Replies to reviewers' comments on "Monitoring and assimilation of S5P/TROPOMI carbon monoxide data with the global CAMS near-real time system" by Antje Inness et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2022-458-RC1, 2022

Thank you very much for the useful comments on our paper. We have tried to address all of them and give our replies in blue below. In particular, we have moved section 3 and several plots into a supplement to shorten the paper thus focussing more on the assimilation results. To reflect this, we have changed the title of the manuscript to: **Assimilation of S5P/TROPOMI carbon monoxide data with the global CAMS near-real time system**

We have added Philippe Nedelec, the IAGOS CO PI, and Luke Jones who processes IAGOS data at ECMWF as co-authors. We have changed the IAGOS plots to use calibrated data.

Comment on acp-2022-458, Anonymous Referee #1
Referee comment on "Monitoring and assimilation of S5P/TROPOMI carbon monoxide data with the global CAMS near-real time system" by Antje Inness et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2022-458-RC1, 2022

This manuscript provides a thorough accounting of the procedure, and changes in performance, of the CAMS NRT system when including TROPOMI observations of carbon monoxide (CO). The authors show that by including the TROPOMI CO observations the CAMS CO columns increase by on average 8%, which result in a much improved match compared to independent observations, as compared to a version of the CAMS system where only MOPITT and IASI observations are assimilated.

The authors furthermore assess more closely the impact of TROPOMI assimilation for fire plumes. While the total column can well be constrained, the TROPOMI data does not provide further constraints on individual plumes that are transported across continents, at altitudes above 500 hPa. This suggests (even though not written explicitly) that modeling aspects, such as accurate emission estimates and transport, remain essential to be able to capture such events.

We have added a sentence to the conclusions to highlight this: *"This suggests that modelling aspects, such as transport and accurate emission estimates, remain essential to be able to capture such events."*

The manuscript is well written, and provides a good description of the current state, and technical details, on the TROPOMI CO assimilation configuration in CAMS.
However I find it rather an accounting of the technical aspects, which would indeed be valuable as a technical report, but less so as a scientific contribution as could be expected in ACP. For that, a more critical assessment of various aspects would be expected. The authors should be more rigorous and selective in the figures that they wish to present, to make their points.
For example, the manuscript contains a long section on the monitoring of TROPOMI data in the CAMS system. While this is valuable, and well-understood from an operational

perspective, the value for the reader is not clear, and does not justify a lengthy reporting. I suggest to condense, or even remove this section, or alternatively move this into supplementary material, also because reported performance statistics over the given time period are difficult to interpret/reproduce, and not generally interesting, given the changes in the CAMS system and observation version.
We have moved the monitoring results (section 3) into a supplement and focus on the assimilation results. We hope this makes it less like a technical report.

Also various figures do not contain any novel information, and could possibly be moved into supplementary material (or removed completely), e.g. Fig. 5, while Fig. 6 and Fig 10 contain also very similar information; as well as Figure 12, 17, 19, 20, and redundancy in Figures 22-25).
To condense the paper, we have moved Section 3 to a supplement (this includes the original Fig. 5, 6, 7, 8). We have also moved the original figures 20, 23, 24 to the supplement. Furthermore, we have moved Tables 2 and 3, which contain the description of CAMS model changes and TROPOMI algorithm changes, to the supplement.

As all the monitoring plots are now in the supplement, we have changed the title of the paper to: "Assimilation of S5P/TROPOMI carbon monoxide data with the global CAMS near-real time system"

Also, discrepancies, e.g. between TROPOMI and MOPITT and IASI, as well as between TROPOMI observations and GFAS fire estimates are indicated, but the implications of these discrepancies are so far essentially disregarded. Such implications are however very important for generating long, consistent time series of TCCO in future reanalyses, and for the accuracy assessment of GFAS fire CO emission estimates, e.g. in terms of its timing, and magnitude. I think exactly such important messages can be drawn from this study, but appear lacking.

Differences between CAMS CO analysis fields and observations can have many reasons (anthropogenic emissions, biomass burning emissions, biogenic emissions, chemistry, deposition…), and cannot purely be traced back to shortcomings in GFAS. Further work is needed to investigate the reason for the negative CO bias in the CAMS model.
We have added in the text in (now) section 3.2.4:

"The differences between CAMS CO fields in CTRL and the satellite observations can have many reasons and cannot purely be traced back to shortcomings in the biomass burning emissions. Underestimation of CO is a common problem with many atmospheric chemistry models (e.g. Gaubert elt al., 2020) and not just the IFS. Other studies have related it to possible overestimation of the hydroxyl radical OH (Strode et al. (2015) as reaction with OH is the main removal of CO, underestimation of anthropogenic emissions and of non-methane volatile organic compounds (NMVOCs) from traffic emissions (Stein et al., 2014) or underestimation of secondary CO sources from the oxidation of methane and NMVOCs (Gaubert et al., 2016). Problems with the deposition fluxes could also play a role. Further work is needed to investigate the reason for the negative CO bias in the CAMS model, not only for the Arctic in this case but also in general. To isolate the impact of the GFAS emissions and to assess their importance for the CAMS CO field one could for example run

two experiments in which no satellite CO data are assimilated, one with and one without GFAS emissions, and then validate the resulting model CO field with the satellite data. This is beyond of the scope of the current paper but would be a very worthwhile future exercise."

So, In an updated version of this manuscript I challenge the authors to provide a more critical of such aspects, and reduced reporting particularly on the monitoring aspects. These elements should ideally also be reflected in an updated version of the abstract, which could well be shortened and made more to-the-point, and the conclusions section.

