming, 45, 503–528, https://doi.org/10.1007/BF01589116, https://doi.org/10.1007/BF01589116, 1989.

Livesey, N.: Earth Observing System (EOS) Version 4 Level data quality and description document, Tech. rep., 2018.

Marécal, V., Peuch, V.-H., Andersson, C., Andersson, S., Arteta, J., Beekmann, M., Benedictow, A., Bergström, R., Bessagnet, B., Cansado, A., Chéroux, F., Colette, A., Coman, A., Curier, R. L., van der Gon, H. A. C., Drouin, A., Elbern, H., Emili, E., Engelen, R. J., Eskes, H. J., Foret, G., Friese, E., Gauss, M., Giannaros, C., Guth, J., Joly, M., Jaumouillé, E., Josse, B., Kadygrov, N., Kaiser, J. W., Krajsek, K., Kuenen, J., Kumar, U., Liora, N., Lopez, E., Malherbe, L., Martinez, I., Melas, D., Meleux, F., Menut, L., Moinat, P., Morales, T., Parmentier, J., Piacentini, A., Plu, M., Poupkou, A., Queguiner, S., Robertson, L., Rouïl, L., Schaap, M., Segers, A., Sofiev, M., Tarasson, L., Thomas, M., Timmermans, R., Valdebenito, Á., van Velthoven, P., van Versendaal, R., Vira, J., and Ung, A.: A regional air quality forecasting system over Europe: the MACC-II daily ensemble production, Geoscientific Model Development, 8, 2777–2813, https://doi.org/10.5194/gmd-8-2777-2015, http://www.geosci-model-dev.net/8/2777/2015/, 2015.

Massart, S., Clerbaux, C., Cariolle, D., Piacentini, A., Turquety, S., and Hadji-Lazaro, J.: First steps towards the assimilation of IASI ozone data into the MOCAGE-PALM system, Atmospheric Chemistry and Physics, 9, 5073–5091, https://doi.org/10.5194/acp-9-5073-2009, http://www.atmos-chem-phys.net/9/5073/2009/, 2009.

Massart, S., Pajot, B., Piacentini, A., and Pannekoucke, O.: On the Merits of Using a 3D-FGAT Assimilation Scheme with an Outer Loop for Atmospheric Situations Governed by Transport, Monthly Weather Review, 138, 4509–4522, https://doi.org/10.1175/2010MWR3237.1, http://journals.ametsoc.org/doi/abs/10.1175/2010MWR3237.1, 2010.

Massart, S., Piacentini, A., and Pannekoucke, O.: Importance of using ensemble estimated background error covariances for the quality of atmospheric ozone analyses, Quarterly Journal of the Royal Meteorological Society, 138, 889–905, https://doi.org/10.1002/qj.971, http://doi.wiley.com/10.1002/qj.971, 2012.

McPeters, R. D., Labow, G. J., and Logan, J. A.: Ozone climatological profiles for satellite retrieval algorithms, Journal of Geophysical Research Atmospheres, 112, 1–9, https://doi.org/10.1029/2005JD006823, 2007.

Migliorini, S.: On the Equivalence between Radiance and Retrieval Assimilation, Monthly Weather Review, 140, 258–265, https://doi.org/10.1175/MWR-D-10-05047.1, http://journals.ametsoc.org/doi/abs/10.1175/MWR-D-10-05047.1, 2012.

Migliorini, S., Piccolo, C., and Rodgers, C. D.: Use of the Information Content in Satellite Measurements for an Efficient Interface to Data Assimilation, Monthly Weather Review, 136, 2633–2650, https://doi.org/10.1175/2007MWR2236.1, http://journals.ametsoc.org/doi/abs/10.1175/2007MWR2236.1, 2008.

Mirouze, I. and Weaver, A. T.: Representation of correlation functions in variational assimilation using an implicit diffusion operator, Quarterly Journal of the Royal Meteorological Society, 136, 1421–1443, https://doi.org/10.1002/qj.643, http://doi.wiley.com/10.1002/qj.643, 2010.

Miyazaki, K., Eskes, H. J., Sudo, K., Takigawa, M., Weele, M. V., Boersma, K. F., and Bilt, D.: Simultaneous assimilation of satellite NO2, O3, CO, and HNO3 data for the analysis of tropospheric chemical composition and emissions, Atmospheric Chemistry and Physics, 12, 9545–9579, https://doi.org/10.5194/acp-12-9545-2012, http://www.atmos-chem-phys.net/12/9545/2012/, 2012.

Mizzi, A. P., Arellano Jr., A. F., Edwards, D. P., Anderson, J. L., and Pfister, G. G.: Assimilating compact phase space retrievals of atmospheric composition with WRF-Chem/DART: a regional chemical transport/ensemble Kalman filter data assimilation system, Geoscientific Model Development, 9, 965–978, https://doi.org/10.5194/gmd-9-965-2016, http://www.geosci-model-dev.net/9/965/2016/, 2016.

Neal, L. S., Dalvi, M., Folberth, G., McInnes, R. N., Agnew, P., O'Connor, F. M., Savage, N. H., and Tilbee, M.: A description and evaluation of an air quality model nested within global and regional composition-climate models using MetUM, Geoscientific Model Development, 10, 3941–3962, https://doi.org/10.5194/gmd-10-3941-2017, https://www.geosci-model-dev.net/10/3941/2017/, 2017.

