Replies to: Anonymous Referee #1. Interactive comment on "Data assimilation of satellite retrieved ozone, carbon monoxide and nitrogen dioxide with ECMWF's Composition-IFS" by A. Inness et al.

Received and published: 16 March 2015

We thank Referee 1 for their useful comments about our paper. We have tried to address all the suggestions and revised the manuscript accordingly. Our replies to their comments are given below in italics and changes to the manuscript in bold italics.

The paper presents an evaluation of the updated ECMWF's chemical composition data assimilation system C-IFS. Multiple chemical species are assimilated (O3, CO and NO2) and results are systematically evaluated in the observation space and in the model space against independents measurements. I think the study is sound and suitable for publication in the Atmospheric Chemistry and Physics journal. However there is a substantial lack of methodological and scientific explanations in some places in the article. Please see below for detailed comments about this point. Also, it is unfortunate that the authors do not emphasize the model response of one assimilated component to another. For example what would be the model response on model O3 of assimilated NO2 and/or CO. I think this might be off topic in the present paper but certainly is a point to discuss in the conclusions and discussions. That would be a first step before assessing the impact of cross correlation in the assimilation of multiple chemical species in your future work.

Concerning the general comments above, we did not emphasize the model response of one assimilated component to another because at the moment the system is set up in a way to treat all the components as independently as possible. We agree with Referee 1 that this feedback would be a very nice future study and have included a sentence in the conclusions. Some time ago, we ran studies with the old coupled MACC system to assess the impact of the assimilation of CO and NO2 on the O3 field and found some improvement in O3 that came from the NO2 assimilation. However, this has not been revisited with the C-IFS system yet. We have reworded the relevant statement in the conclusions:

A future study could look at the model response of one assimilated component to another, e.g. the response of model O_3 to the assimilation of NO_2 and CO data. This could be a first step towards investigating the interactions between the different chemical species before assessing the impact of cross correlations in the assimilation of multiple chemical species. Further plans for the development of the C-IFS data assimilation system include the recalculation of the background error statistics for all MACC control variables with the latest configuration of the model, to include emissions in the control vector so that they can be adjusted in addition to the initial conditions, especially for NO_2, and to investigate the impact of the chemical assimilation on the wind field, which has been suppressed so far.

Comments and suggestions:

P4268, L12-15: All those species are forecasted but not all are assimilated. Please be more specific.

All the listed species, apart from HCHO are assimilated in the MACC NRT system (in different model streams as the greenhouse gases run in a separate experiment). We have changed the first sentence of the second paragraph in the introduction to:

To improve the quality of the MACC forecasts the initial conditions for some of the chemical species (O3, CO, NO2, SO2, CH4, CO2, aerosols) are provided by data assimilation of atmospheric composition observations from satellites (Benedetti et al., 2009; Inness et al., 2013; Massart et al., 2014) in the MACC NRT systems.

P4270, L10-12: Please provide a reference here.

This is just experience gained during GEMS and MACC and there is no peer reviewed reference for this. We have changed the sentence to:

Experience during GEMS and MACC had shown that another disadvantage of the coupled system was that the chemical tendencies were unchanged during the one hour coupling intervals which could lead to problems at the day-night boundary for species with a short chemical lifetime.

P4271, L4: Rephrase, please. It sounds like you improve the model by using assimilation. The model analyzed fields show an improved representation of atmospheric composition.

Done. The sentence now reads:

In this study we will show that by assimilating O3, CO and NO2 observations into C-IFS the analyzed fields show an improved representation of atmospheric composition.

P4272, L24: Suggestion: Please say that the error covariance matrix between chemical species is diagonal

Done. The sentence now reads:

At present, the background errors for the chemical species are univariate, i.e., the error covariance matrix between chemical species is diagonal, in order to ….

P4272, L25: What about the feedback of meteorological variables on chemistry?