We have modified the abstract which now reads:
"The Tropospheric Monitoring Instrument (TROPOMI) on the Copernicus Sentinel 5 Precursor (S5P) satellite, launched in October 2017, provides a wealth of atmospheric composition data, including total columns of carbon monoxide (TCCO) at high horizontal resolution (5.5 km x 7 km). Near-real time TROPOMI TCCO data have been monitored in the global data assimilation system of the Copernicus Atmospheric Monitoring Service (CAMS) since November 2018 to assess the quality of the data. The CAMS system already routinely assimilates TCCO data from the Measurement of Pollution in the Troposphere (MOPITT) instrument and the Infrared Atmospheric Sounding Interferometer (IASI) outside the polar regions.

The assimilation of TROPOMI TCCO data in the CAMS system was tested for the period 2021-07-06 to 2021-12-31, i.e., after the TROPOMI algorithm update to version 02.02.00 in July 2021. By assimilating TROPOMI TCCO observations, the CAMS CO columns increase by on average 8%, resulting in an improved fit to independent observations (IAGOS aircraft profiles, NDACC FTIR tropospheric and total column CO data) compared to a version of the CAMS system where only TCCO from MOPITT and IASI is assimilated. The largest absolute and relative changes from the assimilation of TROPOMI CO are found in the lower and mid troposphere, i.e., that part of the atmosphere that is not already well constrained by the assimilated TIR MOPITT and IASI data. The largest impact near the surface comes from clear-sky TROPOMI data over land, and additional vertical information comes from the retrievals of measurements in cloudy conditions.

July and August 2021 saw record numbers of boreal wildfires over North America and Russia leading to large amounts of CO being released into the atmosphere. The paper assesses more closely the impact of TROPOMI CO assimilation on selected CO plumes. While the CO column can be well constrained by the assimilation of TROPOMI CO data, and the fit to individual IAGOS CO profiles in the lower and mid troposphere is considerably improved, the TROPOMI CO columns do not provide further constraints on individual plumes that are transported across continents and oceans at altitudes above 500 hPa."

More detailed comments
pp5, line 19: You describe why the IASI data is bias-corrected, and not MOPITT. (for historical reasons). Given that IASI appears more consistent with TROPOMI, wouldn't this study be an argument to change this configuration, and apply bias correction to MOPITT instead?

It does indeed and in the next CAMS model upgrade which is due in Q1/2023 we are changing the bias correction settings. IASI-C will be used as anchor and IASI-B and MOPITT will be bias corrected in future. We have added that information in the text:
"*To prepare the CAMS system for the eventual loss of the MOPITT instrument and because IASI and CAMS CO show good agreement (see Figures 7 and 8 below) in the next CAMS model upgrade (planned for Q1/2023) the VarBC settings will be changed to use IASI-C as anchor and to bias correct MOPITT and IASI-B CO.*"

pp5 line 30 and 34: 150 days were used for generating the BGE, both in the old and new configuration. Given the different lifetime, and specifics of CO for different seasons over the globe, it seems to make a difference for which period the 150 days were chosen (summer or winter).
You are right, this might be an issue. Our 150-day sample started at the beginning of February and thus covers the end of winter, spring and the beginning of summer. We agree that in future it would be better to have a sample consisting of 3 NH summer and 3 NH winter months. We have added that information in the text:
"*The 150 days covered the end of NH winter, spring and the beginning of NH summer and thus account for some seasonal differences in CO distributions. For future re-calculation of the background errors it is planned to use a sample composed of a NH summer and winter period to better account for the seasonal differences in CO, including its lifetime and changes in CO emissions, that might affect the background errors.*"

pp6, line 2: "constant in time, globally averaged (…) profile" This seems a gross simplification of the BGE information, given that the tropopause height varies strongly over the globe, and that CO emissions and lifetime features different uncertainties depending on season and region. Could you comment on this? To what extend is such a simplification applied here expected to contribute to overall errors in resulting CO analyses?
This is something we do not know at the moment. We expect this to have an impact, also on other species, e.g. O3. Work has started to use globally varying standard deviation profiles for background errors in the CAMS system and we hope this will improve the CO analysis (as well as the analysis of the other species). Unfortunately, this work is still in progress and we do not have any results yet. We have added in the text:
"*Work is on-going to allow the use of seasonally and geographically varying background standard deviation profiles for the CAMS atmospheric composition fields as this should improve the characterization of background errors that are likely to vary with season and region due to factors such as CO emissions, lifetime and tropopause height.*"

pp10-11: The authors provide a thorough, and useful accounting of the quality of the various TROPOMI CO data streams. My take on this is that the TROPOMI CO data, as used in the CAMS system, is actually biased high by a value of 6.5% (NDACC) to 9% (TCCON) (pp 11, line 9), and even larger positive biases over Antarctica. Is this correct?

Preliminary validation results against rapid delivery NDACC data indicate that the bias is reduced after the upgrade to v 02.02.00 in July 2021 (see validation reports from

https://mpc-vdaf.tropomi.eu/). This is the version we are using in our assimilation experiment. We have added a corresponding sentence in Section 2.2:
"The processor update to version 02.02.00 on in July 2021 included a change in spectroscopic parameters and preliminary results using rapid delivery NDACC data indicate that the bias is reduced to 2.9 % (Lambert et al., 2022)."
We have also added more information in Section 2.3:
Differences between TROPOMI and CAMS are reduced after the upgrade to v02.02.00 in July 2021 (Figure S4 in the supplement) which in agreement with the reduced positive TROPOMI bias reported in validation reports against NDACC observations (Lambert et al., 2022).
And in Section (now) 3.2:
"Assimilation tests with the CAMS system were carried out with the TROPOMI TCCO data for the period 2021-07-06 to 2021-12-31, i.e., after the TROPOMI algorithm upgrade to v02.02.00 (see Table S2 in the supplement) which reduced the positive TROPOMI CO bias against independent observations and led to smaller differences between TROPOMI and CAMS (see Fig. S4 in the supplement)."