Paciorek, C. J. and Schervish, M.: Spatial modelling using a new class of nonstationary covariance functions, Environmetrics, 17, 483–506, https://doi.org/10.1002/env.785, https://onlinelibrary.wiley.com/doi/abs/10.1002/env.785, 2006.

Peiro, H., Emili, E., Cariolle, D., Barret, B., and Le Flochmoën, E.: Multi-year assimilation of IASI and MLS ozone retrievals: variability of tropospheric ozone over the tropics in response to ENSO, Atmospheric Chemistry and Physics, 18, 6939–6958, https://doi.org/10.5194/acp-18-6939-2018, https://www.atmos-chem-phys.net/18/6939/2018/, 2018.

Prates, C., Migliorini, S., Stewart, L., and Eyre, J.: Assimilation of transformed retrievals obtained from clear-sky IASI measurements, Quarterly Journal of the Royal Meteorological Society, 142, 1697–1712, https://doi.org/10.1002/qj.2764, 2016.

Rochon, J. Y., Garand, L., Turner, S. D., and Polavarapu, S.: Jacobian mapping between vertical coordinate systems in data assimilation, Quarterly Journal of the Royal Meteorological Society, 133, 1547–1558, https://doi.org/10.1002/qj.117, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.117, 2007.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, World Scientific Publishing Co, https://doi.org/10.1142/3171, 2000.

Saunders, R., Matricardi, M., and Brunel, P.: An improved fast radiative transfer model for assimilation of satellite radiance observations, Quarterly Journal of the Royal Meteorological Society, 125, 1407–1425, https://doi.org/10.1002/qj.1999.49712555615, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.1999.49712555615, 1999.

Saunders, R., Hocking, J., Rundle, D., Rayer, P., Matricardi, M., Geer, A., Lupu, C., Brunel, P., and Vidot, J.: Rttov-11 Science and Validation Report, pp. 1–62, https://doi.org/NWPSAF-MO-TV-032, https://nwpsaf.eu/deliverables/rtm/docs{_}rttov11/rttov11{_}svr.pdf, 2013.

Sič, B., El Amraoui, L., Piacentini, A., Marécal, V., Emili, E., Cariolle, D., Prather, M., and Attié, J.-L.: Aerosol data assimilation in the chemical transport model MOCAGE during the TRAQA/ChArMEx campaign: aerosol optical depth, Atmospheric Measurement Techniques, 9, 5535–5554, https://doi.org/10.5194/amt-9-5535-2016, http://www.atmos-meas-tech.net/9/5535/2016/, 2016.

Thompson, A. M., Witte, J. C., McPeters, R. D., Oltmans, S. J., Schmidlin, F. J., Logan, J. A., Fujiwara, M., Kirchhoff, V. W. J. H., Posny, F., Coetzee, G. J. R., Hoegger, B., Kawakami, S., Ogawa, T., Johnson, B. J., Vömel, H., and Labow, G.: Southern Hemisphere Additional Ozonesondes (SHADOZ) 1998 2000 tropical ozone climatology 1. Comparison with Total Ozone Mapping Spectrometer (TOMS) and ground-based measurements, Journal of Geophysical Research: Atmospheres, 108, 8238, https://doi.org/10.1029/2001JD000967, http://dx.doi.org/10.1029/2001JD000967, 2003.

van der A, R. J., Allaart, M. A. F., and Eskes, H. J.: Multi sensor reanalysis of total ozone, Atmospheric Chemistry and Physics, 10, 11 277–11 294, https://doi.org/10.5194/acp-10-11277-2010, http://www.atmos-chem-phys.net/10/11277/2010/, 2010.

Waters, J. W., Froidevaux, L., Harwood, R. S., Jarnot, R. F., Pickett, H. M., Read, W. G., Siegel, P. H., Cofield, R. E., Filipiak, M. J., Flower, D. A., Holden, J. R., Lau, G. K., Livesey, N. J., Manney, G. L., Pumphrey, H. C., Santee, M. L., Wu, D. L., Cuddy, D. T., Lay, R. R., Loo, M. S., Perun, V. S., Schwartz, M. J., Stek, P. C., Thurstans, R. P., Boyles, M. A., Chandra, K. M., Chavez, M. C., Chen, G.-S., Chudasama, B. V., Dodge, R., Fuller, R. A., Girard, M. A., Jiang, J. H., Jiang, Y., Knosp, B. W., LaBelle, R. C., Lam, J. C., Lee, K. A., Miller, D., Oswald, J. E., Patel, N. C., Pukala, D. M., Quintero, O., Scaff, D. M., Van Snyder, W., Tope, M. C., Wagner, P. A., and Walch, M. J.:

Table 1. Number of validation profiles for July 2010.

Latitudes	MLS	Radiosoundings
Global	100975	219
90°S-60°S	16967	19
60°S-30°S	17334	9
30°S-30°N	33046	38
30°N-60°N	16669	138
60°N-90°N	16959	15

The Earth observing system microwave limb sounder (EOS MLS) on the aura Satellite, IEEE Transactions on Geoscience and Remote Sensing, 44, 1075–1092, https://doi.org/10.1109/TGRS.2006.873771, 2006.