The chemistry fields are obviously affected by the meteorology (e.g. advection, temperature dependence of chemical reactions) of the IFS model. The main reason we mention only the chemistry to meteorology feedback here is that the chemistry system is still being developed and we do not want to degrade the meteorological analysis. Therefore we have limited the impact the assimilation of atmospheric composition data can have on the meteorological analysis. We have reformulated this part:

At present, the background errors for the chemical species are univariate, i.e., the error covariance matrix between chemical species or between chemical species and dynamical fields is diagonal. Although Miyazaki et al. (2012a) have shown the benefit of including correlations between the background errors of different chemical species, this is not yet included in the C-IFS system. Hence, each compound is assimilated independently from the others. Furthermore, the coupling of tracers and wind field via the adjoint of the tracer continuity equation is also disabled. This restricts the

impact of the tracer assimilation on the meteorological fields and allows us to develop the assimilation of the atmospheric composition data without the fear of degrading the meteorological analysis.

P4273, L4-9: Please explain why CO background error is estimated using and ensemble and but not for O3 and NO2?

This was purely because of practical reasons. All the background errors were re-calculated with the ensemble method, but unfortunately using the O3 and NO2 background errors really degraded the analysis and their use needed more evaluation. We therefore decided to only use the newly calculated background errors for CO and keep the old ones for O3 and NO2. We know this is not ideal and plan to re-calculate all the background errors with the latest C-IFS version shortly, now that the C-IFS data assimilation system is up and running. We have added the following sentence to Section 2.2:

It is planned to recalculate all the background error statistics with the latest version of C-IFS and test these in further assimilation experiments.

P4273, L10-12: Please be more specific here. A correlation length of 5 levels corresponds to what physically (km, hPa)? I guess, the correlation length would be larger over UTLS than toward the surface, where vertical model resolution increases. Please provide physical estimates for LT, MT and UT. What about NO2 vertical correlation length though?

Referee 1 is correct that 5 model levels are further apart in the UTLS and stratosphere than in the lower troposphere. The reason for the limitation to +/- 5 levels was that correlations between UTLS/ stratosphere and levels near the surface degraded the ozone analysis in the lower troposphere, because they led to (unwanted) changes in lower tropospheric ozone when there was a bias in the stratosphere. We thought it would be more consistent to limit the correlations to the levels near a model level. 5 levels correspond to about: 0.2 - 1 km in LT, 1-2 km in MT and about 3 km in the UT.

The vertical correlation matrix for NO2 is diagonal, i.e. there are no correlations with neighbouring levels.

We have added the following sentences to section 2.2:

The vertical correlations of the O3 and CO background errors were restricted to 5 model levels below and above a level to decouple the lower troposphere from the upper troposphere and stratosphere. This corresponds to a physical difference of about 0.2 - 1 km in the lower troposphere, 1-2 km in the mid troposphere and about 3 km in the upper troposphere. The reason for this was that the original background errors had vertical correlations between the upper troposphere/stratosphere and near-surface levels that degraded lower tropospheric ozone when there was a bias in stratospheric ozone. By limiting the vertical correlations to the neighbouring levels this degradation was avoided.

The NO2 background errors were designed to be practically zero in the stratosphere, because only tropospheric NO2 columns are assimilated in this study and the influence of the assimilation is designed to be limited to the troposphere. The vertical correlation matrix for NO2 is diagonal, i.e. there are no correlations between neighbouring levels.

P4273, L19: Typo: Profiles of profiles of : : :

Corrected.

P4273, L20: Is it possible to add errors in relative values in the plot?

Not easily. We also do not think it is necessary because the 4D- Var analysis works on absolute values and not on relative values.

P4273, L25-26: Please clarify and explain why 5%.

This has 'historic' reasons and was introduced when MIPAS data were first assimilated in the ECMWF system, because of instances when observations had unrealistically small error values which led to problems in the minimization. As a safety measure a minimum observation error of 5% was assumed. Also, there is no explicit formulation of representativeness error for the atmospheric composition observations in the MACC system, and this is accounted for by assuming a minimum error of 5%. Most atmospheric composition observations have errors larger than this, but 5% might be a bit large for some total column O3 observations nowadays. We have rephrased the sentence to:

A minimum observation error value of 5 % is used to include any observation operator error and a representativeness error that could arise because of differences in resolution of observation and the model, and that accounts for scales unresolved by the model. This minimum value will need to be reassessed as the model improves and new observational datasets become available.