If so, how
does this relate to percentual differences of monitoring experiments as reported later on in the manuscript (pp 13, line 14-15: departures of 9-11%). This also to provide further justification of the statement given on pp 17, line 1-2 ("this suggests that (…) CAMS is biased low and not that TROPOMI CO is biased high") and pp 18, line 1-2
("In ASSIM TROPOMI TCCO is used without bias correction"). It seems that the definite justification for all this is only provided in sec. 4.2.3, when independent evaluation is introduced.

We have moved this text into section 3.2.3 (former 4.2.3) after discussing Fig 10 (former Fig 14): *"This suggests that large parts of the TROPOMI TCCO analysis departures shown in Fig. S3 and S4 are the result of a low bias of the CAMS model, rather than a high bias of the TROPOMI NRT TCCO product."*

Sec. 3: Please consider to condense this section, or move this to an appendix or so. The only relevant figure is maybe Fig. 7, except that it might be better to present the data for a more consistent model configuration and time period (e.g. only for the full year of 2021). Then also the changes to the CAMS model system (incl. Table 2) could be moved there.

We have moved this section into a supplement.

pp20, line 2-3: "MOPITT departures are increased in ASSIM". I think this is a very important finding that deserves more attention than currently given. Despite the arguments given earlier about the different sensitivities of MOPITT and TROPOMI, I find it worrying that the math with MOPITT decreases, given that I expect that MOPITT TIR retrievals have been thoroughly validated, and improved, over many years (what is their reported uncertainty statistics?). Does this study, and particularly Figure 11, imply that MOPITT TIR retrievals of TCCO are biased low? What are implications for CAMS when developing a new reanalysis product? Your judgement as part of the discussion and conclusions section would be much appreciated.

This is a difficult question. We think the underlying problem is that the CAMS model (if no CO data are assimilated) has a negative CO bias which got worse with the implementation of CY46R1 and the change to CAMS_GLOB emissions in July 2019. CAMS NRT validation reports (e.g. Arola et al., 2022, https://atmosphere.copernicus.eu/eqa-reports-global-services) show this clearly, e.g. in comparison to FTIR data. Assimilating MOPITT CO (without bias correction) and IASI CO (with bias correction) in the CAMS NRT system improves the fit to the independent data, e.g. TCCON or NDACC FTIR, but some negative bias remains, as also seen in our Fig. 3. The CAMS NRT validation reports show that CAMS CO has a larger negative biases wrt to IASI over much of the Globe than MOPITT (i.e. IASI is higher than the CAMS analysis). The reason for this is that there are differences between IASI and MOPITT and we bias correct IASI, but not MOPITT so the resulting analysis is closer to MOPITT.
We do not think there is a problem with the MOPITT data, but there are differences between IASI, TROPOMI and MOPITT (assessing them could make a worthwile paper, but is beyond the scope of this paper, which purely focusses on the impact of the assimilation of TROPOMI and is not a satellite intercomparison paper). George et al. (2015) looked at the differences between MOPITT (V5T) and IASI (v20100815) (older version where MOPITT was higher than IASI) and found that using the same a-priori in the MOPITT and IASI retrievals led to better agreement between the datasets in source regions and during periods of low sensitivity, but differences remained. They attributed this to differences in time and location of the observations, differences in the vertical sensitivity of the instruments and differences in auxiliary parameters used in the retrievals (T, q, cc).

The CAMS analysis is improved if TROPOMI data are assimilated (and a larger bias is found wrt to MOPITT), but it is possible that because TROPOMI has a positive bias it compensates for the negative bias of the underlying CAMS model and gives an analysis that agrees better with observations than if MOPITT is used as the main instrument.
We have used MOPITT in the CAMS reanalysis with very good results. This reanalysis was based on an older model cycle (CY42R1) and used MACCity anthropogenic emissions, i.e. more similar to pre-46R1 NRT configuration where the model bias was smaller. We find good agreement with independent data (e.g. Inness et al., 2019, Wagner et al., 2021).
We have added a paragraph in section 3.2.2 commenting more on the MOPITT departures:
"The increased MOPITT departures in ASSIM (Fig. 7 and 8) do not imply that there are problems with the MOPITT dataset. MOPITT data have been extensively validated (e.g. Deeter et al., 2019 show that MOPITT V8 products generally have biases of less than about 5%) and are used with good results in the CAMS reanalysis where their assimilation leads to a CO analysis that agrees well with independent observations (Inness et al., 2019; Wagner et al., 2021). The underlying problem is the current negative CO bias of the CAMS model, which increased with the implementation of model cycle CY46R1 in July 2019. The CAMS CO analysis in ASSIM is improved when TROPOMI data are assimilated as TROPOMI has a small positive bias with respect to observations (see Section 2.2) which compensates for the negative bias of the underlying CAMS model and gives an analysis that agrees better with independent observations than if MOPITT is used as the main instrument. There are differences between the CO retrievals from TROPOMI, MOPITT and IASI whose investigation is beyond the scope of this paper. A study by George et al. (2015) using older versions of the MOPITT and IASI retrievals showed that that using the same a-priori in the MOPITT and IASI retrievals led to better agreement between the datasets in source regions and during periods

of low sensitivity, but that differences between the retrievals remained. They attributed these to differences in time and location of the observations, differences in the vertical sensitivity of the instruments and differences in auxiliary parameters used in the retrievals (such as temperature, humidity and cloud cover). MOPITT is likely still to be the instrument of choice for a future CAMS2 reanalysis because of the long data set going back to 2000 and the long-term stability (Deeter et al., 2019)."

pp 24, line 7: "leads to changes in surface ozone": Please remove this statement if not further substantiated. Even though O3 is undoubtedly modified, It is in my view much more likely that the changes seen are a direct effect of CO assimilation.
Sorry, this was a typo and should have read CO.