Weaver, A. and Courtier, P.: Correlation modelling on the sphere using a generalized diffusion equation, Quarterly Journal of the Royal Meteorological Society, 127, 1815–1846, http://onlinelibrary.wiley.com/doi/10.1002/qj.49712757518/abstract, 2001.

Weaver, C., da Silva, A., Chin, M., Ginoux, P., Dubovik, O., Flittner, D., Zia, A., Remer, L., Holben, B., and Gregg, W.: Direct Insertion of MODIS Radiances in a Global Aerosol Transport Model, Journal of the Atmospheric Sciences, 64, 808–827, https://doi.org/10.1175/JAS3838.1, 2007.

Wespes, C., Hurtmans, D., Emmons, L. K., Safieddine, S., Clerbaux, C., Edwards, D. P., and Coheur, P.-F.: Ozone variability in the troposphere and the stratosphere from the first 6 years of IASI observations (2008–2013), Atmospheric Chemistry and Physics, 16, 5721–5743, https://doi.org/10.5194/acp-16-5721-2016, https://www.atmos-chem-phys.net/16/5721/2016/, 2016.

Zhang, J., Reid, J. S., Westphal, D. L., Baker, N. L., and Hyer, E. J.: A system for operational aerosol optical depth data assimilation over global oceans, Journal of Geophysical Research, 113, 1–13, https://doi.org/10.1029/2007JD009065, http://www.agu.org/pubs/crossref/2008/2007JD009065.shtml, 2008.

Zhang, Y., Bocquet, M., Mallet, V., Seigneur, C., and Baklanov, A.: Real-time air quality forecasting, part I: History, techniques, and current status, Atmospheric Environment, 60, 632–655, https://doi.org/10.1016/j.atmosenv.2012.06.031, http://dx.doi.org/10.1016/j.atmosenv.2012.06.031, 2012.

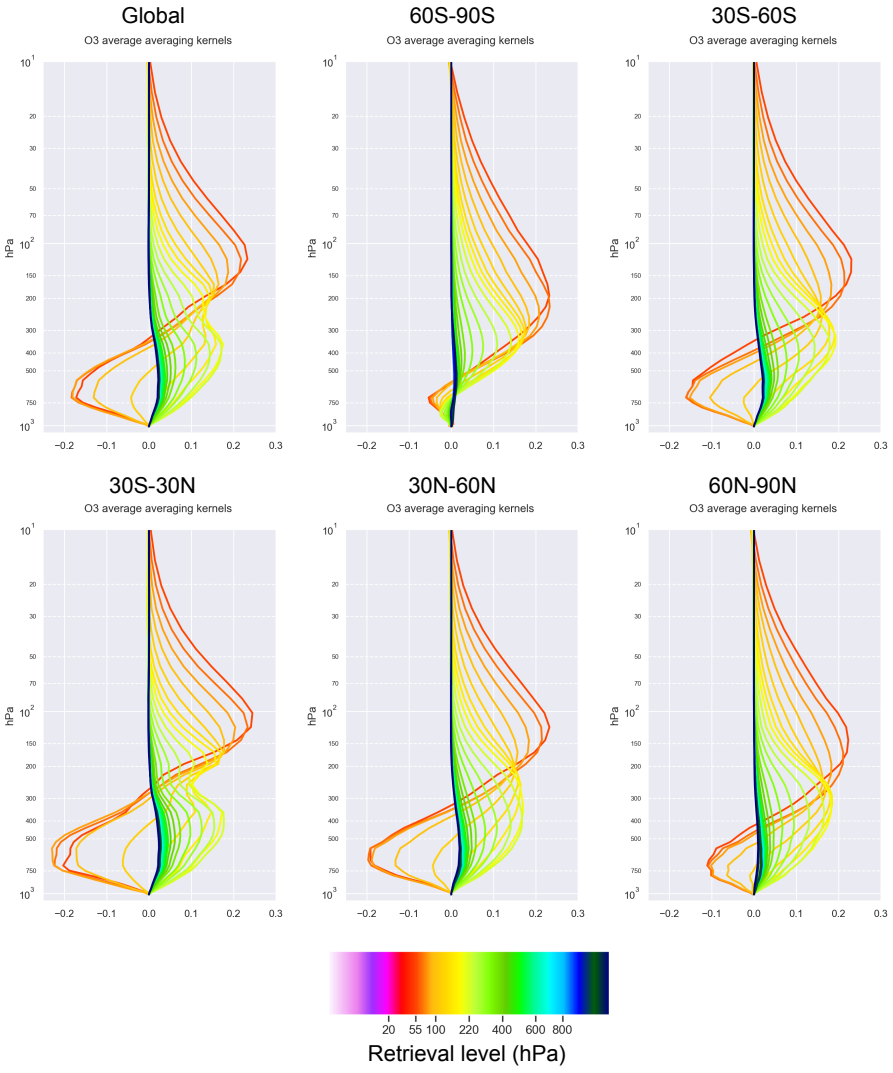

Figure 1. [..[133]]SOFRID O_3 averaging kernels for the [..[134]]month of July 2010 averaged globally (first plot) and for five latitude bands separately (90°S-60°S, 60°S-30°S, 30°S-30°N, 30°N-60°N, 60°N-90°N). Each coloured line corresponds to a retrieval's level, the corresponding pressure is indicated in the colorbar. Only SOFRID levels with a pressure > 50 hPa are displayed for better clarity.