P4274, L18: Please discuss why you use a 12h assimilation window. Is it short enough in time? Provide references. P4274, L19: Explain why you use two minimizations at different resolution. At least provide a reference.

We have added:

The first minimization is run with simplified physics, while the second minimization is performed with improved physics after an update of the model trajectory at high resolution (Mahfouf and Rabier, 2000). Because the parameterizations are computationally expensive the second update carries out fewer iterations of minimization than the first. 12-h assimilation windows are the standard setup of the ECMWF system at present, and it will have to be assessed in further studies if this window length is ideal for the MACC system, or if a shorter window would be better for the assimilation of shorter lived species.

P4275, L1-3: Provide the retrieval equation or a reference

We have included a reference to Inness et al. 2013 who show the equation and have relevant references. We have changed the formulation to:

Averaging kernels were used for the calculation of the model's first-guess fields in the observation operators (see Inness et al. 2013) where available, i.e., for CO data…

P4275, L3-6: I understand you want to avoid averaging the observation within a grid box in order to avoid estimating the correlation of observational errors. However by randomly selecting an observation this might lead to assimilate noisy or unrepresentative observations. Over low polluted

areas the random error or retrieval noise could be higher than the signal itself for certain instruments. Over polluted areas, because of the very heterogeneous nature of the true state and hence of the observations, assimilating randomly selected observation might cause representativeness issues. For those two reasons this method could significantly degrade the analysis compared to averaged assimilated data even with a poor estimation of the error correlation. Could the author discuss on this? Justify why they use this method over the super-observation approach? And add sentence about possible limitations?

We actually carried out some tests assimilating NRT MOPITT L3 data (which are averaged on a 1x1^{o} grid) and compared the results with the assimilation of the MOPITT L2 data for a period in 2014/2015. The results obtained were very similar in both cases and would not alter the findings of the paper. To describe the results from these experiments is beyond the scope of this paper, but they give us confidence that our method gives representative results (at least for CO) and does not degrade the analysis. We also hope to carry out a study with NO2 super-observations in the future to assess what impact the thinning method has on the NO2 assimilation. We have added some clarifying sentences to the manuscript.

A possible limitation of this thinning method is that it might lead to the assimilation of noisy or unrepresentative observations in areas of low background concentrations or to representativeness errors over polluted areas where the true state might be very heterogeneous. However, tests carried out assimilating MOPITT CO data averaged on a 1^{o}x1^{o} grid (not shown in this paper) gave very similar results to assimilating the thinned MOPITT CO data, giving us confidence that our thinning method performs well. The assimilation of averaged NO2 'super-observations' will be tested in the future.

P4275, L15-16: Why did you choose those instruments as anchors?

We have added the following to the manuscript:

The SBUV/2 data were chosen as anchor because they are a high quality reprocessed dataset. The MLS and MIPAS profile data were not bias corrected because experience in REAN had shown that the SBUV/2 data could not anchor all the layers of the higher resolved profile data and that drifts in individual layers could lead to problems in the vertical O3 distribution (Inness et al. 2013).

P4275, L22: The authors should use the word evaluation instead validation in some places. You validate a method and you evaluate results, this is not exactly the same.

We have changed validation to evaluation at several places in the manuscript.

P4277, L16 – P4278, L2: Could you scientifically explain why increase and decrease of CO occur? Why the UTLS CO mostly decreases while the extra tropical CO in the free troposphere increases?

We have added the following paragraph at the end of section 4.1.1:

The most likely reason for the underestimation of CO in CIFS-CTRL in the NH Extratropics is an underestimation of the anthropogenic emissions. This is also discussed in Flemming et al. (2015). It should be noted that low CO values are found by most of the CTMs regardless of the emission inventory used (e.g. Shindell et al., 2006; Kopacz et al., 2010; Fortems-Cheiney et al., 2011), and

that the MACCity anthropogenic emissions are in the same range as the emissions provided by the few other emission inventories available for the post-2000 period (Granier et al., 2011). A possible reason for the generally overestimation of CO in the Tropics could be too large GFAS biomass burning emissions (Flemming et al. 2015). The only exception is the strong underestimation of CO in the biomass burning maximum in Southern Africa, which points to an underestimation of the GFAS biomass burning emissions in that area (see Figure 5 below).