Furthermore, when presenting evaluation of CO against surface observation networks, please provide information on the quality and representativity of these observations – this is to my knowledge not straight-forward, which makes interpretation of results in Figure 18 difficult, if not impossible.

We have removed this figure from the paper and will do some more assessment on the quality and representativity of the observations. The figure is not essential for this paper as we have plenty of other independent observations.

Sec. 4.2.4 (Boreal Wildfires), particularly Figures 19-20: I believe these figures, and their discussion, even though interesting as an assessment on its own, are out of scope wrt the topic of this manuscript. When presenting these results from the GFAS system in a manuscript, I rather expect a scientific judgement, e.g. with respect to their quality, and/or impact in the CAMS system, particularly as TROPOMI CO data really enables such quantification. This connection is currently essentially missing, which therefore does not justify the figures shown here. To the least the authors should be more selective.
pp 27, line 5: "These high values are better captured in ASSIM than CTRL". What are the implications of this finding for the GFAS emission amounts reported above?

We don't quite understand some of this comment. The only figure referring to the GFAS system is Fig. 15. It is purely used to document and quantify the exceptional CO emissions due to boreal fires in 2021, allowing us to put some numbers to the biomass burning emissions and show the persistence in the values above the typical values based on the previous years in the dataset. GFAS is one of the emission input datasets used in the CAMS system (mentioned in Section 2.1.1). The GFAS system uses fire radiative power (FRP) observations from satellite-based sensors (currently MODIS but more observations are being implemented) to produce daily estimates of emissions from wildfires and biomass burning. It does not depend on the CAMS CO analysis and is the same in ASSIM and CTRL. Differences between CAMS CO analysis fields and observations can have many reasons (anthropogenic emissions, biomass burning emissions, biogenic emissions, chemistry, deposition…), and can not purely be traced back to short comings in GFAS. When we say ' These high values are better captured in ASSIM than CTRL', this is because of the TROPOMI CO assimilation and not because of any changes in GFAS.

We think a negative CO bias is still a common problem with many atmospheric chemistry models and not just IFS (e.g. Gaubert et al., 2020). The GFAS data is shown to give some context to the fires during the period of interested. There are uncertainties in the GFAS emissions (which are also valid for other biomass burning emissions datasets):

- Detection limit of the sensors for observing smaller fires, or low temperature fires (such as in peatlands). VIIRS, for example, sees more smaller fires but also more spurious signals.
- Knowledge of fuel types (i.e. vegetation types and peat) and associated emission factors – particularly in Eurasia/Siberia.

We have re-phrased the paragraph to clarify that GFAS provides input data for the CAMS system. It now reads:
"Increased flammability and wildfire risk, due to high temperatures and drought conditions, were manifested by the development of large-scale and persistent wildfires between June and August 2021 in the boreal regions of North America, particularly in the western United States and Canada, and Eurasia, particularly in the Sakha Republic of Russia. To illustrate the increased emissions from these fires, Fig. 15 shows that daily total biomass burning emissions of CO for Canada, the US and Siberia in July and August 2021 from the Global Fire Assimilation System (GFAS; Kaiser et al., 2012) were persistently above the typical daily values based on the 2003-2020 data in July and August 2021. The annual total biomass burning CO emissions for 2021 of 46 Tg (Siberia), 17 Tg (Canada) and 9 Tg (US) were considerably larger than the multi-year mean (2003-2020). GFAS uses fire radiative power observations from the two Moderate Resolution Imaging Spectrometer (MODIS) instruments on the NASA Terra and Aqua satellites to produce daily global estimates of emissions from biomass burning, with a spatial resolution of 10 km x 10km, and is one of the emission input datasets for the CAMS system (see Section 2.1). GFAS emission estimates are likely to be an underestimation of the real biomass burning emissions, and uncertainties (which are also valid for other biomass burning emission datasets) arise from (1) the detection limit of the sensors for observing smaller fires, or low temperature fires (such as in peatlands) and (2) the knowledge of fuel types (i.e. vegetation types and peat) and associated emission factors – particularly in Eurasia/Siberia."

We have moved Figure 20 into the supplement.

pp 29: "horizontal mismatch". Considering that the modeled plume is present one day too late, could it possibly also be a mismatch in time of the GFAS emissions in this configuration?
This is also a possibility. We have added to: "This might be because of a horizontal or temporal mismatch…"

pp 29, l16 "can be considered a success": I quite do not agree with this statement. Of course TROPOMI CO is very successful in correcting the total column, much better than was the case in the CTRL configuration, but from the given evidence it appears simply not able to significantly improve the placement of the plumes in the right altitude, which was the topic of exactly this analysis.

We have changed that sentence to:

While TROPOMI is very successful in correcting the total CO column it cannot completely correct deficiencies in the CAMS vertical CO profiles, suggesting that modelling aspects such as transport and accurate emission estimates remain essential to capture such events in the absence of vertically higher-resolved satellite data for use in the assimilation.

pp32, line 1 (46 Tg, 17 Tg, 9 Tg) - please consider to remove the statement on quantification of CO emissions here. The current evaluations actually suggest that these are under-estimates (but unclear by how much)

We have changed the sentence to: "*In particular, the NH summer of 2021 saw strong wildfires in North America and Russia that released record amounts of CO into the atmosphere.*"

Conclusions section: Please provide an overview of open issues, and current limitations.