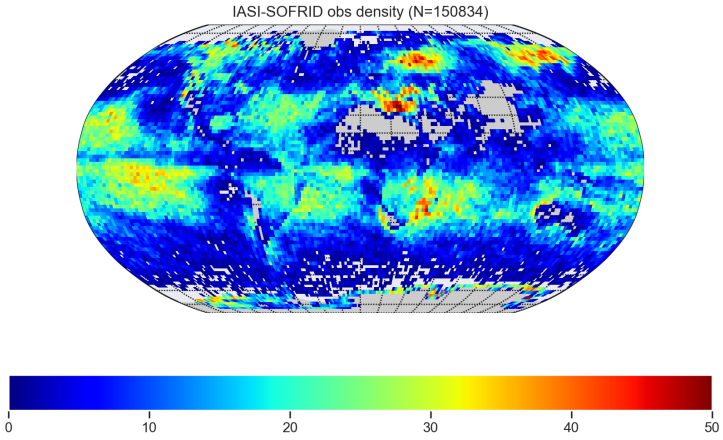

Figure 2. Total number of IASI observations per model grid box (2°x2° degrees) retained after the selection procedure described in Sec. 2.4 and further assimilated in this study for the month of July 2010. The total number of assimilated observation for the entire globe (N) is given on top of the map.

Table 2. Summary of the configuration of SOFRID L2 retrievals and MOCAGE L1 assimilation.

	L2 retrieval	L1 assimilation
Radiative transfer model	RTTOV [..[145]]V11.1	RTTOV V11.3
Algorithm	1D-Var	3D-Var
Spectral window	980 - 1100 cm^{-1}	980 - 1100 cm^{-1}
Measurements error	0.7 (mW m^{-2} sr^{-1} cm)	0.7 (mW m^{-2} sr^{-1} cm)
Control vector	O_3 (1D) + Surface Skin Temperature (SST)	O_3 (3D) + Surface Skin Temperature (SST)
Vertical grid	43 pressure levels (1013-0.1 hPa)	60 hybrid sigma-pressure levels (surface-0.1 hPa)
O_3 prior	MLS+Ozonesondes global climatology	3D-hourly model forecasts
O_3 error covariance	MLS+Ozonesondes climatological covariance	3D-hourly (standard deviation), parameterized (correlations)
SST prior (+ 2m U,V,P,T)	[..[146]]ECMWF IFS analysis	[..[147]]ECMWF IFS forecast
	6 hours timestep, 0.25°x0.25° degrees	3 hours timestep, 0.125°x0.125° degrees
SST error covariance	4°C	4°C
[..[148]]T, H_2O profiles	[..[149]]ECMWF IFS analysis	ECMWF IFS forecast
	6 hours timestep, 0.25°x0.25° degrees, 43 levels [..[150]]	[..[151]]3 hours timestep, 2°x2° degrees, 60 levels [..[152]]
IR emissivity	(Borbas and Ruston, 2010)	(Borbas and Ruston, 2010)

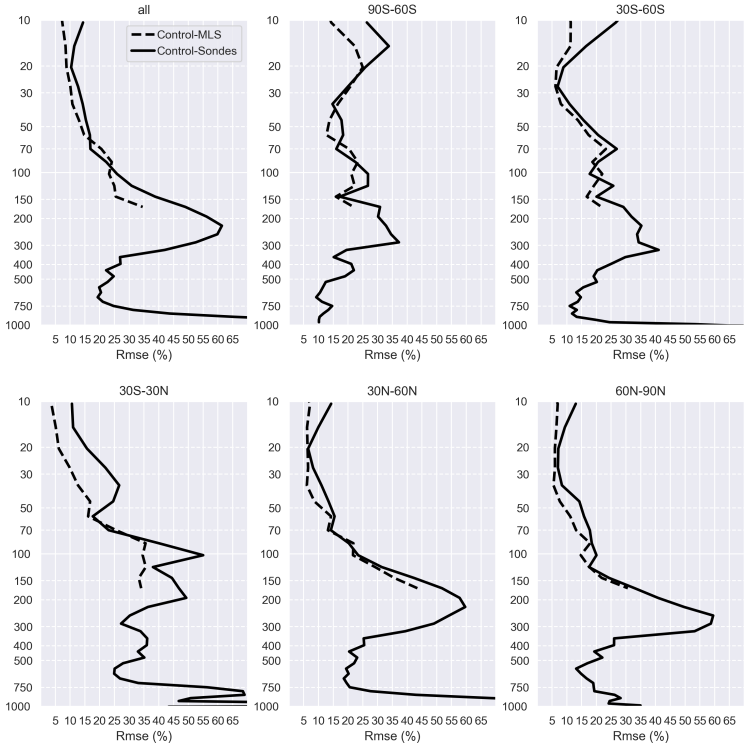

Figure 3. Relative Root Mean Square Error (RMSE) of the control simulation with respect to radiosoundings (solid lines) and MLS (dotted lines) averaged globally (first plot) and for five latitude bands separately (90°S-60°S, 60°S-30°S, 30°S-30°N, 30°N-60°N, 60°N-90°N). To compute the percentage the RMSE statistics have been divided by the correspondent average profile of the observations (radiosoundings or MLS) for each band.