P4278, L14-16: How CO column data assimilation can change the CO profile. The authors need to provide explanation about transport processes here.

In the Extratropics CO columns in CIFS-CTRL are lower than the assimilated MOPITT CO data (Fig. 2), while in the Tropics CO concentrations in CIFS-CTRL are higher than the assimilated MOPITT observations (Fig. 2). A closer look at analysis increments at the beginning of the experiment shows that the assimilation leads to reduced CO values throughout the troposphere in the Tropics and to increased CO concentrations in the Extratropics, with largest absolute changes in the LT. After a while polewards transport of low CO air from the Tropics in the UT leads to lower CO values in CIFS-AN in the extratropical UT.

We have added the following statement in section 4.1.1:

It should be noted that even though TCCO data are assimilated in CIFS-AN transport processes lead to a change in the vertical CO profiles. The assimilation of TCCO data leads to increased CO columns in the Extratropics and to decreased CO columns in the Tropics (Figure 2) with corresponding positive and negative analysis increments throughout the troposphere. Poleward transport from the Tropics in the upper troposphere then leads to the lower CO concentrations in the Extratropical upper troposphere in CIFS-AN seen in Figure 4.

P4284, L8-10: Please provide a reference for this statement.

We have added a reference and changed the sentence to:

Note that the tropopause is higher in the Tropics and that O3 in UT is more influenced by the modelling of tropospheric processes, and hence differences in the chemistry schemes, than at higher latitudes where downward O3 transport from the stratosphere is larger (e.g. Škerlak et al., 2014).

P4284, L19-20: Do you have an explanation of why this is happening. Is this due to long-range transport, stratosphere-troposphere exchanges or bias in the assimilated data? The authors should add a couple sentences about this or refer to the latter explanation about NO2 in the text.

It is not entirely clear to us why the assimilation does not improve O3 in the MT over the US to the same extent as in Europe. We looked at the North American ozone sondes that go into the mean. The low bias seems to be largely coming from stations in the North West/ North Central US.

P4284, L28: are

Changed.

P4285, L3: What is MRT? I guess this is a typo and it should be MT.

*Changed to **MT**.*

P4285, L15: "so badly" : : : use more formal English please.

*Changed this to: **REAN did not perform well….***

P4286, L3: Give the definition of MNMB or a reference.

We have included a definition of MNMB and the correlation coefficient in the supplement, where the evaluation against GAW data is described in more detail. We have rephrased the sentence to:

Figure 16 shows modified normalized mean biases (MNMBs) and correlation coefficients (see supplement for definitions) from the 3 experiments…

P4288, L27-28: Please rephrase. The experiments do not give estimates of satellite values.

We have rephrased this and the sentence now reads:

However, all experiments tend to be lower than GOME-2 NO2 over Europe during the summer, but the differences might be within the error bars of the retrieval …

P4289, L8: Be more specific about "uncertainties in the chemistry".

We have added:

….and uncertainties in the chemistry, e.g. regarding photolysis rates or modelling of wet and dry deposition.

Figures: Please add a title on each subplot of each figure.

This can be done, but we do not think it is necessary as everything is described in the figure captions. If the editor is of the opinion that we should add these subtitles we will do so. Please let us know.

Figure 1: What is the altitude range of TRC and PC calculations.

We do not understand this comment in relation to Fig. 1. Does the referee perhaps mean Table 1? For OMI TRC NO2 data we use the top pressure level given by the data providers in the dataset and calculate the observation equivalent of the model accordingly. The SBUV/2 PC are a six-layer data set with the lowest layer spanning from the surface to 16 hPa.

Figure 3: Please provide the unit next to the colorbar.

Again, we leave it to the editor to decide if this is necessary, because all the information is in the caption.Please let us know.

Figure 6: Provide latitude and longitude of each location

This information is already given in Table S3 in the supplement.