We have added:
"While the CO column can be well constrained, TROPOMI CO data do not provide further constraints on individual plumes that are transported across continents or oceans at altitudes above 500 hPa. Modelling aspects such as transport and accurate emission estimates remain essential to succesfully capture such events. Furthermore, the total column TROPOMI data do not have the vertical resolution to introduce such fine-scale structures in the CAMS analysis. For this, data with higher vertical resolution would be needed.

One shortcoming of the current CAMS system is a low CO bias in the free troposphere. This is a common problem with many atmospheric chemistry models (e.g. Gaubert elt al., 2020) and not just the IFS. Further work is needed to understand the reason for this bias and to improve the model. While the negative bias increased with the change to model version CY46R1 and the change to the CAMS_GLOB emission inventory in July 2019, negative biases were seen before then and might not purely be a result of underestimated anthropogenic emissions. Factors such as the distribution of the hydroxyl radical OH, secondary CO sources from the oxidation of methane and NMVOSs, and deposition processes might also be important. CAMS is in the process of developing an inversion prototype using ECMWF's 4D-Var system (McNorton et al., 2022), which should help to address shortcomings in emission inventories, but work on modelling aspects is also needed."

Technical corrections
pp1 line 28: "already … already" please rewrite removing one 'already'.
Done

pp3 line 16: "published" change to "reported", remove "but not in a peer reviewed publication".
Done

line 18 "additional work", please specify "work" here, ore rewrite this sentence. "Additional work" does not justify a manuscript per se.
We have removed that part of the sentence, which now reads:

"The current work makes use of a much longer timeseries and more mature retrieval version of TROPOMI CO data, and documents the preparation of the global CAMS NRT system for the routine NRT assimilation of TROPOMI total columns CO (TCCO) data, by presenting results from assimilation tests with the NRT TROPOMI TCCO data for the period 6 July to 31 December 2021."

line 39: "CAMS-GLOB-ANT"
Done

pp4, line 1 "CAMS-GLOB-BIO"

done

pp6, line 6/7: please provide a more generic reference, this is not really readable. Alternatively consider to move this technical information, which is not essential to the subject of TROPOMI CO assimilation, towards an appendix or so (see also generic comments above)

We have moved this table into the supplement.

pp 7, line 8: IAGOS: Provide here the explanation of this acronym, and not in line 12.
Done

line 15 "in this paper" -> "here"
done

line 28: "station"->"stations"
done

pp 9, Table 3: Check ordering of rows: one date stamp (2019-08-06) should come after 2019-07-03?
corrected

pp 13, line 15: "medium"->"median"
done

Comment on acp-2022-458
Anonymous Referee #2
Referee comment on "Monitoring and assimilation of S5P/TROPOMI carbon monoxide data
with the global CAMS near-real time system" by Antje Inness et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2022-458-RC2, 2022

General Comments:
This manuscript presents the assimilation results of NRT CAMS by assimilating TROPOMI CO observations with the following contents: 1) assimilation result of CAMS by assimilating MOPITT and IASI CO; 2) difference between TROPOMI CO and CAMS; 3) difference between CAMS-TROPOMI with control run; 4) validation with respect to independent observations; 5) impact of boreal fires. The experiment results, shown in this manuscript, are helpful for better assimilations of tropospheric CO. I recommend the paper for publication after consideration of the points below.

In general comments, the detail of this manuscript should be improved. There are different experiment configurations with complicated names such as "CY47R3"; The titles of subsections need to be improved, for example, "4.2.1 Difference plots ASSIM minus CTRL experiments";

We have changed the title of this section to: "*Differences between assimilation and control experiments*"

the fonts of some figures are too small to see;

We hope that the quality of the figures is good enough when the original files are uploaded for the final publication. The quality of some of them was reduced when including the in the word document. We ask the editor to please let us know if any figures are not good enough for publication when we have uploaded the figures and we will improve them at that stage.

the titles of some
figures seem from original experiment records and should be re-designed;

Changed the titles this in (now) Fig. 3, 10, 11. Re-did (now) Fig. 14 at higher resolution.

the term of
"departure" is widely used in figure captions and should be clearly defined: whether they represent model minus observation or observation minus model.

We do already define what departures are in the introduction (penultimate paragraph): *"The differences between the observations and the model equivalents are called departures. We distinguish between first-guess departures (observations minus model first-guess field) and analysis departures (observations minus analysed field)."*

While the assimilation
result is worthy of publication, the style of the current version is a little different from common research articles.

We hope by moving section 3 to a supplement (see also replies to referee 1) we have improved this.

Specific Comments:

Page 4, Line 20: What is the purpose of “randomly selecting one observation in each grid cell”? Is it better to use the average of observations within each grid cell to reduce the effect of random errors in satellite observations?

Yes, it is and this is indeed what we do when creating super-observations for TROPOMI. The other method was coded a long time ago and is still applied to IASI and MOPITT data. More work to improve the super-observations software is planned for next year and will then also be applied to more instruments.

Page5, Lines 11-20: I understand there could be some “historical reasons”. However, the description here is uncommon in research articles.

We have changed the sentence to:
“MOPITT is used as the anchor because it was the first instrument assimilated in the early CAMS system, and the assimilation of IASI CO was added later.”