Table 3. Names of experiments and assimilated data.

Experiment's name	IASI L1	IASI L2	MLS L2
Control	no	no	no
L1a	yes	no	no
L2a	no	yes	no
MLSa	no	no	yes
MLS+L1a	yes	no	yes
MLS+L2a	no	yes	yes

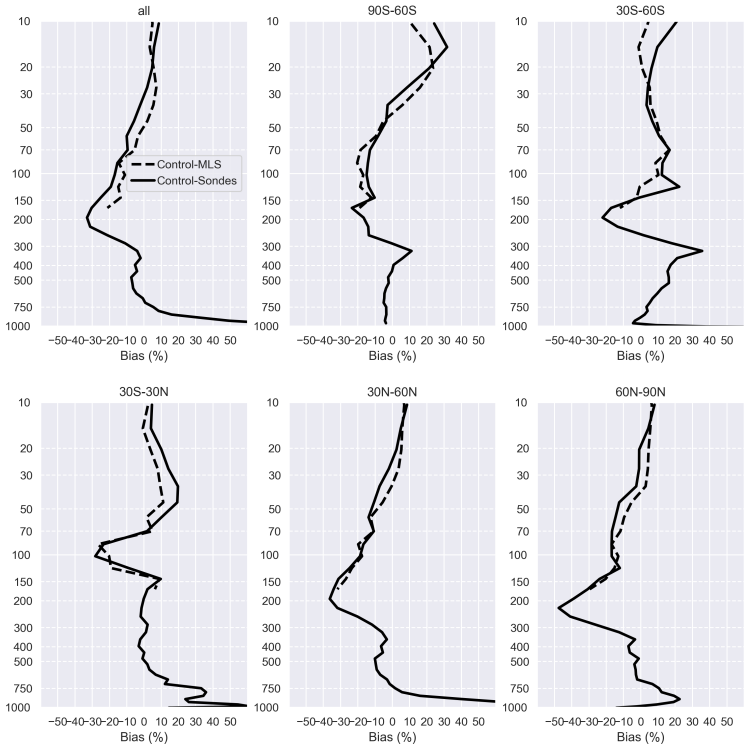

Figure 4. Relative bias of the control simulation with respect to radiosoundings (solid lines) and MLS (dotted lines). Same plots as in Fig 3.

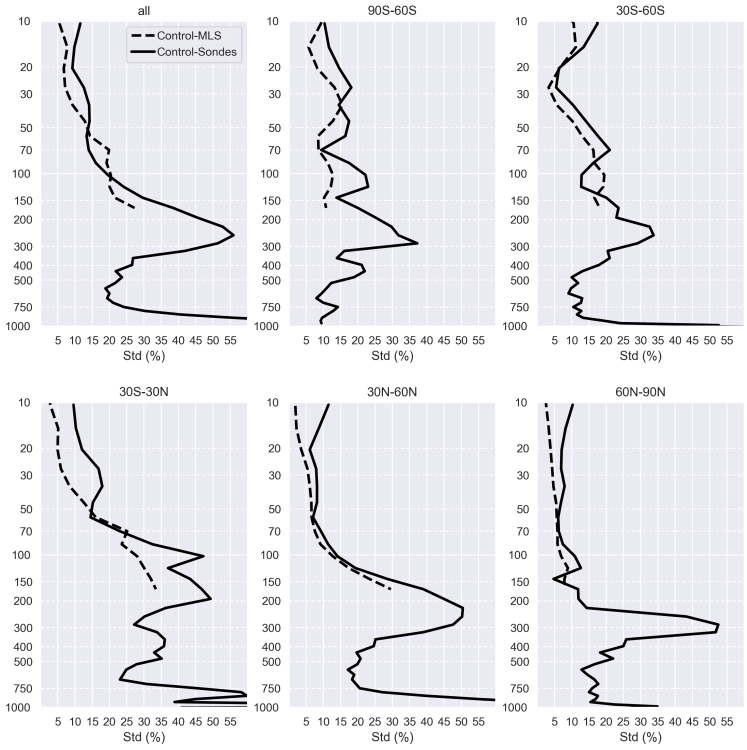

Figure 5. Relative standard deviation of the control simulation with respect to radiosoundings (solid lines) and MLS (dotted lines). Same plots as in Fig 3.