Page11, Line29: The direct averaging of satellite data to the model grid may not be a strict "super-observation" because the effect of observation error at each observation point is not considered, for example, Section 2.3.3 in Miyazaki et al. 2012. In addition, I don’t think the "super-observation" can reduce random errors.
Miyazaki, K., Eskes, H.
J., and Sudo, K.: Global NOx emission estimates derived from an assimilation of OMI tropospheric NO2 columns, Atmos. Chem. Phys., 12, 2263–2288, https://doi.org/10.5194/acp-12-2263-2012, 2012.

Averaging the observations in a grid box will reduce the random error. We average the error in the same way as the observation values. Our current method is a simple average over the obervations in a grid box and not a weighted average as described by Miyazaki et al. (2012) that also takes error correlations into account when calculating the super-obbed observation error. We plan to introduce a more sophisticated super-obbing scheme in the the next year, but this is not in place yet.
We have changed the text to make the difference to the Miyazaki method clearer:
“Our method to create super-observations simply calculates the averages of the observations in a grid-box and hence differs from a method described by Miyazaki et al. (2012) who also weigh the individual observations depending on the data coverage and take error correlations among the data into account. We average the data separately for different surface types (e.g., land, ocean, ice etc.) and for clear and cloudy data, and the observation errors and averaging kernels of the data are averaged in the same way as the observations. This averaging reduces the random errors in the data and also the representativeness errors due to unresolved small-scale features that are seen in TROPOMI data but not resolved in the model.”

Page 17, Lines 26-27: what is the difference between CTRL and CAMS discussed in the previous section?

CTRL is a copy of ASSIM without the assimilation of TROPOMI CO data, to be sure we have a clean comparison. Both ASSIM and CTRL used the model cycle CY47R3. 'CAMS' is the operational configuration that runs routinely every day to produce the NRT CAMS analysis and 5-day forecasts. This operational system changes with time and was upgraded to CY47R3 in October 2021, so that after that date it used the same model cycle that we use in ASSIM and CTRL, before that date CAMS used an older model cycle. We already state in Section 4.2 (3.2 in revised version): "*The model cycle used for the experiments was CY47R3, meaning that the CTRL setup corresponds to the CAMS NRT configuration that was operational from October 2021*."

As shown in Figure 11, the improvement by assimilating TROPOMI CO with respect to the original MOPITT+IASI is mainly due to the underestimated CO columns in MOPITT TIR data. However, why the newer multi-spectral TIR+NIR MOPITT data were not considered?

We acquire the MOPITT data in near-real time from LANCE, and the product available there is the TIR product. Also, we use the TIR data because it is thought to have the highest stability, which is important for a NRT operational application like CAMS. We agree that it could be interesting to explore the use of the combined TIR+NIR product for offline applications.

In addition, the a priori biases are comparable between MOPITT (0.6) and IASI (1.5), then why the a posteriori bias is much larger in MOPITT (-3.3) than IASI (1.2)?

We think the underlying problem is that the CAMS model (if no CO data are assimilated) has a negative CO bias which got worse with the implementation of CY46R1 and the change to CAMS_GLOB emissions in July 2019. CAMS NRT validation reports (e.g. Arola et al., 2022; https://atmosphere.copernicus.eu/eqa-reports-global-services) show this clearly, e.g. in comparison to FTIR data. Assimilating MOPITT CO (without bias correction) and IASI CO with bias correction in the CAMS NRT system improves the fit to the independent data, e.g. TCCON or NDACC FTIR, but some negative bias remains, as also seen in our Fig. 3. The CAMS NRT validation reports show that CAMS CO has a larger negative biases wrt to IASI over much of the Globe than MOPITT (i.e. IASI is higher than the CAMS analysis). The reason for this is that there are differences between IASI and MOPITT and we bias correct IASI, but not MOPITT so the resulting analysis is closer to MOPITT. When TROPOMI data are assimilated the analysis is drawing to those data which have a small positive bias, and this results in more negative MOPITT departures. See also our reply to Referee 1 on this.

Page 25, Figure 18: It is interesting to see the large difference between Europe/North America and China. What could be the possible explanations? A recent study found a similar small influence of MOPITT CO assimilations on surface CO concentrations in China.
In addition, what is the meaning of "Marco Polo" in the title of Figure 18c?
We have removed this figure from the paper and will do some more assessment on the quality and representativity of the observations. The figure is not essential for this paper as we have plenty of other independent observations.

Tang, Z., Chen, J., and Jiang, Z.: Discrepancy in assimilated atmospheric CO over East Asia in 2015–2020 by assimilating satellite and surface CO measurements, Atmos. Chem. Phys., 22, 7815–7826, https://doi.org/10.5194/acp-22-7815-2022, 2022.

Technical Comments:
Page 1, Line 22: the abbreviation of "NRT" should be defined.
We already define it in the introduction but had not defined it in the abstract. To avoid having too many acronyms in the abstract we have removed NRT there.

Page 5, Line24: change "2021-07-09" to "2019-07-09"
Done

Page 17, Line8: change "3 10 18" to "4.3 x10 18"
Done

Page 21, Figure 12 and Figure 13: the definitions for the abbreviations of "FgDep" and "AnDep" should be provided.
We have added this in the captions:
Fig. 12 (now Fig 8): "*Analysis departures (abbreviated as AnDep) from ASSIM are in red, for CTRL in magenta. First-guess departures (abbreviated as FgDep) from ASSIM are in blue…"*
Fig 13 (now Fig 9): "Analysis departures (abbreviated AnDep)/ values from ASSIM are in red, for CTRL in magenta. First-guess departures (abbreviated FgDep)/ values from ASSIM are in blue, for CTRL in black."

Page 22, Figure 14: I assume the left and right columns are relative biases for the assimilated and control runs, respectively. But it should be clarified in the figure caption.
We already say in the caption: "… from ASSIM (left) and CTRL (right)…". We have now changed this to "*from ASSIM (left column) and CTRL (right column*)" and hope this makes it clearer.

Comment on acp-2022-458
Anonymous Referee #3
Referee comment on "Monitoring and assimilation of S5P/TROPOMI carbon monoxide data
with the global CAMS near-real time system" by Antje Inness et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2022-458-RC3, 2022

In their paper, Inness et al descibe the assimilation of TROPOMI CO in the world-leading CAMS analysis system. The combined assimilation of TROPOMI, MOPITT and IASI CO is described in detail, and major advances are reported by using TROPOMI. The paper is well written, is presenting important results, and is complemented by validation results with independent observations. I am much in favour of publishing these results. Several questions came up when reading the paper, and I would ask the authors to answer those before the paper is published.

IAGOS data is used prominently in the paper so it seems reasonable to invite an IAGOS team member as co-author, or at least mention the programme explicitly in the acknowledgement.
We have added Philippe Nedelec as co-author and added in the acknowledgement: *MOZAIC/CARIBIC/IAGOS data were created with support from the European Commission, national agencies in Germany (BMBF), France (MESR), and the UK (NERC), and the IAGOS member institutions (http://www.iagos.org/partners). The participating airlines (Lufthansa, Air France, Austrian, China Airlines, Hawaiian Airlines, Iberia, Cathay Pacific, Air Namibia, Sabena) supported IAGOS by carrying the measurement equipment free of charge since 1994. The data are available at http://www.iagos.fr thanks to additional support from AERIS.*

We have also re-done the plots showing IAGOS data with L2 calibrated data.

The paragraph starting with "The differences between the observations and the model equivalents are called departures" (page 3) is general and rather long to my taste. The sensitivity of the departures to detect small changes is an important point to make I agree.

We have decided to keep this paragraph because Referee 2 requested a definition of departures, and this paragraph does define the term.

"Initial work about the use of early TROPOMI CO data in the CAMS system was published in an ECMWF technical memorandum (Inness et al., 2019c) but not in a peer reviewed publication."
The early paper of Borsdorff (a co-author) from 2018 already showed striking good agreement between early TROPOMI retrievals and the CAMS system. Reading this paper it is a bit surprising that "a more mature retrieval version" seems to be needed for the assimilation to be successful, as is suggested by the text of the introduction? The good initial comparisons are discussed on page 10.
"contains a lot of additional work," (page 3, line 18) If this is relevant, it should be specified what the additional work consists of.

The initial work was submitted to ACP in 2019 but was rejected with the comment that the early TROPOMI CO data were not mature enough. We then published it as an ECMWF Tech memo to at least document the work. The sentence is a justification why we now publish a paper that has similarities with the old Tech memo, something which was requested by the editor before the paper was published in the ACP discussion. We have re-phrased the sentence now
"Initial work about the use of early TROPOMI CO data in the CAMS system was reported in an ECMWF technical memorandum (Inness et al., 2019c). The current work makes use of a longer timeseries and more mature retrieval version of TROPOMI CO data, and documents the preparation of the global CAMS NRT system for the routine NRT assimilation of TROPOMI total columns CO (TCCO) data, by presenting results from assimilation tests with the NRT TROPOMI TCCO data for the period 6 July to 31 December 2021."

Could you please comment why data with LAT < 60S is blacklisted?
Because TROPOMI validation shows large differences with observations there, i.e. at Arrival Height. We thought it was better to not use those data. We have added at the beginning of (now) Section 3.2:
"…not over Antarctica where comparison with NDACC data showed larger biases in the TROPOMI validation reports (https://mpc-vdaf.tropomi.eu/),…"

Equation 2, 3: I would expect to see error terms here to reflect measurement and retrieval uncertainties, and forecast errors, especially since the "true" atmosphere is introduced here. The equations suggest that the true state can be quantified which is not the case. I would recommend that the error terms are explicitly added to these equations to avoid confusion. "d" is resulting from these errors.

We have added an error term to equations 2, 3, 4 and added in the text:
"… ε an error term for measurement errors and errors in the forward model".

"we remove the influence of the a-priori profile" (page 5, line 9) In general the a-priori does not only appear explicitly in these equations, but also implicitly through the dependence of the retrieval "y" and retrieval error on the a-priori, since a badly chosen apriori will generally lead to larger retrieval errors and larger d values.
Thanks for that comment. We have changed the relevant text to:
"… we remove the explicit influence of the a-priori profile in the calculation of the departures, but knowledge of the a-priori profile is still needed in the observation operator calculations for IASI and MOPITT TCCO retrievals. Also, the impact of the a-priori remains implicitly through the dependence of the retrieval y and the retrieval error on the a-priori, since a badly chosen a-priori will generally lead to larger retrieval errors and larger departures."

"the thickness of the 1000-300 hPa layer," is a bias predictor (page 5, line 15) What is the physical relation between the bias and this quantity?
This is used to give some latitudinal variation in the bias correction. We have added in the text: *"The thickness parameter was chosen to allow latitudinal variations of the bias correction."*

NEWBGE and OLDBGE. In many data assimilation applications the background error is scaled using chi-square statistics, which basically requires that the background errors are consistent with the computed "d" values and observation errors. Is something similar done in the CAMS system? Phrased differently: is there statistical evidence that the NEWBGE is more realistic than OLDBGE?
We have not scaled the background errors using chi-square statistics but agree it might be worth exploring in the future. The background errors were created with the NMC method using the model and we update them periodically to remain consistent with the model as it evolves. So, NEWBGE are more realistic with the current model because they are based on the actual statistics of the that model, whereas the OLDBGE were the realistic statistics of the old model. We have added in the text:
"By updating the background errors, we remain consistent with the model as it evolves."