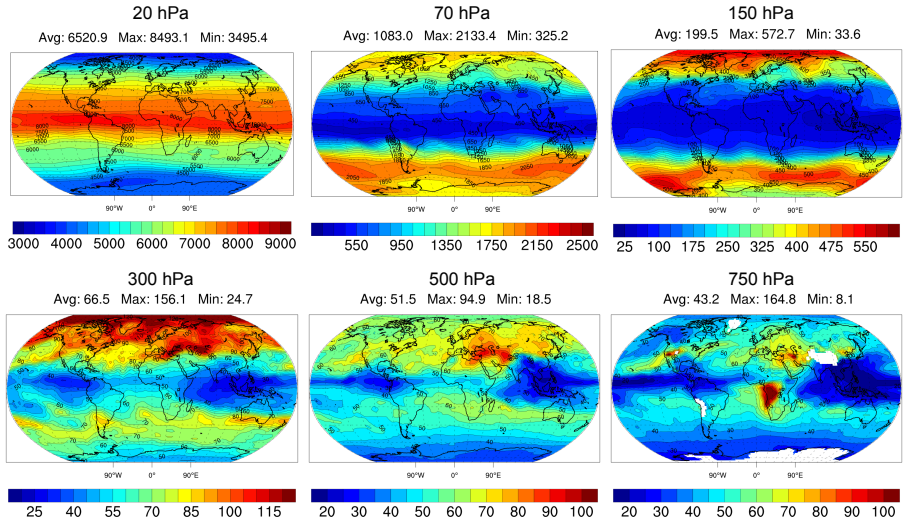

Figure 6. Average O_3 values of the control simulation in parts per billion (ppb) units for July 2010. From left to right different pressure levels are displayed covering the stratosphere (top) and the free troposphere (bottom). Average, maximum and minimum values of the displayed fields are given on top of each map.

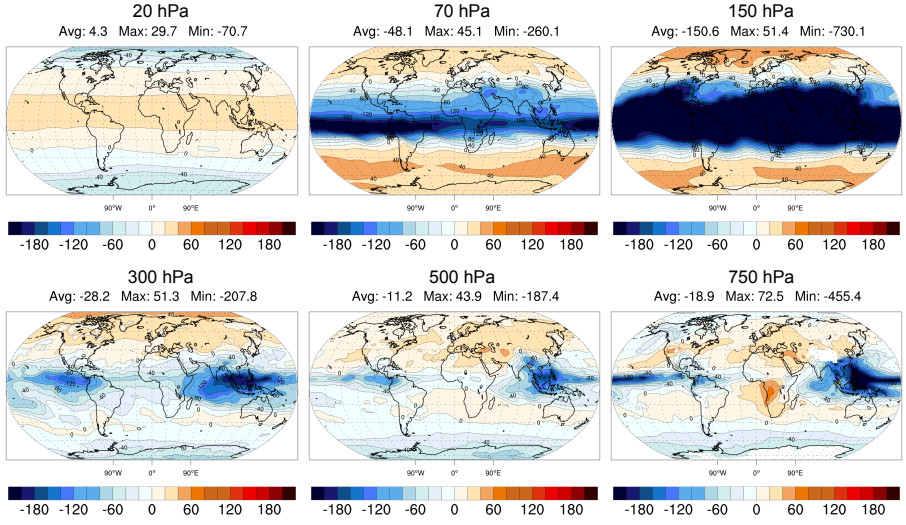

Figure 7. Relative average differences between the control simulation and the SOFRID a-priori on July 2010. Values are given as % of the control simulation (Fig. 6). Same plots as in Fig. 6.

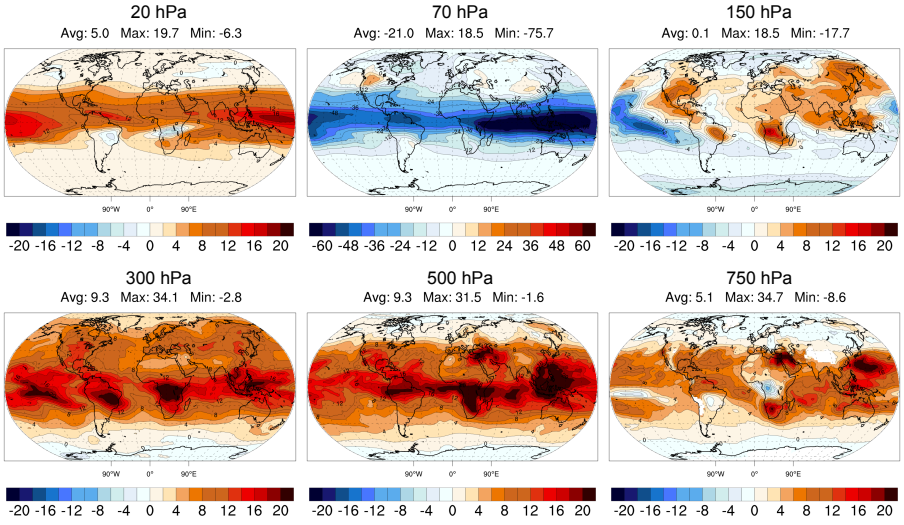

Figure 8. Relative average differences (%) between radiances and Level 2 assimilation (L1a minus L2a divided by the correspondent O_3 values of the control simulation in Fig. 6) [..[135]]on July 2010. [..[136]]Same plots as in Fig. [..[137]]6.