Also for equation 4 (page 10) I would suggest to explicitly add the error terms.

Done.

"with a mean difference between the data sets of 3.2±5.5 %" (page 10, line 21) Is CAMS higher or lower?

TROPOMI is higher. We have added in the text: …" *with a mean difference between TROPOMI and CAMS of 3.2±5.5%...*"

"9.22 ± 3.45% against standard TCCON XCO data and 2.45±3.38% against TCCON unscaled XCO " (page 10, line 27) Please explain why there are two numbers and which one is reflecting the actual bias more.

A calibration is done on the TCCON XCO (standard XCO = calibrated columns, unscaled XCO = no calibration applied) which Sha et al. (2021) think is wrong (the calibration factor is corrected in the latest TCCON processing which was not available for validation at the time). To avoid confusion we have removed the bias value calculated with the standard XCO from the paper and now simply say:
"Sha et al. (2021) reported a bias of 2.45±3.38% against TCCON unscaled XCO…". We have also slightly rephrased the following paragraph to give more weight to the NDACC validation rather than TCCON in the validation reports:
"The routine quarterly TROPOMI validation reports available from https://mpc-vdaf.tropomi.eu/ (last access 22/8/2022; Lambert et al., 2022) show that the S5P L2_CO (NRT or RPRO concatenated with OFFL) carbon monoxide total column data is in good overall agreement with co-located measurements from the NDACC, TCCON and Collaborative Carbon Column Observing Network (COCCON) FTIR monitoring networks. They showed on average a positive bias of approximately +10 % (NRT, before July 2019) or +6.5 % (OFFL and NRTI after July 2019). The validation reports found no latitudinal dependence of the bias and a slight increase of the bias during local winter. Biases at most individual NDACC stations were well below 10%, but slightly larger at mountain stations and also at Eureka (Arctic) and Arrival Height (Antarctica). The biases increased with solar zenith angle by about 10% between 10˚ and 80˚. The processor update to version 02.02.00 on in July 2021 included a change in spectroscopic parameters and preliminary results using rapid delivery NDACC data indicate that the bias is reduced to 2.9 % (Lambert et al., 2022). "

From the abstract I got the (maybe wrong) impression (I realise it is not explicitly mentioned in this way) that the retrieval upgrade to version 02.02.00 was quite important for the success. From section 2.2, however, it is not so clear how much the retrieval versions 1 and 2 differ. The validation results are discussed, but the versions are not separately reported. Is this information available? Is there a clear indication of changes/improvements in version(s) 2 compared to retrievals before Summer 2021? It would be good to have some quantitative information of the differences before the comparisons with CAMS are presented.

Preliminary validation results against rapid delivery NDACC data is now available in the TROPOMI validation reports and indicate that the bias is reduced after the upgrade to v 02.02.00 in July 2021 (see validation reports from https://mpc-vdaf.tropomi.eu/). This is the version we are using in our assimilation experiment. We have added a corresponding sentence in Section 2.2:
"The processor update to version 02.02.00 on in July 2021 included a change in spectroscopic parameters and preliminary results using rapid delivery NDACC data indicate that the bias is reduced to 2.9 % (Lambert et al., 2022)."

"It is not clear yet if this is the direct result of the assimilation of TROPOMI data leading to increased CO in the upper troposphere, or it could be the result of convective transport. " (page 18) I guess this could be easily studied by looking at the vertical profile of the analysis increments.

That's a good idea. We have plotted cross sections of analysis increments and can confirm that this increase is the result of the TROPOMI assimilation. We have replaced the sentence in the text with:
"Cross sections and maps of analysis increments (not shown) illustrate that the increased CO in the upper troposphere is a direct result of the TROPOMI assimilation, rather than a result of convective transport redistributing CO in the vertical, with the largest impact over Africa and South America."

"illustrating that the assimilation of TROPOMI CO improves the fit of the CAMS analysis to the IASI-BC data globally." (page 20 top) A really nice result!

The results for West Africa in Fig 16 could be linked to panel b in Fig. 10, and provides evidence I guess that the increments around 200 hPa are a good result. Please add a comment in the text.
We have added:
"The reduced bias at the West African airports in the upper troposphere confirms that the increase in CO seen in in ASSIM in the upper tropical troposphere in Fig. 6b is a good result."

Figure 16 caption: IASI should be IAGOS
Corrected

What is the relation between "Marco Polo" and China surface observations in Fig 18? Should this be CNEMC?
This was a mistake. Also, we have removed this figure from the paper and will do some more assessment on the quality and representativity of the observations. The figure is not needed for this paper as we have plenty of other independent observations.

The results suggest that the low biases in CAMS are linked to a low bias in emissions. Are there future approaches developed to improve this situation, e.g. a bias correction applied to the emissions or emission inversions? Please comment at the end of the conclusions section.

We have added in the conclusion:
"One shortcoming of the current CAMS system is a low CO bias in the free troposphere. This is a common problem with many atmospheric chemistry models (e.g. Gaubert elt al., 2020) and not just the IFS. Further work is needed to understand the reason for this bias and to improve the model. While the negative bias increased with the change to model version CY46R1 and the change to the CAMS_GLOB emission inventory in July 2019, negative biases were seen before then and might not purely be a result of underestimated anthropogenic emissions. Factors such as the distribution of the hydroxyl radical OH, secondary CO sources from the oxidation of methane and NMVOSs, and deposition processes might also be important. CAMS is in the process of developing an inversion prototype using ECMWF's 4D-Var system (McNorton et al., 2022), which should help to address shortcomings in emission inventories, but work on modelling aspects is also needed."