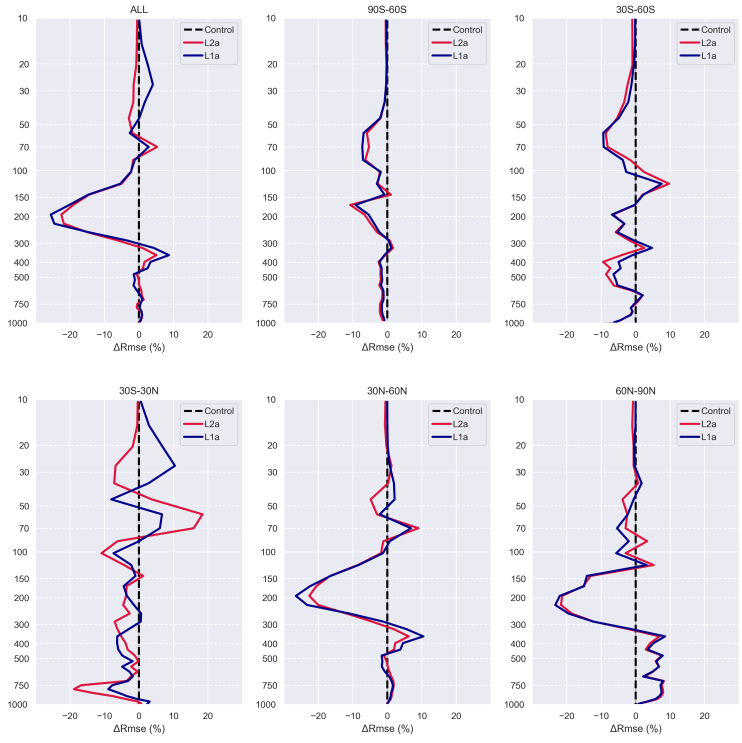

Figure 9. Relative difference of RMSE (ΔRMSE) with respect to radiosoundings for L1a (blue) and L2a (red). The difference is computed by subtracting the RMSE of L1a (L2a) against radiosoundings from the RMSE of the control simulation (Fig. [..[138]]3). Negative values mean that the assimilation improved (decreased) the RMSE of the control simulation, positive values indicate degradation (increase) of the RMSE. The statistics are computed for the same latitudes as in Fig. [..[139]]3.

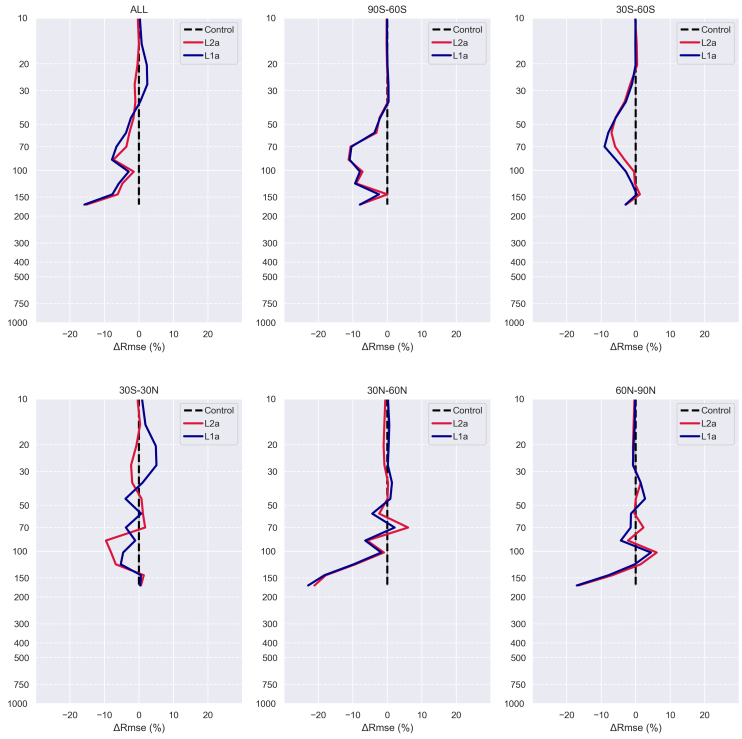

Figure 10. Relative difference of RMSE (ΔRMSE) with respect to MLS profiles for L1a (blue) and L2a (red). Same plots as in Fig. 9.

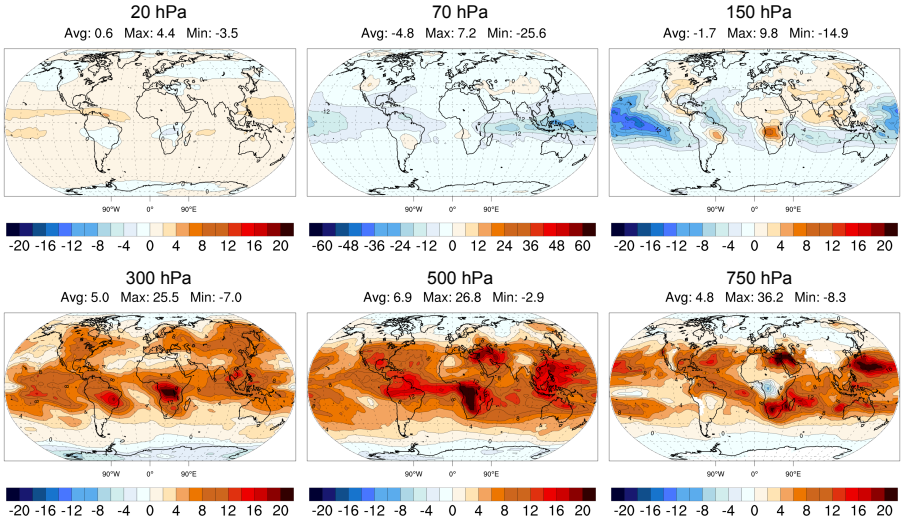

Figure 11. Relative average differences between MLS+L1a and MLS+L2a (MLS+L1a minus MLS+L2a divided by the correspondent O_3 [..[140]]values of the control [..[141]]simulation in Fig. 6) [..[142]]on July 2010. [..[143]]Same plots as in Fig. [..[144]]8.

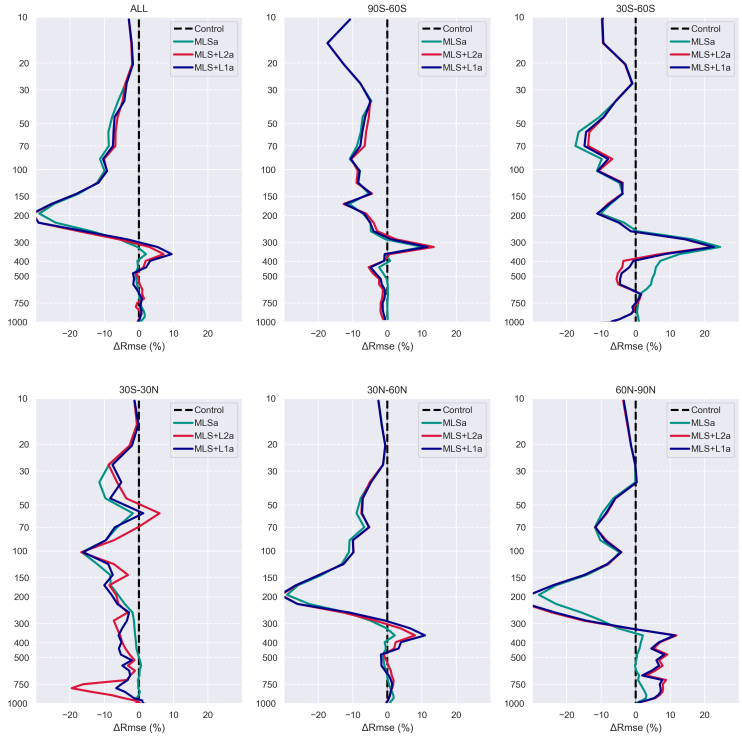

Figure 12. Relative difference of RMSE (ΔRMSE) with respect to radiosoundings for MLS-a (teal), MLS+L1a (dark blue) and MLS+L2a (red). Same plots as in Fig. 9.

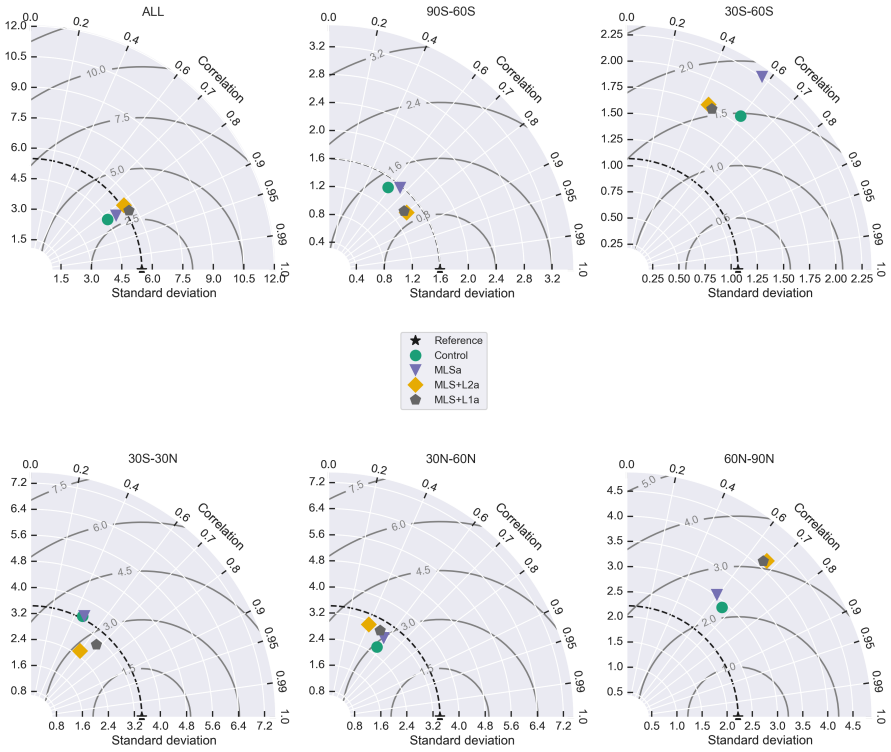

Figure 13. Taylor diagrams of modeled tropospheric ozone columns (340-750 hPa) for the Control simulation (green), MLS-a (violet), MLS+L1a (grey) and MLS+L2a (yellow) averaged globally and for five latitude bands separately. The Taylor statistics are computed against radiosoundings.